



Statistics for Industry and Technology

Series Editor

N. Balakrishnan
McMaster University
Department of Mathematics and Statistics
1280 Main Street West
Hamilton, Ontario L8S 4K1
Canada

Editorial Advisory Board

Max Engelhardt
EG&G Idaho, Inc.
Idaho Falls, ID 83415

Harry F. Martz
Group A-1 MS F600
Los Alamos National Laboratory
Los Alamos, NM 87545

Gary C. McDonald
NAO Research & Development Center
30500 Mound Road
Box 9055
Warren, MI 48090-9055

Kazuyuki Suzuki
Communication & Systems Engineering Department
University of Electro Communications
1-5-1 Chofugaoka
Chofu-shi
Tokyo 182
Japan

Advances in Degradation Modeling

Applications to Reliability, Survival Analysis, and Finance

M.S. Nikulin
Nikolaos Limnios
N. Balakrishnan
Waltraud Kahle
Catherine Huber-Carol
Editors

Birkhäuser
Boston • Basel • Berlin

Editors

M.S. Nikulin
l'Institut de Mathématiques de Bordeaux, IMB
Université Victor Segalen Bordeaux 2
33076 Bordeaux Cedex
France
nikou@sm.u-bordeaux2.fr

Nikolaos Limnios
Laboratoire de Mathématiques Appliquées
Université de Technologie de Compiègne
60205 Compiègne Cedex
France
nikolaos.limnios@utc.fr

N. Balakrishnan
Department of Mathematics and Statistics
McMaster University
1280 Main Street West
Hamilton, Ontario L8S 4K1
Canada
bala@univmail.cis.mcmaster.ca

Waltraud Kahle
Institut für Mathematische Stochastik
Otto-von-Guericke-Universität
D-39016 Magdeburg
Germany
waltraud.kahle@ovgu.de

Catherine Huber-Carol
Laboratoire de Statistique Théorique et
Appliquée
Université Paris Descartes et U780 INSERM
45 rue des Saints-Pères
75270 Paris Cedex 06
France
catherine.huber@parisdescartes.fr

ISBN 978-0-8176-4923-4 e-ISBN 978-0-8176-4924-1
DOI 10.1007/978-0-8176-4924-1

Library of Congress Control Number: 2009939145

Mathematics Subject Classification (2000): 60K10, 62Mxx, 90B25

© Birkhäuser Boston, a part of Springer Science+Business Media, LLC 2010

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Birkhäuser Boston, c/o Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Birkhäuser Boston is part of Springer Science+Business Media (www.birkhauser.com)



William Q. Meeker in 2009

Contents

Preface	XVII
William Q. Meeker – Career and Accomplishments	XIX
List of Contributors	XXIX
List of Tables	XXXIII
List of Figures	XXXV

Part I Review, Tutorials, and Perspective

1 Trends in the Statistical Assessment of Reliability	
<i>William Q. Meeker</i>	3
1.1 Background and Purpose	3
1.2 Traditional Reliability Data and Data Analysis	4
1.3 Product Design and Reliability Budgeting	5
1.4 Accelerated Testing	6
1.5 Multiple Failure Modes	6
1.6 Field and Warranty Data	8
1.7 Degradation Reliability Data	8
1.8 Recurrence Data	9
1.9 The Next Generation of Reliability Data	10
1.10 Software for Statistical Analysis of Reliability Data	11
1.11 Use of Bayesian Methods in Reliability	12
1.12 Concluding Remarks	14
References	14
2 Degradation Processes: An Overview	
<i>Mohamed Abdel-Hameed</i>	17
2.1 Introduction	17
2.2 Lévy and Pure Jump Processes	18
2.2.1 Lévy processes	18
2.2.2 Pure jump Markov processes	20
2.3 Life Distributions and Dependence Between Random Variables	20
2.3.1 Classes of life distributions	20
2.3.2 Dependence between random variables	21

2.4	The Degradation Process and Properties of the Resulting Life Distribution	21
2.4.1	Non-stationary gamma degradation process	21
2.4.2	Increasing Lévy and pure jump degradation processes	22
2.4.3	Brownian motion like degradation processes	22
2.5	Maintenance Policies of Devices Subject to Degradation	23
	References	24

3 Defect Initiation, Growth, and Failure – A General Statistical Model and Data Analyses

	<i>Wayne B. Nelson</i>	27
3.1	Introduction	27
3.2	An Application	28
3.3	A Population Model for Defect Initiation and Growth	30
3.4	Model Fitting by Computer	34
	References	37

4 Properties of Lifetime Estimators Based on Warranty Data Consisting only of Failures

	<i>Kazuyuki Suzuki, Watalu Yamamoto, Takashi Hara, and Md. Mesbahul Alam</i>	39
4.1	Introduction	39
4.2	A Model for Failure Data in Warranty Database	41
4.3	Maximum Likelihood Estimation	42
4.4	Properties of Maximum Likelihood Estimators	44
4.4.1	Fisher information matrix	44
4.4.2	One failure mode case	44
4.4.3	Two failure modes case	44
4.5	Effect of Sales Amount on Precision	46
4.6	Conclusion	48
A	Gradients and Hessians of Log-Likelihood (4.3)	48
A.1	Gradients	49
A.2	Hessians	49
B	Expectation of the Elements of Hessians	51
C	Gauss–Hermite Quadrature	54
	References	55

Part II Shock Models

5 Shock Models

	<i>Allan Gut and Jürg Hüsler</i>	59
5.1	Introduction	59
5.2	Cumulative Shock Models	60
5.3	Extreme Shock Models	60
5.4	Mixed Shock Models	61
5.4.1	An auxiliary result	62
5.4.2	Some examples	63
5.5	More Realistic Model	63
5.5.1	Delayed sums	64

5.5.2	A generalized cumulative shock model	65
5.5.3	A generalized extreme shock model	66
5.5.4	A generalized mixed shock model	68
5.5.5	Extreme shock models with varying threshold	69
5.5.6	The exact distribution of ν	70
5.5.7	The asymptotic distribution	71
5.6	An Extension to Markovian Shock Models	73
5.6.1	Cumulative shock models	74
5.6.2	Extreme shock models	74
5.6.3	Mixed shock models	75
	References	75

6 Parametric Shock Models

	<i>Waltraud Kahle and Heide Wendt</i>	77
6.1	Introduction	77
6.2	Modeling Degradation by Marked Point Processes	78
6.3	Characteristics of the Degradation Process	82
6.3.1	The counting process $(\Phi(t))$	82
6.3.2	The sequence (X_n) and the degradation process (Z_t)	83
6.3.3	The cumulative degradation at time t	84
6.3.4	The first passage time Z^h	88
6.4	Maximum Likelihood Estimations	90
6.5	The Large Sample Case	97
6.6	Moment Estimates	99
6.7	Comparison of Maximum Likelihood and Moment Estimates	101
6.8	Conclusion	103
	References	103

7 Poisson Approximation of Processes with Locally Independent Increments and Semi-Markov Switching – Toward Application in Reliability

	<i>V.S. Koroliuk, N. Limnios, and I.V. Samoilenko</i>	105
7.1	Introduction	105
7.2	Main Results	107
7.3	Proof of Theorem 1	110
	References	115

8 On Some Shock Models of Degradation

	<i>Maxim Finkelstein and Ji Hwan Cha</i>	117
8.1	Introduction	117
8.2	Asymptotic Properties for Two Special Shock Models	119
8.3	Terminating and “Accumulating” Shocks	121
8.4	Concluding Remarks	124
	References	124

Part III Degradation Models

9 The Wiener Process as a Degradation Model: Modeling and Parameter Estimation

<i>Waltraud Kahle and Axel Lehmann</i>	127
9.1 Introduction.....	127
9.2 Increments of the Degradation Process Are Observed	129
9.3 Observation of Failure Times	134
9.4 Observation of Both Degradation Increments and Failure Times	135
9.5 An Example.....	141
9.6 Simulation Study	144
References	145

10 On the General Degradation Path Model: Review and Simulation

<i>Firoozeh Haghighi, Nazanin Noorae, and Narges Nazeri Rad</i>	147
10.1 Introduction.....	147
10.2 Degradation Model.....	148
10.2.1 Degradation model with noise	149
10.2.2 Degradation model without measurement error	150
10.3 Simulation	151
References	154

11 A Closer Look at Degradation Models: Classical and Bayesian Approaches

<i>Marta A. Freitas, Thiago R. dos Santos, Magda C. Pires, and Enrico A. Colosimo</i>	157
11.1 Introduction.....	157
11.1.1 Background and literature	157
11.1.2 The problem	159
11.2 Train Wheel Degradation Data	161
11.3 Statistical Methods for Degradation Data Analysis	162
11.3.1 Methods based on “classical” inference.....	162
11.3.2 Bayesian inference	166
11.4 The Wheel Degradation Data Revisited	167
11.4.1 Estimation of $F_T(t)$ using the approximate method.....	168
11.4.2 Estimation of $F_T(t)$ using the analytical method	170
11.4.3 Estimation of $F_T(t)$ using the two-stage method	171
11.4.4 Estimation of $F_T(t)$ using the numerical method	172
11.4.5 Comparison of the results generated by the methods based on “classical” inference	172
11.4.6 Bayesian inference	172
11.5 Conclusions	177
References	178

12 Optimal Prophylaxis Policy Under Non-monotone Degradation

<i>S.S. Rasova and B.P. Harlamov</i>	181
12.1 Setting of the Problem	181
12.2 Optimization Problem	183
12.3 Trajectories of Choice	184
12.4 Degradation Process of Diffusion Type	192
References	194

13 Deterioration Processes with Increasing Thresholds

<i>S. Zacks</i>	195
13.1 Introduction	195
13.2 Preliminaries	196
13.3 The Reliability and Hazard in Case I	197
13.4 The Reliability and Hazard in Case II	199
13.5 Reliability and Hazard in Case III	200
13.6 Reliability and Hazard in Case IV	201
13.7 Exponential Deterioration	202
13.7.1 Case I with $K = 1$	202
13.7.2 Case I with $K = 2$	203
13.7.3 Case II with $K = 1$	204
13.7.4 Case III, $K = 1$	206
13.7.5 Case IV: Linear threshold	207
References	208

14 Failure Time Models Based on Degradation Processes

<i>Axel Lehmann</i>	209
14.1 Introduction	209
14.1.1 Models for degradation data	210
14.1.2 Models relating degradation and failure	213
14.2 Degradation–Threshold–Shock Models	214
14.2.1 The general model	214
14.2.2 Degradation–threshold models	216
14.2.3 Degradation–shock models	217
14.2.4 Likelihood function	218
14.2.5 Estimation of the survival function and the failure rate of T ..	222
14.3 DTS Models with Covariates	222
14.3.1 Maximum likelihood estimation	225
14.3.2 Semiparametric estimation	227
14.4 A DTS Model for Repairable Items	229
14.5 Application of DTS Models	230
References	231

15 Degradation and Fuzzy Information

<i>R. Viertl</i>	235
15.1 Introduction	235
15.2 Material Degradation	236
15.3 Fuzzy Initial Conditions	237
15.4 Fuzzy Distribution of $q(0)$	238

15.5 Application to Accelerated Life Testing	240
References	240

16 A New Perspective on Damage Accumulation, Marker Processes, and Weibull's Distribution

<i>Nozer D. Singpurwalla</i>	241
16.1 The Hazard Potential	242
16.2 A Stochastic Process Model for Damage and Its Markers	243
16.3 Introduction	244
16.4 The Weibull Distribution in Material Failure: Some History	244
16.5 Preliminaries and Notation	244
16.6 The Weakest Link Principle and an Application	245
16.7 Weibull's Approximation and Analysis	246
16.8 Critique of Weibull's Analysis	247
16.9 The Theory of Extreme Values	248
16.10 Comments on Using the Weibull Distribution Motivated via WLP and EVT	248
References	249

Part IV Reliability Estimation and ALT

17 Reliability Estimation of Mechanical Components Using Accelerated Life Testing Models

<i>Fabrice Guérin, M. Barreau, A. Charki, A. Todoskoff, S. Cloupet and D. Bigaud</i>	253
17.1 Introduction	253
17.2 Accelerated Life Testing Model	254
17.3 Regression Test Plan	256
17.3.1 Introduction	256
17.3.2 Parametric ALT model	257
17.3.3 Generalized proportional hazard (GPH) model	257
17.3.4 Semi-parametric ALT model	259
17.3.5 Application to ball bearings	260
17.4 Reliability Test With Previous Accelerated Damage	266
17.4.1 Principle	266
17.4.2 Test plan definition	267
17.4.3 Parametric model	268
17.4.4 Nonparametric model	269
17.4.5 Simulation example	270
17.5 Conclusions	272
References	273

18 Reliability Estimation from Failure-Degradation Data with Covariates

<i>V. Bagdonavičius, I. Masiulaityté, M.S. Nikulin</i>	275
18.1 Introduction	275
18.2 Modelling Simultaneous Traumatic Events and Degradation Data Under Covariates	276

18.3	Estimation of Model Parameters	280
18.3.1	The data	280
18.3.2	Likelihood function construction	280
18.3.3	Example 1: Time-scaled gamma process	281
18.3.4	Example 2: Shock processes	284
18.3.5	Example 3: Path models	286
18.3.6	Modified loglikelihood	288
18.4	Estimation of Reliability Characteristics	288
	References	289

19 Asymptotic Properties of Redundant Systems Reliability Estimators

	<i>V. Bagdonavičius, I. Masiulaitytė, M.S. Nikulin</i>	293
19.1	Introduction	293
19.2	Point Estimators of the c.d.f. of Redundant Systems	294
19.2.1	Nonparametric estimation	294
19.2.2	Parametric estimation	296
19.3	Asymptotic Distribution of \hat{K}_j and Confidence Intervals for $K_j(t)$	296
19.3.1	Nonparametric case	297
19.3.2	Parametric case	302
19.4	Power of Goodness-of-Fit Tests	306
	References	310

20 An Approach to System Reliability Demonstration Based on Accelerated Test Results on Components

	<i>Léo Gerville-Réache and Vincent Couallier</i>	311
20.1	Introduction	311
20.2	Global Reliability Demonstration from k “Zero-Failure” Component Testing Procedures	313
20.2.1	Equal component test times	314
20.2.2	Equal reliability targets	314
20.2.3	Integrating failed demonstration procedures	315
20.3	Designing a Global Demonstration Test for the Reliability of a Series System Under a Success Probability Constraint	316
20.3.1	Basic principle of the demonstration test planning	316
20.3.2	“ n/j failures” demonstration test	318
20.4	Conclusion	319
	References	319

Part V Survival Function Estimation

21 Robust Versus Nonparametric Approaches and Survival Data Analysis

	<i>Catherine Huber</i>	323
21.1	Introduction	324
21.2	Motivation for Robustness	324
21.2.1	Instability of usual tests and estimators	324
21.2.2	Sensitivity of rank test	327

21.3	Robustness Concepts	328
21.3.1	Robust versus nonparametric approach	328
21.3.2	Regularity of the parametric model	329
21.3.3	Extension of the underlying model	330
21.3.4	Measures of robustness	330
21.4	Robustness in Survival Analysis	331
21.4.1	Measure of robustness of Kaplan–Meier estimator	331
21.4.2	Perspectives	336
	References	336

22 Modelling Recurrent Events for Repairable Systems Under Worse Than Old Assumption

	<i>G. Babykina and V. Couallier</i>	339
22.1	Introduction	339
22.2	A New Model of Imperfect Repair: The LEYP Model	342
22.2.1	Some useful properties	344
22.3	Taking Covariates into Account	345
22.4	Statistical Estimation and Data Description	345
22.5	Numerical Example	347
22.5.1	Data description	347
22.5.2	Parameter estimation	350
22.5.3	Predictions	352
22.6	Conclusion	353
	References	353

23 Survival Models for Step-Stress Experiments With Lagged Effects

	<i>N. Kannan, D. Kundu, and N. Balakrishnan</i>	355
23.1	Introduction	355
23.2	Model Description	357
23.2.1	Step-stress models with latency	357
23.3	Maximum Likelihood Estimators for the CRM	358
23.4	Least Squares Estimators	360
23.5	Data Analysis	361
23.6	Simulation Results	362
23.7	Conclusions	367
	References	368

24 Estimation of Density on Censored Data

	<i>V. Solev</i>	371
24.1	Introduction	371
24.2	Approximating of Parametric Set	374
24.3	Hellinger Distance	376
24.4	Main Result	378
	References	379

Part VI Competing Risk and Chaotic Systems

25 Toward a Test for Departure of a Trajectory from a Neighborhood of a Chaotic System

<i>M. LuValle</i>	383
25.1 Introduction	383
25.1.1 Terminology and the Lorenz attractor	385
25.1.2 The alternative for the simulation	386
25.2 The Test Statistic and Supporting Theory	387
25.3 Computer Experiments	391
25.3.1 Description of the computer experiments	392
25.4 Directions for Future Work	394
References	394

26 Probability Plotting with Independent Competing Risks

<i>Francis G. Pascual and Christopher Gast</i>	397
26.1 Introduction	397
26.1.1 Competing risks	397
26.1.2 Probability plotting	398
26.1.3 Outline	399
26.2 Notation and Model Assumptions	399
26.2.1 Distributions of individual risks	399
26.2.2 Distribution of subject lifetime	400
26.2.3 The likelihood function	400
26.3 The Kaplan–Meier Estimator And Probability Plotting	401
26.3.1 Kaplan–Meier estimator	401
26.3.2 Linearizing the Cdf under one risk	402
26.3.3 Probability plotting and competing risks	402
26.4 Proposed Method	403
26.5 Applications	404
26.5.1 Breast cancer study	404
26.5.2 Shock absorber failure data	407
26.5.3 Simulated data set	410
26.6 Conclusions	412
References	413

Index	415
------------------------	-----

Preface

William Q. Meeker has made pioneering and phenomenal contributions to the general area of reliability and, in particular, to the topics of degradation and accelerated testing. His research publications and the numerous citations he has received over the past three decades provide an ample testimony to this fact.

Statistical methods have become critical in analyzing reliability and survival data. Highly reliable products have necessitated the development of accelerated testing and degradation models and their analyses. This volume has been put together in order to (i) review some of the recent advances on accelerated testing and degradation, (ii) highlight some new results and discuss their applications, and (iii) suggest possible directions for future research in these topics.

With these specific goals in mind, many authors were invited to write a chapter for this volume. These authors are not only experts in lifetime data analysis, but also form a representative group from former students, colleagues, and other close professional associates of William Meeker. All contributions have been peer reviewed and organized into 26 chapters. For the convenience of readers, the volume has been divided into the following six parts:

- Review, Tutorials, and Perspective
- Shock Models
- Degradation Models
- Reliability Estimation and ALT
- Survival Function Estimation
- Competing Risk and Chaotic Systems

It needs to be emphasized here that this volume is not a proceedings, but a carefully and deliberately planned volume comprising chapters consistent with the editorial goals and purposes mentioned above.

Our thanks go to all the authors who have contributed to this volume. Thanks are also due to Mrs. Debbie Iscoe for the excellent typesetting of the entire volume. Special thanks go to Ms. Regina Gorenshteyn and Mr. Tom Grasso (Editor, Birkhäuser, Boston) for their interest and support for this project.

The volume was difficult, it is clear, but Leah (Project Manager at Integra Software Services), Brian, Tom, and Dubby helped us very much to prepare this nice volume!

XVIII Preface

Like us, all the authors who contributed to this volume have great admiration for the work and accomplishments of William Meeker and therefore provided us with hearty cooperation during the preparation of this volume. It is a great pleasure and honor for all of us to dedicate this volume to William Meeker.

June 2009

M.S. Nikulin
Nikolaos Linnios
N. Balakrishnan
Waltraud Kahle
Catherine Huber-Carol

William Q. Meeker – Career and Accomplishments

William Meeker received his B.S. degree in industrial management from Clarkson College of Technology in 1972 and M.S. degree in operations research and Ph.D. degree in administrative and engineering systems, both from Union College (Schenectady, New York) in 1973 and 1975, respectively. Soon after getting his Ph.D., he joined as an assistant professor in the Department of Statistics at Iowa State University, Ames. After getting promoted to the ranks of associate professor and professor in 1978 and 1981, respectively, he was appointed there as a distinguished professor of liberal arts and sciences in 1996, and he has been there in this position since. During this period, he also held visiting positions at Global Research Center of General Electric Company (Schenectady, New York), Quality Theory and Systems Department of Bell Laboratories (Holmdel, New Jersey), Louisiana State University (Baton Rouge, Louisiana), and University of Waterloo (Waterloo, Ontario, Canada). He is also a faculty affiliate at Los Alamos National Laboratory (Nevada) since 1999.

At Iowa State University, William Meeker has made invaluable contributions. Since 1989, he has been a principal investigator at the Center for Nondestructive Evaluation. He excelled in teaching a wide array of courses and in fact received the Iowa State University Teaching Excellence Award in 1989 and 1991 for his efforts. He has guided 71 M.S. projects and 11 Ph.D. dissertations and is supervising six graduate students at present.

William Meeker has provided distinguished service to the statistical community at large by his activities in various capacities for professional societies. These include secretary-treasurer of ASA Business and Economics Section (1981–1982), member of Advisory Board of ASA Section on Physical and Engineering Sciences (1984–1986), member of ASA Committee on Publications (1987–1989), member of ASA Ad Hoc Committee for *Journal of Computational and Graphical Statistics* (1987–1990), president of the Iowa Chapter of ASA (1989–1991), chair of the IMS Committee on Statistical Tables (1990–1994), member of ASQC Publications Management Board (1991–1993), chair of *Technometrics* Management Committee (1991–1993), COPSS visiting lecturer (1991–1995), member of ASA Journals Management Committee (1992–1993), representative of Iowa Chapter to ASA Council of Chapters (1995–1997), member of the ISI Committee for Statistics in Business and Industry (1997–2005), member of NSF SCREMS Proposal Review Panel (1998, 2005), program chair of Spring Research Conference (1999), member of ASA Fellows Committee (2001–2003), member of NSF Statistics Research Proposal Review Panel (2001), member of NSF CAREER Proposal

Review Panel (2002), member of NRC Panel on Operational Test Design and Evaluation of the Interim Armored Vehicle (2001–2003), chair of ASA Task Force on Journals Marketing (2003), member of ASA Task Force on Electronic Publications (2002–2004), member of ASA Publications Committee (1998–2002), chair of ASA Publications Committee (2003–2006), and member of NSF Research Experiences for Undergraduates Proposal Review Panel (2006).

In addition, William Meeker has provided tremendous service to many research journals in various capacities. Included in this list are associate editor of *Technometrics* (1979–1986), editor of *Technometrics* (1987–1989), editorial board member of *Selected Tables in Mathematical Statistics* (1981–1989), co-editor of *Selected Tables in Mathematical Statistics* (1990–1994), editorial board member of *International Statistical Review* (1995–1999), editorial board member of *Lifetime Data Analysis* (2001–2009), advisory editor of *Quality Technology & Quality Management* (2003–2009), and guest editor of *Journal of Statistical Planning and Inference* Special Issue on Accelerated Testing (JSPI, 2009).

William Meeker has received numerous distinctions and awards throughout his career. He has been elected a fellow of the American Statistical Association and the American Society for Quality, and an elected member of the International Statistical Institute. Some other notable awards include ASA Outstanding Statistical Application (2001), Frank Wilcoxon Prize for the best practical application paper in *Technometrics* (1987, 1995, 1999), W.J. Youden Prize for the best expository paper in *Technometrics* (1996, 1998, 1999, 2002), William G. Hunter Award from the Statistics Division of the ASQ (2003), and ASQ Shewhart Medal (2006). Moreover, his book *Statistical Methods for Reliability* (co-authored with L. Escobar), published in 1998 by John Wiley & Sons, received the Professional/Scholarly Publishing Division of the Association of American Publishers Award for Excellence and Innovation in Engineering.

William Meeker, through his pioneering and phenomenal research in the area of reliability over the last 35 years, has influenced deeply the trend of research in this area and has provided guidance, inspiration, and encouragement to numerous young researchers. For his exemplary career and immense contributions, he was chosen “Statistician of the Year” in 2006 by the Chicago Chapter of the ASA. It is our sincere hope and wish that he will continue his contributions to the area of reliability and the statistical profession in general with renewed vigor and energy.

This volume includes a number of chapters on the topics of degradation and accelerated testing written by experts who know and appreciate William Meeker and all his contributions!

Books and Booklets

1. Meeker, W.Q., Cornwell, L., and Aroian, L.A. (1981), The Product of Two Normally Distributed Random Variables. Volume 7 of *Selected Tables in Mathematical Statistics*. Providence, Rhode Island: American Mathematical Society.
2. Meeker, W.Q., and Hahn, G.J. (1985), How To Plan An Accelerated Life Tests—Some Practical Guidelines. Volume 10 in the *American Society for Quality Control Basic References in Quality Control: Statistical Techniques*. Milwaukee, Wisconsin: American Society for Quality Control.

3. Hahn, G.J., and Meeker, W.Q. (1991), *Statistical Intervals: A Guide for Practitioners*. New York: John Wiley and Sons.
4. Meeker, W.Q., and Escobar, L.A. (1998), *Statistical Methods for Reliability Data*. New York: John Wiley and Sons.

Book Chapters

1. Meeker, W.Q. (1979), Nites Rest Inc.—A Box-Jenkins Time Series Analysis Case Study. Chapter 12 in *Forecasting, Time Series, and Regression: An Applied Approach*, edited by Bruce L. Bowerman and Richard T. O'Connell, Duxbury Press, North Scituate, MA.
2. Meeker, W.Q., and Escobar, L.A. (1994), Maximum Likelihood Methods for Fitting Parametric Statistical Models to Censored and Truncated Data. Chapter 8 in *Probabilistic and Statistical Methods in the Physical Sciences*, edited by John Stanford and Stephen Vardeman, New York: Academic Press.
3. Meeker, W.Q., and Escobar, L.A. (1999), Accelerated Life Tests: Concepts and Data Analysis. Chapter 10 in *A Systems Approach to Service Life Prediction of Organic Coatings*, edited by D.R. Bauer and J.W. Martin, Washington: American Chemical Society.
4. Meeker, W.Q., Escobar, L.A., Doganaksoy, N., and Hahn, G.J. (1999), Reliability Concepts and Data Analysis. Section 48 in the *Juran's Handbook on Quality*, 5th Edition, edited by J. M. Juran and A. B. Godfrey, New York: McGraw Hill.
5. Meeker, W.Q., Escobar, L.A., and Chan, V. (2002), Using Accelerated Tests to Predict Service Life in Highly-Variable Environments. Chapter 19 in *Service Life Prediction Methodology and Metrologies*, edited by J. W. Martin and D.R. Bauer, Washington: American Chemical Society.
6. Meeker, W.Q., Escobar, L.A., and Zayac, S.A. (2003), Use of Sensitivity Analysis to Assess the Effect of Model Uncertainty in Analyzing Accelerated Life Test Data. Chapter 6 in *Case Studies in Reliability and Maintenance*, edited by W.R. Blischke and D.N.P. Murthy, New York: John Wiley & Sons.
7. Meeker, W.Q., and Escobar, L.A. (2003), Use of Truncated Regression Methods to Estimate the Shelf Life of a Product from Incomplete Historical Data. Chapter 12 in *Case Studies in Reliability and Maintenance*, edited by W.R. Blischke and D.N.P., Murthy, New York: John Wiley & Sons.
8. Doganaksoy, N., Hahn, G.J., and Meeker, W.Q. (2005), Assuring Product Reliability and Safety, a chapter in *Statistics - A Guide to the Unknown*, 4th Edition, edited by Roxy Peck and others, Brooks-Cole and the American Statistical Association.
9. Pascual, F.G., Meeker, W.Q., and Escobar, L.A. (2006), Accelerated Life Test Models and Data Analysis. Chapter 19 in *Handbook for Engineering Statistics*, New York: Springer.
10. Gu, X., Stanley, D., Byrd, W., Dickens, B., Vaca-Trigo, I., Meeker, W. Q., Nguyen, Chin, J., and Martin, J. (2009), Linking Accelerated Laboratory Test with Outdoor Performance Results for a Model Epoxy Coating System. In *Service Life Prediction of Polymeric Materials*, edited by Jonathan Martin, Rose A. Ryntz, Joannie Chin, Ray A. Dickie, New York: Springer.
11. Vaca-Trigo, I., and Meeker, W.Q. (2009), A Statistical Model for Linking Field and Laboratory Exposure Results for a Model Coating. In *Service Life Prediction of*

Polymeric Materials, edited by Jonathan Martin, Rose A. Ryntz, Joannie Chin, Ray A. Dickie, New York: Springer.

Refereed Publications

1. Meeker, W., Hahn, G., and Feder, P. (1975), A Computer Program for Evaluating and Comparing Experimental Designs and Some Applications. *The American Statistician*, **29**, No. 1, 60–64.
2. Meeker, W.Q., and Nelson, W. (1975), Optimum Accelerated Life Tests for Weibull and Extreme Value Distributions. *IEEE Transactions on Reliability*, **24**, No. 5, 321–332.
3. Meeker, W.Q., and Nelson, W. (1976), Weibull Percentile Estimates and Confidence Limits from Singly Censored Data by Maximum Likelihood. *IEEE Transactions on Reliability*, **25**, No. 11, 20–24.
4. Hahn, G., Meeker, W.Q., and Feder, P. (1976), The Evaluation and Comparison of Experimental Designs for Fitting Regression Relationships. *Journal of Quality Technology*, **8**, No. 3, 140–157.
5. Kamen, A., Schmee, J., and Meeker, W.Q. (1976), Propriety of Using Percentages in Reporting Anticariogenic Studies. *Journal of Dental Research*, **55**, No. 4, 703.
6. Meeker, W.Q., and Nelson, W. (1976), Tables for the Weibull and Extreme Value Distributions. *The Relia-Com Review*, **1**, No. 3, 1–5.
7. Meeker, W.Q., Hahn, G.J., and Feder, P.I. (1977), New Bias Evaluation Features of EXPLOR-A Program for Assessing Experimental Design Properties. *The American Statistician*, **31**, No. 2, 95–96.
8. Meeker, W.Q., and Hahn, G.J. (1977), Asymptotically Optimum Over-Stress Tests to Estimate the Survival Probability at a Condition with a Low Expected Failure Probability (with discussion). *Technometrics*, **19**, No. 4, 381–399.
9. Meeker, W.Q., and Nelson, W. (1977), Confidence Limits for the Weibull Distribution from Censored Data. *Technometrics*, **19**, No. 4, 473–476.
10. Hahn, G.J., Feder, P.I., and Meeker, W.Q. (1978), Evaluating the Effect of Incorrect Specification of a Regression Model, Part I: Basic Concepts and Example. *Journal of Quality Technology*, **10**, No. 2, 61–72.
11. Meeker, W.Q. (1978), Sequential Tests of Independence for 2x2 Contingency Tables. *Biometrika*, **65**, No. 1, 85–90.
12. Nelson, W., and Meeker, W.Q. (1978), Theory for Optimum Accelerated Life Tests for the Weibull and Extreme Value Distributions. *Technometrics*, **20**, No. 2, 171–177.
13. Hahn, G.J., Feder, P.I., and Meeker, W.Q. (1978), Evaluating the Effect of Incorrect Specification of a Regression Model, Part II: Further Example and Discussion. *Journal of Quality Technology*, **10**, No. 3, 93–98.
14. Meeker, W.Q. (1978), TSERIES-A User-Oriented Computer Program for Time Series Analysis. *American Statistician*, **32**, No. 3, 111–112.
15. Meeker, W.Q., and Hahn, G.J. (1978), A Comparison of Accelerated Test Plans to Estimate the Survival Probability at a Design Stress. *Technometrics*, **10**, No. 3, 245–247.
16. Meeker, W.Q., and Hahn, G.J. (1980), Prediction Intervals for the Ratios of Normal Distribution Sample Variances and Exponential Distribution Sample Means. *Technometrics*, **22**, No. 3, 357–366.

17. Meeker, W.Q. (1981), A Conditional Sequential Test for the Equality of Two Binomial Proportions. *Applied Statistics*, **30**, No. 2, 109–115.
18. Meeker, W.Q., and Duke, S.D. (1981), CENSOR—A User-Oriented Computer Program for Life Data Analysis. *The American Statistician*, **35**, No. 2, 112.
19. Hahn, G.J., and Meeker, W.Q. (1982), Pitfalls and Practical Considerations in Product Life Analysis, Part 1: Basic Concepts and Dangers of Extrapolation. *Journal of Quality Technology*, **14**, No. 3, 144–152.
20. Hahn, G.J., and Meeker, W.Q. (1982), Pitfalls and Practical Considerations in Product Life Analysis, Part 2: Mixtures of Product Populations and More General Models. *Journal of Quality Technology*, **14**, No. 4, 177–185.
21. Meeker, W.Q., and Hahn, G.J. (1982), Sample Sizes for Prediction Intervals. *Journal of Quality Technology*, **14**, No. 4, 201–206.
22. Meeker, W.Q. (1984), A Comparison of Accelerated Life Test Plans for Weibull and Lognormal Distributions and Type I Censored Data. *Technometrics*, **26**, 157–171.
23. Hahn, G.J., and Meeker, W.Q. (1984), An Engineer's Guide to Books on Statistics and Data Analysis. *Journal of Quality Technology*, **16**, No. 3, 196–218.
24. Escobar, L.A., and Meeker, W.Q. (1986), Optimum Accelerated Life Tests with Type II Censored Data. *Journal of Statistical Computation and Simulation*, **23**, 273–297.
25. Escobar, L.A., and Meeker, W.Q. (1986), Elements of the Fisher Information Matrix for the Smallest Extreme Value Distribution and Censored Data. *Applied Statistics*, **35**, 80–86.
26. Meeker, W.Q. (1986), Planning Life Tests in which Units are Periodically Inspected for Failure. *IEEE Transactions on Reliability*, **35**, 571–578.
27. Meeker, W.Q. (1987) Limited Failure Population Life Tests: Application to Integrated Circuit Reliability. *Technometrics*, **29**, 151–165.
28. Ostrouchov, G., and Meeker, W.Q. (1988), Accuracy of Approximate Confidence Bounds from Interval Censored Weibull and Lognormal Data. *Journal of Statistical Computation and Simulation*, **29**, 43–76.
29. Vander Weil S., and Meeker, W.Q. (1990), Accuracy of Approximate Confidence Bounds Using Censored Weibull Regression Data from Accelerated Life Tests. *IEEE Transactions on Reliability*, **39**, 346–351.
30. Weston, S.A., and Meeker, W.Q. (1991), Coverage Probabilities of Nonparametric Simultaneous Confidence Bands for a Survival Function. *Journal of Statistical Computation and Simulation*, **32**, 83–97.
31. Meeker, W.Q., Escobar, L.A., and Hill, D.A. (1992), Sample Sizes for Estimating the Weibull Distribution Hazard Function from Censored Samples. *IEEE Transactions on Reliability*, **41**, 133–138.
32. Escobar, L.A., and Meeker, W.Q. (1992), Assessing Local Influence in Regression Analysis with Censored Data. *Biometrics*, **48**, 507–528.
33. Kernan, W.J., and Meeker, W.Q. (1992), A Statistical Test to Assess Changes in Spontaneous Behavior of Rats Observed with a Computer Recognition System. *Journal of Biopharmaceutical Statistics*, **2**, 115–135.
34. Hahn, G.J., and Meeker, W.Q. (1993) The Assumptions of Statistical Inference. *The American Statistician*, **47**, 1–11.
35. Lu, C.J., and Meeker, W.Q. (1993), Using Degradation Measures to Estimate a Time-to-Failure Distribution. *Technometrics*, **35**, 161–174.
36. Meeker, W.Q., and Escobar, L.A. (1993), A Review of Recent Research and Current Issues in Accelerated Testing. *International Statistical Review*, **61**, 147–168.

37. Arnold, B.C., Beaver, R., Groeneveld, R.A., and Meeker, W.Q. (1993), The Non-truncated Marginal of a Truncated Bivariate Normal Distribution. *Psychometrika*, **58**, 471–488.
38. Meeker, W.Q., and Escobar, L.A. (1994), An Algorithm to Compute the cdf of the Product of Two Normal Random Variables. *Communications in Statistics*, **23**, 271–280.
39. Meeter, C.A., and Meeker, W.Q. (1994), Optimum Accelerated Life Tests with Nonconstant Scale Parameter. *Technometrics*, **36**, 71–83.
40. Escobar, L.A., and Meeker, W.Q. (1994), Fisher Information Matrix for the Extreme Value, Normal, and Logistic Distributions and Censored Data. *Applied Statistics*, **43**, 533–540.
41. Garrigoux, C.G., and Meeker, W.Q. (1994), A Reliability Model for Planning In-Service Inspections for Components Subject to Degradation Failure. *Pakistan Journal of Statistics*, **10**, 79–98.
42. Garrigoux, C.G., and Meeker, W.Q. (1995), Assessing the Effect of In-Service Inspections on the Reliability of Degrading Components. In: *Recent Advances in Life-Testing and Reliability*, N. Balakrishnan (Editor), CRC Press, Boca Raton.
43. Meeker, W.Q., and LuValle, M.J. (1995), An Accelerated Life Test Model Based on Reliability Kinetics. *Technometrics*, **37**, 133–146.
44. Meeker, W.Q., and Escobar, L.A. (1995), Teaching About Approximate Confidence Regions Based on Maximum Likelihood Estimation. *The American Statistician*, **49**, 48–53.
45. Meeker, W.Q., and Hamada, M. (1995), Statistical Tools for the Rapid Development & Evaluation of High-Reliability Products. *IEEE Transactions on Reliability*, **44**, 187–198.
46. Moore, D.S., Cobb, G.W., Garfield, J., and Meeker, W.Q. (1995), Statistics Education Fin de Siècle. *The American Statistician*, **49**, 250–260.
47. Escobar, L.A., and Meeker, W.Q. (1995), Planning Accelerated Life Tests with Two or More Factors. *Technometrics*, **37**, 411–427.
48. Olin, B.D., and Meeker, W.Q., (1996), Applications of Statistics in Nondestructive Evaluation (with discussion). *Technometrics*, **38**, 95–112.
49. Cannon, A.R., and Meeker, W.Q. (1996), Statistical Tests for Signals in Categorical Temporal Data. *Biometrical Journal*, **38**, 39–59.
50. Lu, C.J., Meeker, W.Q., and Escobar, L.A. (1996), A Comparison of Degradation and Failure-Time Analysis Methods of Estimating a Time-to-Failure Distribution. *Statistica Sinica*, **6**, 531–546.
51. Marasinghe, M., Meeker, W.Q., Cook, D., and Shin, T. (1996), Using Graphics and Simulation to Teach Statistical Concepts. *American Statistician*, **50**, 342–351.
52. Field, D., and Meeker, W.Q. (1996), Optimizing Product Design Based on Time to Failure Distributions. *Quality and Reliability Engineering International*, **12**, 429–438.
53. Pascual, F.G., and Meeker, W.Q. (1997), Regression Analysis of Fatigue Data with Runouts Based on a Model with Nonconstant Standard Deviation and a Fatigue Limit Parameter. *Journal of Testing and Evaluation*, **25**, 292–301.
54. Liu, S., Lu, J.C., Kolpin, D.W., and Meeker, W.Q. (1997), Analysis of Environmental Data with Censored Observations. *Environmental Science and Technology*, **31**, 3358–3362.

55. Escobar, L.A., and Meeker, W.Q. (1998), The Asymptotic Covariance Matrix for Maximum Likelihood Estimators with Models based on Location-Scale Distributions Involving Censoring, Truncation, and Explanatory Variables. *Statistica Sinica*, **8**, 221–237.
56. Sarkar, P., and Meeker, W.Q. (1998), A Bayesian On-Line Change Detection Algorithm with Process Monitoring Applications. *Quality Engineering*, **10**, 539–549.
57. Meeker, W.Q., Escobar, L.A. , and Lu, C.J. (1998), Accelerated Degradation Tests: Modeling and Analysis. *Technometrics*, **40**, 89–99.
58. Meeker, W.Q., and Escobar, L.A. (1998), Pitfalls of Accelerated Testing. *IEEE Transactions on Reliability*, **47**, 114–118.
59. Pascual, F.G., and Meeker, W.Q. (1998), Planning Life Tests with a Limited Number of Test Positions. *Journal of Testing and Evaluation*, **26**, 434–443.
60. Escobar, L.A., and Meeker, W.Q. (1999), Statistical Prediction Based on Censored Life Data. *Technometrics*, **41**, 113–124.
61. Hahn, G.J., Doganaksoy, N., and Meeker, W.Q. (1999), Reliability Improvement. *Quality Progress*, **32**, 133–139.
62. Pascual, F.G., and Meeker, W.Q. (1999), Estimating Fatigue Curves with the Random Fatigue-Limit Model (with discussion). *Technometrics*, **41**, 277–302.
63. Chan, V., and Meeker W.Q. (1999), A Failure-Time Model for Infant Mortality and Wearout Failure Modes. *IEEE Transactions on Reliability*, **48**, 678–682.
64. Doganaksoy, N., Hahn, G.J., and Meeker, W.Q. (2000), Product Life Analysis: A Case Study. *Quality Progress*, **33**, 115–122.
65. Jeng, S.L., and Meeker W.Q. (2000), Comparisons of Weibull Distribution Approximate Confidence Intervals Procedures for Type I Censored Data. *Technometrics*, **42**, 135–148.
66. Escobar, L.A., and Meeker, W.Q. (2001), A Note on the Asymptotic Equivalence of the Fisher Information Matrices for Type I and Type II Censored Data from Location–Scale Families. *Communications in Statistics*, **30**, 2211–2225.
67. Jeng, S.L., and Meeker W.Q. (2001), Parametric Simultaneous Confidence Bands for Cumulative Distributions from Censored Data. *Technometrics*, **43**, 450–461.
68. Meeker, W.Q., Doganaksoy, N., and Hahn, G.J. (2001), Using Degradation Data for Product Reliability Analysis. *Quality Progress*, **34**, 60–65.
69. Meeker, W.Q., and Escobar, L.A. (2002), Software for Reliability Data Analysis and Test Planning. *Brazilian Journal of Statistics*, **15**, 169–200.
70. Nordman, D., and Meeker, W.Q. (2002), Weibull Prediction Intervals for a Future Number of Failures. *Technometrics*, **44**, 15–23.
71. Doganaksoy, N., Hahn, G.J., and Meeker, W.Q. (2002), Reliability Analysis by Failure Mode. *Quality Progress*, **35**, 47–52.
72. Wu, H., and Meeker, W.Q. (2002), Early Detection of Reliability Problems Using Information from Warranty Data Bases. *Technometrics*, **44**, 120–133.
73. Doganaksoy, N., Hahn, G.J., and Meeker, W.Q. (2003), Speedier Reliability Analysis. *Quality Progress*, **36**, 58–64.
74. Escobar, L.A., Meeker, W.Q., Kugler, D.L., and Kramer, L.L. (2003), Accelerated Destructive Degradation Tests: Data, Models, and Analysis. In: *Mathematical and Statistical Methods in Reliability*, B.H. Lindqvist and K.A. Doksum (Editors.) World Scientific Publishing Company.
75. Meeker, W.Q., and Escobar, L.A. (2004), Reliability: The Other Dimension of Quality. *Quality Technology & Quality Management*, **1**, 1–25.

76. Meeker, W.Q., and Escobar, L.A. (2004), Discussion of “Failure Augmentation Method: An Information Maximization Approach to Categorical Response Optimization”. *Technometrics*, **46**, 15–16.
77. Chan, V., Lahiri, S.N., and Meeker, W.Q. (2004), Block Bootstrap Estimation of the Distribution of Cumulative Outdoor Degradation. *Technometrics*, **46**, 215–224.
78. Meeker, W.Q., Hahn, G.J., and Doganaksoy, N. (2004), Planning Life Tests for Reliability Demonstration. *Quality Progress*, **37**, 80–82.
79. Jeng, S.L., Lahiri, S.N., and Meeker W.Q. (2005), Asymptotic Properties of Bootstrap Likelihood Ratio Statistics for Time Censored Data. *Statistica Sinica*, **15**, 35–57.
80. McKane, S.W., Escobar, L.A., and Meeker, W.Q. (2005), Sample Size and Number of Failure Requirements for Demonstration Tests with Log-Location-Scale Distributions and Type II Censoring. *Technometrics*, **47**, 182–190.
81. Meeker, W.Q., Hahn, G.J., and Doganaksoy, N. (2005), Planning Reliability Assessment. *Quality Progress*, **38**, 90–93.
82. Zhang, Y., and Meeker, W.Q. (2005), Bayesian Life Test Planning for the Weibull Distribution with Given Shape Parameter. *Metrika*, **61**, 237–249.
83. Zhang, Y., and Meeker, W.Q. (2005), Bayesian Optimum Planning for Accelerated Life Tests. *Technometrics*, **48**, 49–60.
84. Doganaksoy, N., Hahn, G.J., and Meeker, W.Q. (2006), How to Analyze Reliability Data for Repairable Products. *Quality Progress*, **39**, 93–95.
85. Escobar, L.A., and Meeker, W.Q. (2006), A Review of Accelerated Test Models. *Statistical Science*, **21**, 552–577.
86. Chan, V., and Meeker, W.Q. (2007), Estimation of Degradation-Based Reliability in Outdoor Environments, *Communications in Statistics*, **37**, 408–424.
87. Doganaksoy, N., Hahn, G.J., and Meeker, W.Q. (2007), Reliability Assessment by Use-Rate Acceleration, *Quality Progress*, **39**, 74–76.
88. Hong, Y., Meeker, W.Q., and Escobar, L.A. (2008), Avoiding Problems With Normal Approximation Confidence Intervals for Probabilities. *Technometrics*, **50**, 64–68.
89. Zuo, J., Meeker, W.Q., and Wu, H. (2008), Analysis of Window-Observation Recurrence Data. *Technometrics*, **50**, 128–143.
90. Hong, Y., Meeker, W.Q., and Escobar, L.A. (2008), The Relationship Between Confidence Intervals for Failure Probabilities and Life Time Quantiles. *IEEE Transactions on Reliability*, **57**, 260–266.
91. Ma, H., and Meeker, W.Q. (2008), Optimum Step-Stress Accelerated Life Test Plans for Log-Location-Scale Distributions. *Naval Research Logistics*, **55**, 551–562.
92. Doganaksoy, N., Hahn, G.J., and Meeker, W.Q. (2008), The Pros of Proactive Product Servicing. *Quality Progress*, **40**, 60–62.
93. Shi, Y., Escobar, L.A., and Meeker, W.Q. (2009), Planning Accelerated Destructive Degradation Tests. *Technometrics*, **51**, 1–13.
94. Hong, Y., Meeker, W.Q., and McCalley, J.D. (2009), Prediction of Remaining Life of Power Transformers Based on Left Truncated and Right Censored Lifetime Data. *Annals of Applied Statistics*, **3**, No. 2, 857–879.
95. Meeker, W.Q., Escobar, L.A., and Hong, Y. (2009), Using Accelerated Life Tests Results to Predict Field Reliability. *Technometrics*, **51**, No. 2, 146–161.
96. Escobar, L.A., Hong, Y., and Meeker, W.Q. (2009), Simultaneous Confidence Bands and Regions for Log-Location-Scale Distributions with Censored Data. *Journal of Statistical Planning and Inference*, **139**, No. 9, 3231–3245.

97. JSPI (2009), Special Issue on Degradation, Damage, Fatigue and Accelerated Life Models in Reliability Testing. *Journal of Statistical Planning and Inference*, **139**, No. 5, 1575–1820. Edited by Luis A. Escobar, Fabrice Guerin, William Q. Meeker and Mikhail Nikulin.

List of Contributors

Mohamed Abdel-Hameed

Department of Statistics
College of Business and Economics
United Arab Emirates University
United Arab Emirates
mohamed@uaeu.ac.ae

G. Babykina

Cemagref
50 Avenue de Verdun
33620 Cestas Cedex, France
genia.babykina@cemagref.fr

V. Bagdonavičius

Vilnius University
Vilnius, Lithuania
Vilijandas.Bagdonavicius@maf.vu.lt

N. Balakrishnan

Department of Mathematics and
Statistics
McMaster University, Canada
bala@univmail.cis.mcmaster.ca

M. Barreau

Institut des Sciences et Techniques
de l'Ingénieur d'Angers
LASQUO EA3858
F-49000 Angers France
mihaela.barreau@univ-angers.fr

D. Bigaud

Institut des Sciences et Techniques
de l'Ingénieur d'Angers
France

Ji Hwan Cha

Ewha Womans University
Seoul, Korea
jhcha@ewha.ac.kr

A. Charki

Institut des Sciences et Techniques
de l'Ingénieur d'Angers
LASQUO EA3858 Angers
F-49000 France
abderafi.charki@univ-angers.fr

S. Cloupet

Institut des Sciences et Techniques
de l'Ingénieur d'Angers
LASQUO EA3858 Angers
F-49000 France
sylvain.cloupet@univ-angers.fr

Enrico A. Colosimo

Department of Statistics
Federal University of Minas
Gerais, Brazil
enricoc@est.ufmg.br

Vincent Couallier

IMB University of Bordeaux
146 rue Léo Saignat
33076 Bordeaux, France
couallier@sm.u-bordeaux2.fr

Thiago R. dos Santos

Department of Statistics
Federal University of Minas Gerais
Brazil

Maxim Finkelstein

University of the Free State Bloemfontein
Republic of South Africa
and Max Planck
Institute for Demographic Research
Rostock, Germany
FinkelM.SCI@ufs.ac.za

Marta A. Freitas

Department of Industrial Engineering
Federal University of Minas
Gerais, Brazil
marta@dep.ufmg.br

Christopher Gast

Axio Research Acquisition Co. LLC
2601 4th Avenue Suite 200
Seattle, WA 98121 USA
cmgast@gmail.com

Léo Gerville-Réache

Université de Bordeaux
UMR 5251
Bordeaux F-33000, France
leo.gerville@u-bordeaux2.fr

Fabrice Guérin

Institut des Sciences et Techniques
de l'Ingenieur d'Angers
LASQUO EA3858 Angers
F-49000 France
fabrice.guerin@istia.univ-angers.fr

Allan Gut

Dept. of Mathematics
Uppsala University
Box 480, SE-751 06 Uppsala, Sweden
allan.gut@math.uu.se

Firoozeh Haghighi

University of Tehran
Department of Mathematics,
Statistics, and Computer Sciences
Tehran, Iran
haghighi@khayam.ut.ac.ir

Takashi Hara

Department of Systems Engineering
University of Electro-Communications
1-5-1 Chofugaoka, Chofu
Tokyo 182-8585 Japan

B.P. Harlamov

Institute of Problems of Mechanical
Engineering
Russian Academy of Science
St. Petersburg, Russia
b.p.harlamov@gmail.com

Catherine Huber

Université Paris Descartes
45 rue des Saints-Pères
75 270 Paris Cedex 06
and U780 INSERM
94 800 Villejuif, France
catherine.huber@parisdescartes.fr

Jürg Hüsler

Dept. of Statistics
University of Bern
Sidlerstrasse 5
CH-3012 Bern, Switzerland
huesler@stat.unibe.ch

Waltraud Kahle

Otto-von-Guericke-University
Magdeburg, Germany
waltraud.kahle@ovgu.de

N. Kannan

Department of Management Science and
Statistics
The University of Texas at San Antonio
San Antonio, Texas, USA
nandini.kannan@utsa.edu

V.S. Koroliuk

Institute of Mathematics
Ukrainian National Academy of Science
Kiev, Ukraine
korol@imath.kiev.ua

D. Kundu

Department of Mathematics and
Statistics
Indian Institute of Technology Kanpur
Pin 208016, India
kundu@iitk.ac.in

Axel Lehmann

Hochschule für Technik Wirtschaft und
Kultur Leipzig
University of Applied Sciences
Department of Computer Science
Mathematics and Natural Sciences
04251 Leipzig, Germany
lehmann@imn.htwk-leipzig.de

N. Limnios

Laboratoire de Mathématiques
Appliquées
Université de Technologie de Compiègne
60205 Compiègne Cedex
France
nikolaos.limnios@utc.fr

M. LuValle

OFS Laboratories
Somerset, NJ USA
mj1@ofsoptics.com

I. Masiulaitytė

Vilnius University
Vilnius, Lithuania

Md. Mesbahul Alam

Department of Systems Engineering
University of Electro-Communications
1-5-1 Chofugaoka Chofu
Tokyo 182-8585, Japan

William Q. Meeker

Department of Statistics and
Center for Nondestructive Evaluation
Iowa State University
Ames, Iowa 50010 USA
wqmeeker@iastate.edu

Narges Nazeri Rad

University of Tehran
Department of Mathematics,
Statistics, and Computer Sciences
Tehran, Iran

Wayne B. Nelson

Consultant
Schenectady, NY USA
WNconsult@aol.com

M.S. Nikulin

IMB
Victor Segalen University
Bordeaux, France
nikou@sm.u-bordeaux2.fr

Nazanin Nooraee

University of Tehran
Department of Mathematics
Statistics and Computer Sciences
Tehran, Iran

Francis G. Pascual

Department of Statistics
Washington State University
Pullman, WA 99164 USA
jave@wsu.edu

Magda C. Pires

Department of Statistics
Federal University of Minas Gerais
Brazil

S.S. Rasova

Institute of Problems of Mechanical
Engineering
Russian Academy of Science St.
Petersburg
Russia

I.V. Samoilenko

Institute of Mathematics
Ukrainian National Academy of Science
Kiev, Ukraine

Nozer D. Singpurwalla

The George Washington University
Washington, D.C. 20052 USA
nozer@gwu.edu

V. Solev

Steklov Institute of Mathematics
at St. Petersburg
nab. Fontanki 27 St.Petersburg
191023 Russia
solev@pdmi.ras.ru

Kazuyuki Suzuki

Department of Systems Engineering
University of Electro-Communications
1-5-1 Chofugaoka Chofu
Tokyo 182-8585 Japan
suzuki@se.uec.ac.jp

A. Todaskoff

Institut des Sciences et Techniques
de l'Ingenieur d'Angers
LASQUO EA3858 Angers
F-49000 France

Watalu Yamamoto

Department of Systems Engineering
University of Electro-Communications
1-5-1 Chofugaoka Chofu
Tokyo 182-8585, Japan
watalu@se.uec.ac.jp

R. Viertl

Institute of Statistics and Probability
Theory
Vienna University of Technology
1040 Wien, Austria
R.Viertl@tuwien.ac.at

Filia Vonta

Department of Mathematics and
Statistics
University of Cyprus P.O. Box 20537
CY-1678 Nicosia Cyprus
and Department of Mathematics
National Technical University of Athens
Athens, Greece
vonta@ucy.ac.cy
vonta@math.ntua.gr

Heide Wendt

Otto-von-Guericke-University
Magdeburg, Germany

S. Zacks

Binghamton University
Department of Mathematical Sciences
Binghamton, NY 13902-6000 USA
shelly@math.binghamton.edu

List of Tables

3.1	Dendrite age and size data	29
3.2	ML estimates of parameters and confidence limits for dendrite data	35
4.1	Information available from warranty database	41
4.2	Parameter setting for examining properties of estimators (one failure mode)	45
4.3	Asymptotic variances of estimators (one failure mode)	45
4.4	Changes in the asymptotic variances with change in observational period, $t = 1, 3, 5$ ($m_a = 2.0$)	45
4.5	Changes in the Asymptotic variances with change in observational period, $t = 1, 3, 5$ ($m_a = 1.0$)	45
4.6	Changes in the asymptotic variances with change in observational period, $t = 1, 3, 5$ ($m_a = 0.7$)	46
4.7	Parameter setting for examining properties of estimators (two failure modes)	46
4.8	Asymptotic variances of parameters (two failure modes)	46
6.1	Empirical moments of maximum likelihood (MLE) and moment (ME) estimators ($\theta_0^Y = (2.4, 0.5, 1.2)$)	102
6.2	Empirical moments of maximum likelihood (MLE) and moment (ME) estimators ($\theta_0^Y = (2.4, 0.5, 1.2)$)	102
9.1	Mean-squared errors of maximum likelihood estimators in 1000 simulation runs	145
11.1	Pseudo failure distances	168
11.2	Interval and point estimates obtained by each method	170
11.3	Bayesian estimates of the quantities of interest and 95% HPD regions (prior: Weibull distribution)	175
11.4	Bayesian estimates of the quantities of interest and 95% HPD regions (prior: lognormal distribution)	175
17.1	Ball bearing parameters	261
17.2	Simulation parameters	261
17.3	Simulation results (N in hours)	261
17.4	MLE estimations of model parameters	262
17.5	Mean and standard deviation of MLE estimations after 20 repeatings	262

17.6	Mean and standard deviation of MLE estimations after 20 repeatings for different censoring times	264
17.7	Simulation data	270
17.8	Mean and standard deviation of MLE estimations after 20 repeatings	271
19.1	Confidence level for finite samples ($n_1 = n_2 = 100$)	302
19.2	Confidence level (parametric estimation) for finite samples ($n_1 = n_2 = 100$)	304
19.3	Significance level	308
19.4	Power of the test	309
19.5	Significance level	309
19.6	Power of the test	309
19.7	Significance level	309
19.8	Power of the test	310
20.1	Ten component demonstration tests with confidence 70%. Equal test times $TT = 10,000$ hours provide different objective values for the $T10\%$ targets	314
20.2	Ten component demonstration tests with confidence 70%. Equal reliability targets require different test times	315
20.3	Component demonstration times and minimum true $t_{10\%}$ for guaranteeing a 50% global probability of success for 0/1 and 3/1 demonstration plans	318
22.1	Raw pipes data format	347
22.2	Raw failure data format	347
22.3	Distribution of the number of failures	350
22.4	Estimation results	351
23.1	Estimates and confidence intervals for the control group	361
23.2	Estimates and confidence intervals for the test group	361
23.3	$\tau_1 = 100, \tau_2 = 150, \theta_1 = 1/100, \theta_2 = 1/50$	363
23.4	Known Ratio	363
23.5	$\tau_1 = 100, \tau_2 = 150, \theta_1 = 1/100, \theta_2 = 1/25$	363
23.6	Known ratio	364
23.7	$\tau_1 = 100, \tau_2 = 125, \theta_1 = 1/100, \theta_2 = 1/50$	364
23.8	Known ratio	364
23.9	$\tau_1 = 100, \tau_2 = 125, \theta_1 = 1/100, \theta_2 = 1/25$	364
23.10	Known ratio	367
25.1	Discovery rates for reconstruction procedures 1,2, and 3	393
26.1	Maximum likelihood results for the cancer data set	406
26.2	Maximum likelihood results for the shock absorber data set	409
26.3	Maximum likelihood results for a simulated data set assuming that risk 1 and risk 2 are independent and distributed, respectively, lognormal with $\mu_1 = 6$ and $\sigma_1 = 1$ and Weibull with $\mu_2 = 5$ and $\sigma_2 = 0.7$	410

List of Figures

3.1	Display of dendrite size data	29
3.2	Blister size $Y_i(t)$ over time t	30
3.3	Dendrite size $Y_i(t)$ over time t	31
3.4	Basic model with areas below 0 for the population fraction not initiated	32
3.5	Basic model with initiation time and failure time distributions . . .	33
4.1	Effect of number of sold products N_t and observational period t on $\sqrt{AVar\left(\frac{\hat{m}_a}{m_a}\right)}$	47
4.2	Effect of number of sold products N_t and observational period t on $\sqrt{AVar\left(\frac{\hat{\eta}_a}{\eta_a}\right)}$	47
5.1	Realization of a sequence of shocks with strengthening and weakening load limits, depending on the values X_i . Here $\nu = 38, W = 18, N_+(\nu) = 5, N_-(\nu) = 3$	70
6.1	A realization of $\Phi(t)$ and $Z(t)$	79
6.2	Mean cumulative degradation process for different markings of (T_n)	85
6.3	Density of the first passage time Z^h for different η	89
6.4	Distribution function of Z^h for different X_0	91
6.5	Confidence estimates for $\theta^X = (p, \delta)$ and different η	98
6.6	Confidence estimates for $\theta^Y = (\mu, \beta)$ and $\theta^T = (\alpha, \gamma)$	99
9.1	Sample for observations of process increments for the case $n = 2, m_1 = 3, m_2 = 2$	130
9.2	Confidence estimates for (μ, σ^2)	142
9.3	Confidence estimates for (μ, σ^2)	143
9.4	Confidence estimates for (μ, h)	143
9.5	Confidence estimates for (σ^2, h)	144
10.1	Simulated observed paths for 20 units. Above <i>left</i> ($\sigma = 0.01$), above <i>right</i> ($\sigma = 0.0055$), and <i>Bottom</i> ($\sigma = 0.0025$) [model (10.2)]	152
10.2	True survival function (<i>solid line</i>) versus the nonparametric estimator of survival function for $n = 20, 50, 100$ [model (10.2)]	152
10.3	Simulated sample paths without measurement error [model (10.8)]	153

10.4	Parametric estimators of $S(t)$ and its 90% confidence intervals [model (10.8)]	153
10.5	The comparison between parametric (<i>dotted line</i>) and nonparametric (<i>solid line</i>) estimators of $S(t)$ for different sample sizes [model (10.8)]	154
11.1	Plot of the wheel degradation data	162
11.2	Probability plots for the pseudo failure distances	169
11.3	Comparison of the parametric and nonparametric (Kaplan–Meier) estimates of $R_T(t)$ at each evaluation point	169
11.4	Point estimates (MTTF, $t_{0.10}, t_{0.50}$) and confidence intervals obtained by each method of degradation data analysis (“classical” inference). Weibull distribution	173
11.5	Point estimates (MTTF, $t_{0.10}, t_{0.50}$) and confidence intervals obtained by each method of degradation data analysis (“classical” inference). Lognormal distribution	173
11.6	Point estimates of $R(300,000)$ and confidence intervals obtained by each method of degradation data analysis (“classical” inference). Weibull and lognormal distributions	174
11.7	Histograms of the posterior distributions for $R(300,000)$, $t_{0.10}$, and the mean covered distance, respectively. Weibull model	176
11.8	Histograms of the posterior distributions for $R(300,000)$, $t_{0.10}$ and, the mean covered distance, respectively. Lognormal model	177
12.1	A trivial solution	187
12.2	A trivial solution of the second kind	187
12.3	Trajectories of choice with moving on SO^e and on NW	191
12.4	Derivative of $G_r(a)$ with respect to r as $r = 0$ for different drift . .	193
13.1	The reliability function $R_1(t; 15, 10, 15)$	203
13.2	The hazard function $A_1(t; 15, 10, 15)$ with $\mu = 2$ and $\lambda = 1$	204
13.3	Reliability function $R_2(t; \tau_1, \tau_2, \beta)$ with $K = 2$	205
13.4	The reliability function $R_1^*(t; k, \beta_0, \beta_1)$	206
13.5	The reliability function $R_c(t; \gamma_0, \beta_0, \beta_1)$	207
13.6	The reliability function $R_L(t; \beta)$	208
14.1	Wiener process degradation model	212
14.2	Gamma process degradation model	212
14.3	Failure rates of Inverse Gaussian distribution ($x^* = h$)	217
14.4	Failure rates of the distribution (14.7)	219
14.5	Wiener process DTS model with exponential timescale	224
14.6	Gamma process DTS model with exponential timescale	224
15.1	Degradation of a quality parameter $q(t)$	236
15.2	Solutions of Equation (15.4)	237
15.3	Fuzzy-valued function	238
15.4	Fuzzy distribution function	240
17.1	Definition of time transformation function	255
17.2	Definition of the studied test plan	256
17.3	Record of failure times	258
17.4	Reliability functions by the different ALT models (N in hours) . .	262
17.5	Reliability function by parametric model (N in hours)	263
17.6	Reliability function by GPH model (N in hours)	263

17.7	Reliability function by semi-parametric model (N in hours)	264
17.8	Reliability functions by parametric model with censoring at 1,000 hours (a) and 300 hours (b)	265
17.9	Reliability functions by GPH model with censoring at 1,000 hours (a) and 300 hours (b)	265
17.10	Reliability functions by semi-parametric model with censoring at 1,000 hours (a) and 300 hours (b)	265
17.11	Example of error propagation	266
17.12	Principle of test plan with previous accelerated damage	267
17.13	Definition of test plan with previous accelerated damage	268
17.14	Definition of pdf under different stresses	268
17.15	Estimation of the reliability function by parametric and nonparametric models	271
19.1	Graphs of the trajectories of the nonparametric estimators \hat{F}_1, \hat{K}_i (Weibull distribution)	295
19.2	Graphs of the trajectories of the nonparametric estimators \hat{F}_1, \hat{K}_i (loglogistic distribution)	296
20.1	Probability of failure of a “zero-failure” demonstration with 70% confidence and $\beta = 2.08$	317
22.1	Conditional intensities of two sets of simulated times (<i>plain lines</i>), ROCOF function (<i>thick dashed line</i>), initial intensity (<i>thin dashed line</i>) for $\alpha = 1.5$ and $\delta = 1$. $\lambda(t) = (1 + \alpha i)e^{-2.1\delta t^{\delta-1}}$	343
22.2	Conditional intensities of two sets of simulated times (<i>plain lines</i>), ROCOF function (<i>thick dashed line</i>), initial intensity (<i>thin dashed line</i>) for $\alpha = 1.5$ and $\delta = 1.5$. $\lambda(t) = (1 + \alpha i)e^{-2.1\delta t^{\delta-1}}$	343
22.3	Data collection schema	346
22.4	Observation and prediction schema	347
22.5	Bar chart of installation dates	348
22.6	Bar chart of replacement dates	348
22.7	Mean annual failure rate per pipe during 1976–1985: <i>black</i> , 1986–1995: <i>darkgrey</i> , 1996–2006: <i>lightgrey</i>	349
22.8	Mean annual failure rate per km of length during 1976–1985: <i>black</i> , 1986–1995: <i>darkgrey</i> , 1996–2006: <i>lightgrey</i>	349
22.9	Number of monthly failures over the period 1976–2006, the failures of the year 1996 in <i>black</i>	350
22.10	Cumulat number of observed failures (in <i>grey</i>), predicted failures using the complete model (<i>plain black line</i>) and the reduced model (<i>dotted line</i>)	352
22.11	Prediction error in 10-day periods (real failures – predicted failures) calculated using the complete model (<i>plain line</i>) and reduced model (<i>dotted line</i>)	353
23.1	Cumulative hazard functions of cumulative exposure and cumulative risk models	359
23.2	Empirical and predicted CH functions: test group	362
23.3	Sampling distribution of the MLE of a	365
23.4	Sampling distribution of the MLE of b	365
23.5	Sampling distribution of the MLE of θ_1	366
23.6	Sampling distribution of the MLE of θ_2	366

25.1	Revealing projection for Lorenz attractor	386
25.2	Null and alternative overlay in revealing projection	387
25.3	Large neighborhood of attractor in parameters for NULL plus alternative overlay	391
25.4	False discovery rate plots, alternative and null hypothesis, Chi square, and computational null distributions, points below line are rejections of null hypothesis	393
26.1	Plot of the KM estimate of the cumulative distribution function F with pointwise confidence intervals for the cancer data set	405
26.2	Lognormal probability plot of survival times for deaths due to cancer with estimates of the cdf assuming that cancer is the only risk, Other is the only risk, and cancer and Other are independent lognormal and Weibull, respectively	405
26.3	Weibull probability plot of survival times for deaths due to Other with estimates of the cdf assuming that cancer is the only risk, Other is the only risk, and cancer and Other are independent lognormal and Weibull, respectively	406
26.4	Probability plots for the cancer data set with pointwise confidence intervals for failure-time cdf F assuming independent lognormal cancer deaths and Weibull other deaths. Plot (a) is for deaths due to cancer. Plot (b) is for deaths due to other causes.	407
26.5	Probability plots for the cancer data set with pointwise confidence intervals for failure-time cdf F assuming independent lognormal risks. Plot (a) is for deaths due to cancer. Plot (b) is for deaths due to other causes	408
26.6	Plot of the KM estimate of the cumulative distribution function F with pointwise confidence intervals for the shock absorber data set	408
26.7	Probability plots for the shock absorber data set with pointwise confidence intervals for failure-time cdf F assuming independent Weibull M1 failures and lognormal M2 failures. Plot (a) is for failures due to mode M1. Plot (b) is failures due to mode M2.	409
26.8	Probability plots for the shock absorber data set with pointwise confidence intervals for failure-time cdf F assuming independent Weibull failure times. Plot (a) is for failures due to mode M1. Plot (b) is failures due to mode M2.	410
26.9	Probability plots for the simulated data set (lognormal risk 1, Weibull risk 2) with pointwise confidence intervals for failure-time cdf F assuming independent lognormal risks. Plot (a) is for failures due to risk 1. Plot (b) is failures due to risk 2.	411
26.10	Probability plots for the simulated data set (lognormal risk 1, Weibull risk 2) with pointwise confidence intervals for failure-time cdf F assuming independent lognormal risk 1 failures and Weibull risk 2 failures. Plot (a) is for failures due to risk 1. Plot (b) is failures due to risk 2	412

Review, Tutorials, and Perspective

Trends in the Statistical Assessment of Reliability

William Q. Meeker

Department of Statistics and Center for Nondestructive Evaluation, Iowa State University,
Ames, Iowa

Abstract: Changes in technology have had and will continue to have a strong effect on changes in the area of statistical assessment of reliability data. These changes include higher levels of integration in electronics, improvements in measurement technology and the deployment of sensors and smart chips into more products, dramatically improved computing power and storage technology, and the development of new, powerful statistical methods for graphics, inference, and experimental design and reliability test planning. This chapter traces some of the history of the development of statistical methods for reliability assessment and makes some predictions about the future.

Keywords and phrases: Accelerated testing, Bayesian methods, degradation data, maximum likelihood, multiple failure modes, nonparametric estimation, product design, recurrence data, statistical software, warranty data

1.1 Background and Purpose

Today's manufacturers need to develop new, higher technology products in record time while improving quality, reliability, and productivity. Much of this need has been driven by the expanding global marketplace and the resulting increased competition. Manufacturers of high-quality and high-reliability products have a strong competitive advantage. These manufacturers are, however, facing

- Need for rapid product development
- Changing technologies/new materials
- More complicated products with more components
- Higher customer expectations for better reliability

As suggested by Condra [5], reliability can be defined as *quality over time*. The improvements in quality of manufactured products that we have seen over the past 30 years (e.g., in automobiles manufactured in the United States) have also had the effect of improving product reliability. It has been recognized, however, that achieving

and improving high reliability requires tools that lie beyond the standard tools used in quality improvement. Generally achieving high reliability requires careful focus on the time dimension.

Reliability is a highly quantitative engineering discipline. Probability theory and statistical models and methods play an important role in reliability. In particular, many reliability-related decisions require the analysis of reliability data, either from past field performance or from laboratory tests (usually accelerated tests). One important consideration, that separates reliability from many other applications of statistics, is that extrapolation is present in almost all applications. For example, we extrapolate in time when we have 1 year of data but have to predict warranty returns going out 3 years. We extrapolate in temperature when we test units at high temperatures and then estimate life at use conditions. We extrapolate from past experience when we use knowledge about a model that worked well in the past to describe a new situation. After conducting a laboratory test, we extrapolate to behavior in the field.

The purposes of this non-technical paper are to outline some trends in the use of statistics in reliability and connect these with changes in technology and to predict what we can expect to see in the future, indicating areas where more research will be needed.

1.2 Traditional Reliability Data and Data Analysis

Reliability data arise from a number of different sources, including laboratory life tests, field-tracking studies, and warranty databases. Traditional reliability data have consisted of failure times for units that failed and running times for units that had not failed. Interestingly (as can be seen by reading old papers in the engineering literature), there was a long period of time when many engineers thought that it was necessary that all units fail so that the life data could be analyzed. Of course, methods for computing estimates of failure distributions from what is essentially censored data have been used for centuries in medical and insurance applications. Methods for analyzing censored data (nonparametric estimation and maximum likelihood) were further developed in the 1950s and the 1960s and became well known to most statisticians by the 1970s. The bibliographic guide provided by Buckland [4] nicely outlines the important references in this area up to that point in time. Books written by Mann, Schafer, and Singpurwalla [27], Lawless ([22] second edition [23]), Nelson [37], Cox and Oakes [7], and Crowder et al. [8] document many of the important advances made during the 1960s and 1970s. Statistical methods for the analysis of censored data began to appear in commercial software by the mid-1980s (starting with SAS) and are commonplace today.

The most popular tool for life data analysis is the probability plot, used to assess distributional goodness of fit, detect data anomalies, and display the results of fitting parametric distributions in the presence of censored data. It should be more widely recognized, however, that a probability plot is a valuable tool for general data analysis, even when there is not any censoring. A community of engineers has long championed what has been called *Weibull analysis* which implies fitting a Weibull distribution to failure data. Weibull analysis, for example, was described in Abernethy, Breneman, Medlin and Reinman [1]. The Weibull distribution is not, however, always the best

distribution to use, and modern software allows fitting a number of different parametric distributions. The vast majority of applications in reliability use either the Weibull or lognormal distribution. One reason for this is that there are strong mechanistic justifications that suggest these distributions, much as the central limit theorem can sometimes be used to explain why some random variables should be well described by a normal distribution.

The method of maximum likelihood (ML) and likelihood-based inference methods are at the core of almost all reliability applications. For example, all of the statistical methods (even the nonparametric methods) used in Meeker and Escobar [31] are based on ML. There are two primary reasons for using ML. First, ML is highly versatile; it is hard to find statistical problems where ML cannot be used. Second, under mild conditions, ML methods are known to be optimum in large samples. Even when these mild conditions do not hold, it is hard to find competitors that are consistently better.

In most applications of reliability, it is important to quantify statistical uncertainty (i.e., uncertainty due to limited data). Confidence intervals are the most commonly used method to do this. Normal theory-based large-sample approximate confidence intervals are still the most commonly used method to construct confidence intervals, in spite of the fact that there have been examples and studies to show that these intervals can be horribly inadequate (i.e., having coverage probability far from the nominal value), even with large-sample sizes. Either likelihood-based confidence intervals or simulation-based confidence intervals dominate the crude normal approximations (e.g., Efron and Tibshirani [12], Meeker [29], Vander Weil and Meeker [45], and Jeng and Meeker [21]). Before too long, improving computer power and clever implementation schemes should, for many applications, eliminate the use of normal-approximation intervals.

1.3 Product Design and Reliability Budgeting

The use of probabilistic modeling for reliability is becoming more common in product design. Historically, engineers used deterministic models for product design and applied *factors of safety* to assure that a product would be reliable. In a highly competitive market, however, being overly conservative may require a price that is too high for the market. In recent years, many companies have instituted *Reliability by Design* or *Design for Six Sigma* programs that involve some kind of reliability assurance process based on probabilistic system reliability models. Such models are described in books on system reliability such as Rausand and Hoyland [42]. These models require, as inputs, the reliability of individual components and interfaces. Engineers need to obtain component reliability information in a timely and economical manner. The most common and least expensive sources of such information are standard values in handbooks (e.g., MIL-HDBK-217F [34]) and previous field experience. If the information is not available from these relatively inexpensive sources (e.g., because the component is new or an old component is being used in a different application or environment that could affect its reliability), engineers may have to conduct their own tests in the laboratory. Most such tests would be *accelerated* so that the information is obtained in a timely manner.

Another source of component life information that is becoming more widely used is physical/chemical or other mechanistic models for failure. For example, finite element

models can sometimes be combined with other engineering/physical knowledge to predict reliability of a component. Such methods are described, for example, in Haldar and Mahadevan [18]. Of course, in many applications, some actual physical testing is still required to provide inputs to the models or to verify parts of the models. The use of such models in reliability often requires large amounts of time and effort in development. The hope is that such investments will result in lower costs in the future due to savings from not having to do as much actual physical testing.

1.4 Accelerated Testing

When a component must last for years or even decades, it may be possible to conduct an *accelerated test* to obtain information in a timely manner. Virtually all life tests done to evaluate reliability are accelerated in one way or another. The basic idea is to test units at high levels of cycling rate, temperature, voltage, stress, or another accelerating variable to get reliability information quickly. Then a physically motivated model is used to extrapolate to use conditions. Nelson [40], originally published in 1990, is the most important reference on the statistical aspects of accelerated testing. This book describes models, statistical methods for analysis, and statistical methods for test planning.

It is sometimes said, jokingly, that engineers are very good at conducting accelerated tests that cause units to fail quickly. Hitting a sample of integrated circuits with a hammer will certainly make them fail, but will provide no useful information about product reliability. The important, serious, question is whether the failures generated in the accelerated test provide useful information about how the product will behave at actual use conditions. There are numerous examples (e.g., two of the three examples in Chapter 19 of Meeker and Escobar [31]) where using too much stress or a temperature that was too high caused a new failure mode that, if not recognized and accounted for, would lead to overly optimistic estimates of life at use conditions. Thus, the most important challenge of accelerated testing is deciding how to accelerate, how much to accelerate, and to find an adequate model to relate the results from the accelerated test to actual use conditions.

1.5 Multiple Failure Modes

For some applications, when analyzing reliability data, it is important to distinguish among different product failure modes. The reasons for this are as follows:

- When the failure modes behave differently (e.g., some are defect or infant mortality related and others are caused by wear out), it is generally easier to find a well-fitting failure-time distributions for the individual failure modes (e.g., Example 15.6 in Meeker and Escobar [31]).

- When forecasting warranty costs, some failure modes are much more expensive to fix than others (e.g., replacing a mother board versus replacing a defective battery in a computer). In some applications, there is one failure mode that is of critical importance (e.g., a failure mode that could cause serious harm) and others that are innocuous, leading to end of life of the product and thus eliminating the possibility of the critical failure mode.
- Predictions are often needed for the number of replacement parts that will be needed to effect repairs.
- Knowledge of the relative frequency of different failure modes and the effect of eliminating one or more of the individual failure modes is important for engineers who need to make design changes that will improve product reliability and reduce warranty costs.

When failure-mode information is available for all failed units and when the different failure modes can be assumed to be statistically independent, the analysis of multiple failure-mode data is, technically, not much more difficult than it is for analyzing a single failure mode.

Competing risk theory provides the appropriate statistical model for life data with multiple failure modes. The first book-length treatment of the theory of competing risk was provided by David and Moeschberger [10]. For single distribution applications in reliability, these methods are described and illustrated with examples in Chapter 5 of Nelson [37], Chapter 15 of Meeker and Escobar [31], and Crowder [9]. Nelson [38] describes methods for analyzing accelerated life test data with multiple failure modes. Life data analysis with multiple failure modes is, however, greatly facilitated when software tools have been designed to make the needed operations easy. Today, several statistical packages provide capabilities for estimating separate distributions for each failure mode and for making assessments of improvement in product reliability by eliminating one or more of the failure modes.

When failure-mode information is missing for some or all of the failed units (known as masked causes of failure) or when the failure modes cannot be described by the simple independence model, the analysis is more complicated and special methods, not generally available in software, have to be employed. Flehinger, Reiser, and Yashchin [15], for example, describe statistical methods for dealing with masked data.

When the failure times for different failure modes are not independent, one approach is to use some kind of multivariate failure-time model. In practice, however, there is not enough information to identify such a model. When it is important to estimate the marginal distributions for each failure mode (e.g., to predict the effect of removing a failure mode), it is often possible to collapse a large number of related failure modes into several groups such that the failure modes are approximately independent. Such collapsing of failure modes is common in practice (e.g., p. 34 of David and Moeschberger [10]). Meeker, Escobar, and Hong [32] describe an application where many failure modes were collapsed into two groups, engineering judgment was used to make an assumption about the degree of dependency, and sensitivity analysis was used to assure that inferences were reasonably accurate over a plausible range of assumptions. In some applications, estimation of the sub-distribution functions (e.g., p. 4 of Crowder [9]), which are identifiable, is sufficient.

1.6 Field and Warranty Data

Although laboratory reliability testing is often used to make product design decisions, the “real” reliability data come from the field, often in the form of warranty returns or, specially designed field-tracking studies. Warranty databases were initially created for financial-reporting purposes, but more and more companies are finding that warranty data are a rich source of reliability information. Perhaps 6–8 months after a product has been introduced into the market (sooner if costs have already been higher than expected), managers begin to ask about warranty costs over the life cycle of the product. Two common problems with warranty data are that good failure-mode information is rarely available (there is usually some kind of code in the database, but it is usually of limited use to determine the actual cause of failure) and warranty data are generally heavily censored, because information that is obtained after a product is out of warranty is limited. Thus, even though companies should be concerned about reliability of their products far beyond the end of a warranty period, operationally, little data is available.

For some products, careful field tracking provides good reliability data. Examples include medical devices and a company’s fleet of assets (e.g., information about the reliability for a fleet of automobiles).

There are sometimes large gaps between predictions made from product design models (supplemented by limited reliability testing) and reality. These differences are often caused by unanticipated failure modes. Algorithms for early detection of emerging reliability issues (e.g., Wu and Meeker [46]) are being implemented in software and have the potential to save companies large amounts of money. Once a new emerging issue has been identified, statistical methods (e.g., Escobar and Meeker [13], and Lawless and Fredette [25]) can be used to produce forecasts for the additional warranty costs. Field data also provide important feedback that can be used to assess the adequacy of reliability predictions methods and to give engineers information on how to design future products.

1.7 Degradation Reliability Data

In modern high-reliability applications, we might not expect to see failures in our reliability testing, resulting in limited information about reliability needed for product design. Suppose that 500 units had been put on test and run for 1000 hours of operation for a product that is required to last 5000 hours. If there are no failures at the end of the test, there would be little or no information for quantifying reliability (depending on assumptions that one might be willing to make). If, however, we could monitor, over time, a degradation (or a performance) variable that is closely related to failure (e.g., length of a fatigue crack or light output from a laser) on all of the test units, there would be a large amount of reliability information.

There are a number of other advantages for using such repeated-measures degradation data for reliability assessment. For example, degradation data provide information

that is much richer for building and assessing the adequacy of physical/chemical models used for test acceleration.

Today the term *degradation* refers to either performance degradation (e.g. light output from an LED) or some measure of actual chemical degradation (e.g., concentration of a harmful chemical compound). Over the past 30 years, we have seen many different kinds of applications where degradation data were available. Some early ideas on the use of degradation models in reliability were given in Gertsbakh and Kordonsky [17]. Through the 1990s and continuing today, statistical methods have been developed for making reliability inferences from degradation data. Initially these were developed by researchers or engineers in need of the methods. Statistical methods for the analysis of degradation data are, however, now beginning to be deployed in commercial statistical software. Some engineers (e.g., Murray [36]) had been using informal *simple* methods of analysis that fit models to the sample path for individual units and extrapolated these until some failure level, providing *pseudo failure* data that could be analyzed by common life data analysis methods. Lu and Meeker [26] and Meeker, Escobar, and Lu [33] used a more sophisticated random effects model to describe unit-to-unit variability and showed how the degradation model, along with a failure definition, induces a failure-time distribution. In some areas of application, it is necessary to model the stochastic behavior in the sample paths over time. Lawless and Crowder [24], for example, use such a model.

There have been a number of examples where the natural response in a reliability test is a degradation variable, but the analysts (at least initially) turned the degradation data into failure-time data, because all of the test books and software known to them dealt only with the analysis life data. The application described in Meeker, Escobar, and Lu [33] is one such application. In these examples, the limited number of failures provided only limited reliability information and the results of a degradation analysis were more informative.

When an appropriate degradation variable can be measured, degradation data, when properly analyzed, can provide much more information because there are quantitative measurements on all units (not just those that failed). Indeed, it is possible to make powerful reliability inferences from degradation data even when there are no failures. It is, of course, not always possible to find a degradation variable that corresponds to a failure mode of concern.

Even when a repeated measures degradation variable is not available, it might be possible to do destructive tests to evaluate units that have not failed. Examples of such *destructive degradation tests* are given in Chapter 11 of Nelson [38] and in Escobar et al. [14].

1.8 Recurrence Data

The discussion in the previous sections dealt with reliability data analysis for non-repairable components (or devices). Since a nonrepairable component can fail only once, time to failure data from a sample of nonrepairable components consists of the times to first failure for each component. In many applications involving nonrepairable components, the assumption of independent and identically distributed failure times

(or at least deviations from an assumed model that might have more than one component of variability) is a reasonable one and suitable lifetime distributions (such as the Weibull or lognormal) are used to describe the distribution of failure times. In contrast, repairable system data typically consist of the times of multiple repairs (or other events of interest) on the same system. Such data are known as recurrence data. In the discussion here we will, for concreteness, refer to recurrences as *repairs*. The purpose of some reliability studies is to describe the trends and patterns of repairs or failures for an overall system or collection of systems over time. Data consist of a sequence of system repair times for similar systems. When a single component or subsystem in a larger system is repaired or replaced after a failure, the distribution of the time to the next system repair will depend on the nature of the repair, and the overall state of the system at the time just before the current repair and the nature of the repair. Thus, repairable system data, in many situations, should be described with models that allow for changes in the state of the system over time or for dependencies between repairs over time. A number of books have been written that describe the many technical advances in this area over the past 20 years. Nelson [39] describes basic graphical and simple, but effective, nonparametric methods for recurrence data analysis and illustrates the methods with a wide range of applications. These methods have been implemented in several computer packages. The work by Cook and Lawless [6], which was written at a higher technical level, describes methods and applications involving a wide variety of nonparametric, parametric, and semiparametric recurrence data models, including regression models. The authors show how readily available software can be used to implement the methods described in their book.

1.9 The Next Generation of Reliability Data

Due to changes in technology, the next generation of reliability field data will be richer in information. Use rates and environmental conditions are important sources of variability in product lifetimes. The most important differences between carefully controlled laboratory accelerated test experiments and field reliability results are due to uncontrolled field variation (unit to unit and temporal) in variables like use rate, load, vibration, temperature, humidity, UV intensity, and UV spectrum. Historically, use rate/environmental data have, in most applications, not been available to reliability analysts. Incorporating use rate/environmental data into our analyses will provide stronger statistical methods. Today it is possible to install sensors and smart chips in a product to measure and record use rate/environmental data over the life of the product. In addition to the time series use rate/environmental data, we also can expect to see further developments in sensors that will provide information, at the same rate, on degradation or indicators of eminent failure. Depending on the application, such information is also called *system health* and *materials state* information. In some applications (e.g., aircraft engines and power distribution transformers), system health/use rate/environmental data from a fleet of products in the field can be returned in real time to a central location for real-time process monitoring and especially for prognostic purposes. An appropriate signal in these data might provoke rapid action to avoid a serious system failure (e.g., by reducing the load on an unhealthy transformer). Also,

should some issue relating to system health arise at a later date, it would be possible to sort through historical data that have been collected to see if there might have been a detectable signal that could be used in the future to provide an early warning of the problem.

In products that are attached to the Internet (e.g., computers and high-end printers), such use rate/environmental data can, with the owner's permission, be downloaded periodically. In some cases, use/environmental data will be available on units only when they are returned for repair (although monitoring at least a sample of units to get information on unfailed units would be statistically important).

The future possibilities for using use rate/environmental data in reliability applications are unbounded. Lifetime models that use rate/environmental data have the potential to explain much more variability in field data than has been possible before. The information can also be used to predict the future environment lifetimes of individual units. This knowledge can in turn provide more precise estimates of the lifetime of individual products. As the cost of technology drops, cost-benefit ratios will decrease and applications will spread.

1.10 Software for Statistical Analysis of Reliability Data

Newly developed statistical methods will not see widespread use until they are implemented in easy-to-use software. Thus, it is important to have such software to do reliability data analysis. The first major software system for reliability data analysis was designed and developed by Wayne Nelson. This package, called STATPAC, was described in Nelson and Hendrickson [41] and Strauss [44]. STATPAC was far ahead of its time and contained a combination of capabilities for graphical analysis and fitting of general statistical models with censored data that are available today in only the most advanced statistical packages. Meeker and Duke [30] provided a useful computer program for doing reliability data analysis with fewer capabilities. By the mid-1980s, SAS had incorporated some of the most important models and methods described in Lawless [22], Nelson [37], and Nelson [38] into their general purpose system and these capabilities continue to be developed and extended. More recently, JMP, MINITAB, R, and S-PLUS have also incorporated some of the most widely used methods and models for reliability data analysis into their packages. The following list describes some popular packages commonly found on desktop computers.

JMP(www.jmp.com) is a popular, highly sophisticated general-purpose desktop statistical software package. In addition to such standard statistical tools such as single distribution analysis, regression analysis, and tools for experimental design, JMP also has some special tools for reliability data analysis, including the analysis of censored data, competing risk analysis, accelerated life data, and the analysis of recurrence data. JMP has appealing tools for graphical presentation of data and fitted models. JMP also has a sophisticated scripting language that allows extension of the system.

MINITAB (www.minitab.com) is another popular, general-purpose desktop statistical software package. Its reliability analysis capabilities are similar to those of JMP.

The Reliasoft (www.reliasoft.com) suite of programs does not provide general-purpose statistical capabilities, but rather attempts to cover a broad range of needs

of a reliability engineer. WEIBULL ++ does basic analysis of single distribution data. ALTMA can be used to analyze accelerated life test data. BLOCKSIM provides predictions of system reliability based on evaluation of a system specified by a description of system structure and the reliability of individual components. RG can be used to assess reliability growth of a system.

S-PLUS (www.insightful.com) is a general-purpose, highly sophisticated environment for doing graphics and statistical computing, using the S language (which was developed at Bell Laboratories). One of the important features of S-PLUS is that it is extendable. That is, users can add capabilities (including GUIs) at the same level as the developers of the system. SPLIDA (www.public.iastate.edu/splida) is a free add-on to S-PLUS that has extensive capabilities for planning reliability studies and for analyzing reliability data. Almost all of the SPLIDA capabilities are available through the SPLIDA GUI.

R (www.Rproject.org) is a freeware implementation of the S language having many of the same capabilities of S-PLUS, but with only limited GUI capabilities. It is expected that there will soon be an R version of SPLIDA, but because of the limited GUI capabilities in R, this version of SPLIDA will require the use of S-language commands to operate.

1.11 Use of Bayesian Methods in Reliability

The biggest changes in the future of reliability analysis will come in the form of much more widespread use of Bayesian methods.

There are, in the folklore of reliability, stories of how the use of Bayesian methods in reliability applications has led to unreasonably optimistic predictions for products that had poor reliability. Combined with healthy skepticism about the use of prior information to predict reliability and the previously difficult technical challenges to implement Bayesian methods, applications of Bayesian methods in reliability have been limited. Of course, there are also many well-documented stories about serious reliability disasters where Bayesian methods played no role in the faulty decision making.

As mentioned in Section 1.2, ML is at the core of most statistical methods used for reliability data analysis. Ignoring the philosophical differences in the theory of statistical inference, Bayesian methods can be viewed as an extension of likelihood-based methods, where one combines prior information, in the form of a prior probability distribution, with the likelihood. Bayesian methods have the same appealing characteristics as ML (versatile and good statistical properties), but also have a very important advantage. In particular, Bayesian methods provide a formal method for combining data from various different sources, such as information from different studies, information from reliability tests at different levels of product and component integration, prior information based on previous experience, or prior information from general, but imprecise, engineering knowledge.

Martz and Waller [28] provided an early description of the use of Bayesian methods in reliability applications. The applications were, however, limited because of limitations in the technology (both statistical and computing power). Over the past 25 years, however, there has been an explosion in interest and application of Bayesian methods

in a wide range of areas of application. This enthusiasm has been driven by important advances in methods for implementing the Bayesian paradigm (especially MCMC methods) and important advances in computer hardware capabilities. Ibrahim, Chen, and Sinha [20] is an important reference on the analysis of survival data, taking applications from the areas of medicine and public health. Hamada, Wilson, Reese, and Martz [19] is a valuable recent addition to the reliability literature that treats, from a Bayesian point of view, many of the topics described in this chapter. They describe the basic methodology for each area and illustrate the methods with a wide range of applications. Aven [2], Singpurwalla [43], and Garrick [16] use Bayesian frameworks in risk assessment applications.

There are many particular areas of application where the use of Bayesian methods is particularly compelling. As mentioned in Section 1.3, engineers often assume that some of the needed inputs to their reliability models are known, when in fact, there is always some degree of uncertainty. Use of a prior probability distribution to quantify knowledge would be better than incorrectly assuming that something is known without error.

The example in Chapter 14 of Meeker and Escobar [31] shows how bringing in a little prior information about the Weibull shape parameter (based on previous experience with the same failure mechanism in similar components) provided a much improved and more useful inference on reliability.

For another example, when doing temperature-accelerated life testing (like that described in Section 1.4) to accelerate a chemical reaction leading to failure, there are two vastly different approaches that have been used. The use of a *known* activation energy in electronic component reliability modeling is common, especially in temperature-accelerated testing of microelectronic components. This is, in effect, specifying the slope of the relationship between log lifetime and reciprocal absolute temperature (the Arrhenius model from physical chemistry). For example, MIL-STD-883 [35] specifies testing at only one accelerated level of temperature and requires one to input a value of the activation energy (slope) in order to make reliability predictions, as described on page 282 of Nelson [40]. Of course, the activation energy is not known exactly and assuming that it is known gives a false sense of having little statistical uncertainty (e.g., compare the analyses depicted in Figures 19.13 and 19.15 of Meeker and Escobar [31]). At the other extreme, the standard non-Bayesian approach would use the available data to estimate the activation energy as the slope of a linear relationship between life and reciprocal absolute temperature. Actually, there is, in most applications, useful but imprecise information about activation energy. An appropriate compromise analysis would use a prior distribution to describe the available information about activation energy. Section 22.2 of Meeker and Escobar [31] provides a simple example.

In all of these examples, either of the two extremes (assuming that the parameter is known versus assuming that nothing is known about the parameter) gives wrong answers, especially with respect to statistical uncertainty. Bayesian methods, however, provide an appropriate compromise.

Efron [11] asked “Why isn’t everyone a Bayesian?” If we rephrase the question to “Why isn’t everyone using Bayesian statistical methods in reliability applications?” I would assert that there are two primarily reasons:

- Concerns about the use of prior information.
- Lack of user-friendly software that would make the use of Bayesian methods easy.

Many of us are concerned with the use of Bayesian methods when there is the possibility of having wishful thinking or politically driven opinions masquerading as legitimate prior information. Efron [11] concluded that “The high ground of scientific objectivity has been seized by the frequentists.” Specification of a prior distribution would not be a primary concern if those at risk in a decision-making problem always have the opportunity to specify the prior (or to choose a diffuse prior, indicating lack of information to be used in the analysis). When, however, there are multiple parties at risk and lack of agreement on the prior to use, the objectivity that we seek in using data and statistical methods becomes elusive. *Objective Bayesian* methods (e.g., Berger [3]) that have been developed over the past couple of decades may provide a solution to the technical side of this problem. In effect, the objective Bayesian methods specify a diffuse prior that will result in inferential results that are similar to what one would obtain from a non-Bayesian method. The ideas of objective Bayesian methods can be adapted to allow specification of prior information only for those quantities where there is legitimate prior information (e.g., prior information for which all of the risk-takers have a consensus).

The use of Bayesian methods in reliability will quickly become standard practice at the point in time when there is a software package that has

- Bayesian methods for a wide range of important reliability methods
- Default diffuse prior distributions, but the ability to easily specify legitimate prior information in certain dimensions, when it is available
- A carefully designed graphical user interface that makes the package easy to use and hard to abuse.

1.12 Concluding Remarks

Probability modeling and statistical methods are important tools in the reliability discipline. The new focus, for many manufacturers, in probabilistic design assures that this importance will increase with time. Continuing changes in technology will assure that there will be important opportunities for statisticians to continue to make important research contributions to the discipline.

Acknowledgments. I would like to thank Yili Hong, Katherine Meeker, Mikhail S. Nikulin, and Ying Shi for helpful comments on an earlier version of this chapter.

References

1. Abernethy, R.B., Breneman, J.E., Medlin, C.H., and Reinman, G.L. (1983). *Weibull Analysis Handbook*, Air Force Wright Aeronautical Laboratories Technical Report AFWALTR832079. Available from the National Technical Information Service, Washington, DC.

2. Aven, T. (2003). *Foundations of Risk Analysis*, New York: John Wiley & Sons.
3. Berger, J. (2006). The Case for objective Bayesian analysis, *Bayesian Analysis*, **1**, 385–402.
4. Buckland, W.R. (1964). *Statistical Assessment of the Life Characteristic*, London: Griffin.
5. Condra, L.W. (1993). *Reliability Improvement with Design of Experiments*, New York: Marcel Dekker.
6. Cook, R.J. and Lawless, J.F. (2007). *The Statistical Analysis of Recurrent Events*, New York: Springer.
7. Cox, D.R. and Oakes, D. (1984). *Analysis of Survival Data*, London: Chapman & Hall.
8. Crowder, M.J., Kimber, A.C., Smith, R.L., and Sweeting, T.J. (1991). *Statistical Analysis of Reliability Data*, New York: Chapman & Hall.
9. Crowder, M.J. (2001). *Classical Competing Risks*, New York: Chapman & Hall.
10. David, H.A. and Moeschberger, M.L. (1978). *The Theory of Competing Risks*, London: Griffin.
11. Efron, B. (1986). Why isn't everyone a Bayesian? (with discussion), *The American Statistician*, **40**, 1–11.
12. Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*, New York: Chapman & Hall.
13. Escobar, L.A. and Meeker, W.Q. (1999). Statistical prediction based on censored life data, *Technometrics*, **41**, 113–124.
14. Escobar, L.A., Meeker, W.Q., Kugler, D.L., and Kramer, L.L. (2003). Accelerated Destructive Degradation Tests: Data, Models, and Analysis. In: *Mathematical and Statistical Methods in Reliability*, B. H. Lindqvist and K. A. Doksum, Editors, New Jersey: World Scientific Publishing Company.
15. Flehinger, B., Reiser, B., and Yashchin, E. (1998). Survival with competing risks and masked causes of failure, *Biometrika*, **85**, 151–164.
16. Garrick, B.J. (2008). *Quantifying and Controlling Catastrophic Risks*, Amsterdam: Elsevier.
17. Gertsbakh, I.B. and Kordonsky, K.B. (1969). *Models of Failure, English Translation from the Russian Version*, New York: Springer-Verlag.
18. Haldar, A. and Mahadevan, S. (2000). *Reliability Assessment Using Stochastic Finite Element Analysis*, New York: John Wiley & Sons.
19. Hamada, M.S., Wilson, A.G., Reese, C.S., and Martz, H.F. (2008). *Bayesian Reliability*, New York: Springer.
20. Ibrahim, J.G., Chen, M.H., and Sinha, D. (2001). *Bayesian Survival Analysis*, New York: Springer.
21. Jeng, S.L. and Meeker W.Q. (2000). Comparisons of Weibull distribution approximate confidence intervals procedures for Type I censored data. *Technometrics*, **42**, 135–148.
22. Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime Data*, New York: John Wiley & Sons.
23. Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*, Second Edition, New York: John Wiley & Sons.
24. Lawless, J.F. and Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Analysis*, **10**, 213–227.
25. Lawless, J.F. and Fredette, M. (2005). Frequentist prediction intervals and predictive distributions, *Biometrika*, **92**, 529–542.
26. Lu, C.J. and Meeker, W.Q. (1993). Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, **35**, 161–174.
27. Mann, N.R., Schafer, R.E., and Singpurwalla, N.D. (1974). *Methods for Statistical Analysis of Reliability and Life Data*, New York: John Wiley & Sons.
28. Martz, H.F. and Waller, R.A. (1982). *Bayesian Reliability Analysis*, New York: John Wiley & Sons.

29. Meeker, W.Q. (1987). Limited failure population life tests: application to integrated circuit reliability, *Technometrics*, **29**, 151–165.
30. Meeker, W.Q. and Duke, S.D. (1981). CENSOR-A user-oriented computer program for life data analysis, *The American Statistician*, **35**, 112.
31. Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*, New York: John Wiley and Sons.
32. Meeker, W.Q., Escobar, L.A., and Hong, Y. (2009). Using accelerated life tests results to predict field reliability, *Technometrics*, **51**, 146–161.
33. Meeker, W.Q., Escobar, L.A., and Lu, C.J. (1998). Accelerated degradation tests: modeling and analysis, *Technometrics*, **40**, 89–99.
34. MIL-HDBK-217F (1991). *Reliability Prediction for Electronic Equipment*. Available from Naval Publications and Forms Center, 5801 Tabor Ave, Philadelphia, PA 19120.
35. MIL-STD-883 (1985). *Test Methods and Procedures for Microelectronics*. Available from Naval Publications and Forms Center, 5801 Tabor Ave, Philadelphia, PA 19120.
36. Murray, W.P. (1993). Archival life expectancy of 3M magneto-optic media, *Journal of the Magnetics Society of Japan*, **17**, Supplement S1, 309–314.
37. Nelson, W. (1982). *Applied Life Data Analysis*, New York: John Wiley & Sons.
38. Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*, New York: John Wiley & Sons.
39. Nelson, W. (2003). *Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications*, ASASIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, PA.
40. Nelson, W. (2004). *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*, New York: John Wiley & Sons (updated paperback version of the original 1990 book).
41. Nelson, W. and Hendrickson, R. (1972). *1972 User manual for STATPAC—a general purpose program for data analysis and for fitting statistical models to data*. General Electric CR&D Technical Report 72GEN009.
42. Rausand, M. and Hoyland, A. (2004). *System Reliability Theory: Models and Statistics Methods*, Second Edition, New York: John Wiley & Sons.
43. Singpurwalla, N.D. (2006). *Reliability and Risk: A Bayesian Perspective*, New York: John Wiley & Sons.
44. Strauss, S. (1980). STATPAC: a general purpose program for data analysis and for fitting statistical models to data, *The American Statistician*, **34**, 59–60.
45. Vander Weil, S. and Meeker, W.Q. (1990). Accuracy of approximate confidence bounds using censored Weibull regression data from accelerated life tests, *IEEE Transactions on Reliability*, **39**, 346–351.
46. Wu, H. and Meeker, W.Q. (2002). Early detection of reliability problems using information from warranty data bases, *Technometrics*, **44**, 120–133.

Degradation Processes: An Overview

Mohamed Abdel-Hameed

Department of Statistics, College of Business and Economics, United Arab Emirates University, PO BOX 1755, Al Ain, United Arab Emirates, mohamed@uaeu.ac.ae

Abstract: In this chapter we survey research on different types of degradation processes. Over time, a device is subject to degradation, generally described as an increasing stochastic process. The device has a threshold Y , and it fails once the degradation level exceeds the threshold. Life distribution properties and maintenance policies for such devices are discussed.

Keywords and phrases: Degradation processes, life distributions, increasing failure rate, increasing failure rate average, maintenance and replacement policies

2.1 Introduction

Degradation processes received a great attention over the last decades, in this chapter we review some of the research done on these processes. At the beginning of the work done in reliability, engineers described the uncertainties about the failure times using the survival function; knowing the shape of such a function they can determine and study the properties of the failure rate and based on that they can determine the best possible maintenance policies. To estimate the survival function accurately (from statistical point of view), one has to observe the failure times of many items and these failure random variables are assumed to be independent. In practice, it is not always possible to observe many failures, and even if such failure times are possible to obtain, they are not independent as they all might be effected by a common environment. The other approach is to assess the failure of a device based on the characteristics of the process that caused its failure, normally a degradation process. Such an approach is common in assessing the amount of crack, the amount of erosion, and creep, and amount of contamination. Since many devices fail because of degradation, the degradation process is some type of stochastic process. In this chapter we examine different candidates for such processes. Such candidates include Lévy processes, functionals of Wiener processes, as well as pure jump Markov processes.

In Section 2.2, we give different classes of degradation processes. In Section 2.3, we define classes of life distributions and discuss dependence of random variables. In

Section 2.4, we examine the behavior of the failure rate of devices subject to deterioration and give conditions that insure that the distribution of the failure time belongs to the defined classes of life distributions. In Section 2.5, maintenance and replacement policies of devices subject to degradation are discussed. Detailed proofs of the results given in this chapter are found in the references indicated at the end of this chapter.

2.2 Lévy and Pure Jump Processes

Throughout we let R to be the set of real numbers, $R_+ = [0, \infty)$, N be the set of integers, and $N_+ = \{0, 1, \dots\}$. In this section we will define Lévy process, as well as pure jump Markov processes. We will discuss different examples of Lévy processes and give their corresponding Lévy measures. The relationship between Lévy processes and pure jump Markov processes is also examined.

2.2.1 Lévy processes

A stochastic process $X = \{X_t, t \geq 0\}$ is said to be a Lévy process if the following holds:

- (i) The distribution of $X_{t+s} - X_t$, is independent of t .
- (ii) The process is additive, i.e., for every $t, s \geq 0$, $P\{X_{t+s} - X_t \in A \mid X_u, u \leq t\} = P\{X_{t+s} - X_t \in A\}$.
- (iii) X is stochastically continuous, i.e., for every $t \geq 0$ and $\varepsilon > 0$:
 $\lim_{s \rightarrow t} P\{|X_t - X_s| > \varepsilon\} = 0$.

That is to say a Lévy process is a stochastically continuous process with stationary and independent increments.

If $\Phi(z)$ is the characteristic function of a Lévy process, then $\ln \Phi(z)$ is of the form

$$t \left\{ i z a - \frac{z^2 b}{2} + \int_R [\exp(izx) - 1 - izx I_{\{|x| < 1\}}] \nu(dx) \right\},$$

where $a \in R$, $b \in R_+$, and ν is a measure on R satisfying $\nu(\{0\}) = 0$, $\int_R (1 \wedge x^2) \nu(dx) < \infty$.

Examples of such processes include the following:

- (1) **The Brownian motion.** A Lévy process is said to be a Brownian motion with drift μ , and volatility rate σ^2 , if $\mu = a$, $b = \sigma^2$, and $\nu(R) = 0$.

Another way to define Brownian motion is as follows: for every $t, s \in R_+$, $X_{t+s} - X_t$ has a normal distribution with mean $s\mu$, and a variance $s\sigma^2$.

It follows that the sample paths of the Brownian motion are non-differentiable, continuous and has finite variation on every open interval, almost everywhere. Furthermore, such a process is a strong Markov process. For every $x \in R_+$, we define the time of first passage through the threshold x by

$$T_x = \inf\{t \geq 0 : X_t \geq x\}.$$

For any $\alpha > 0$, we define $\gamma = (\mu^2 + 2\alpha\sigma^2)^{1/2}$. Abdel-Hameed and Nakhi [6] show that, for $y < x$, the Laplace transform of T_x is given by the equation

$$E_y(\exp(-\alpha T_x)) = \exp\{-(\gamma - \mu)(x - y)/\sigma^2\}.$$

Inverting the right-hand side of the above equation with respect to α , we get the probability density function of T_x as follows:

$$f_y(x, t | \mu, \sigma^2) = \frac{(x - y)}{\sqrt{\sigma^2 t^3}} \exp\left\{-\frac{(x - y - \mu t)^2}{2t\sigma^2}\right\}.$$

Although the proof of this formula has been obtained by other authors before, however, our proof is completely different.

For any $t \geq 0$, we let

$$M_t = \sup\{X_s, s \leq t\}.$$

Then,

$$P\{M_t < x\} = P\{T_x > t\}.$$

- (2) **Increasing Lévy Processes.** A Lévy process is said to be increasing if its sample paths are increasing.

Every such process must satisfy the following: $b = 0$, $\nu(-\infty, 0) = 0$, $\int (x \wedge 1) \nu(dx) < \infty$, and $d = a - \int_0^1 x \nu(dx) \geq 0$.

The measure ν characterizes the size and frequency of the jumps. If the measure is infinite, then the process has infinitely many jumps of very small sizes in any small interval. The constant d defined above is called the drift term of the process. If $d = 0$, then the process is a pure jump process that changes from one state to another state only through jump.

We now mention some examples of the Lévy process.

The compound Poisson process.

If the process is a compound Poisson process with jump rate λ , and the distribution of the jump sizes is denoted by F , then $\nu(dx) = \lambda F(dx)$, in this case the Lévy measure is finite. Actually, the finiteness of the Lévy measure characterizes the Poisson process.

Stable processes.

Stable processes are increasing Lévy processes, with Lévy measure given by

$$\nu(dy) = \alpha y^{-(1+\beta)},$$

where the constant α is non-negative, β is between 0 to 1, and $d=0$.

The gamma process.

The gamma process is an increasing Lévy process with Lévy measure

$$\nu(dx) = \delta x^{-1} \exp(-\eta x) dx,$$

where δ and η are non-negative real numbers and $d=0$.

2.2.2 Pure jump Markov processes

A Markov process (X) is called a pure jump process if for each $t \geq 0$

$$X_t = X_0 + \sum_{s \leq t} (X_s - X_{s-}).$$

Cinlar and Jacod [10] show that if X is such a process, then there exists a Poisson random measure N on $R_+ \times R_+$ whose mean measure at the point $(s, z) \in R_+ \times R_+$ is $dsdz/z^2$ and a deterministic function c defined on $R_+ \times (0, \infty)$ that is increasing in the first argument such that

$$\sum_{s \leq t} f(X_{s-}, X_s) = \int_{[0, t] \times R_+} N(ds, dz) f(X_{s-}, X_{s-} + c(X_{s-}, z)) \quad (2.1)$$

almost surely for each function f defined on $R_+ \times R_+$, for which $f(x, x) = 0$ for all x in R_+ . In particular, it follows that

$$X_t = X_0 + \int_{[0, t] \times R_+} N(ds, dz) c(X_{s-}, z).$$

The above formula has the following interpretation:

$t \rightarrow X_t(w)$ jumps at s if the Poisson random measure $N(w, \cdot)$ has an atom at (s, z) and the jump is from the left-hand limit $X_{s-}(w)$ to the right-hand limit: $X_s = X_{s-} + c(X_{s-}, z)$.

We note that an increasing Lévy process is a special case of the pure jump Markov processes with $c(x, z) = z$.

2.3 Life Distributions and Dependence Between Random Variables

In this section we discuss classes of life distributions and different notions of dependence between random variables.

2.3.1 Classes of life distributions

Let X be a non-negative random variable describing the lifetime of a given device. Let F be the distribution function of X and let $\bar{F} = 1 - F$ be the survival function, and cumulative hazard function $R = -\ln \bar{F}$. Then \bar{F} is to have, or to be

- (i) increasing failure rate (IFR) if R is a convex function. The IFR property is equivalent to saying that the failure rate is increasing, whenever it exists.
- (ii) increasing failure rate average (IFRA) if $(1/t)R(t)$ is increasing function in t .
- (iii) new better than used if for every non-negative t, s $\bar{F}(t+s) \leq \bar{F}(t)\bar{F}(s)$.

Assuming that $\bar{F}(0) = 1$, the following implications do hold:

$$\text{IFR} \implies \text{IRA} \implies \text{NBU}.$$

There are dual life distribution classes parallel to the above-mentioned classes and are obtained by reversing the direction of inequality or monotonicity in the above definitions. These classes are the decreasing failure rate (DFR), decreasing failure rate average (DFRA), and new worse than used (NWU) classes. Knowing the behavior of the failure rate of any device enables us to determine the appropriate maintenance and replacement policy for such device.

Excellent treatment of the classes of life distributions can be found in Barlow and Proschan [9].

2.3.2 Dependence between random variables

Components of systems exhibit some degree of dependence between their performances, indicated by their lifetimes. These dependence could be positive, as in the failure times of components subject to the same environment. Negative dependence arises in competing risk applications, where items are competing for a fixed amount of resources. The statistical literature is full with references to different measures of dependence. The simplest is the correlation and partial correlation coefficients. The notion of association between random variables has many applications in reliability and statistics. While there are many other measures of dependence between random variables, we will only discuss a few of them that will be needed in the following sections. One of the strongest notion of dependence is given in the following

Definition. A function $f : R^2 \rightarrow R_+$ is said to be totally positive of order r (TP_r) if $\det(f(x_i, y_j)) \geq 0$ for each choice $x_i \leq x_2 \leq \dots \leq x_k$ and $y_i \leq y_2 \leq \dots \leq y_k$, $1 \leq k \leq r, r \geq 1$.

If f is the joint probability density function of two random variables, then f is TP_r implies that the two random variables, loosely speaking, exhibit a very strong positive dependence. More detailed investigation of this matter can be found in Barlow and Proschan [9].

Definition. A function $f : R \rightarrow R_+$ is said to be a Pólya frequency function of order r (PF_r) if $f(x - y)$ is TP_r in x, y .

2.4 The Degradation Process and Properties of the Resulting Life Distribution

2.4.1 Non-stationary gamma degradation process

Abdel-Hameed [1] discusses the case where the degradation process (X) is a non-stationary gamma process, with transition density function given by, for $x, t \geq 0$

$$p(t, x) = \exp(-\lambda x) (\lambda x)^{\Lambda(t)} / x \Gamma(\Lambda(t)),$$

where Γ is the gamma function, and $\Lambda(t)$ is a non-negative increasing function in its argument. Assume that the device has a resistance level random variable (denoted by Y) with right-tail probability \bar{G} and the process X and the random variable Y

are independent of each other. The device fails when the degradation level crosses its resistance level. Let ρ be the failure time of the device, then

$$\rho = \inf\{t : X_t \geq Y\}$$

which is the time of first crossing of the process X to the random boundary Y . It follows that for $t \geq 0$,

$$\begin{aligned} \bar{F}(t) &= P\{\rho > t\} \\ &= P\{X_t \leq Y\} \\ &= E(\bar{G}(X_t)) \\ &= \int_0^\infty P\{X_t \leq y\}G(dy) \\ &= \int_0^\infty \gamma(\Lambda(t), \lambda y)G(dy)/\gamma(\Lambda(t), \infty) \end{aligned}$$

where $\gamma(t, y)$ is the incomplete gamma function defined as $\int_0^y \exp(-x)x^{t-1}dx$. He shows that (under appropriate conditions on $\Lambda(t)$) life distribution properties of the resistance level Y are inherited as corresponding properties of the failure time ρ . In particular, if the resistance level has increasing failure rate and $\Lambda(t)$ is a convex function, then the failure-time distribution has increasing failure rate as well. As a by-product, if the resistance level is constant and if $\Lambda(t)$ is convex, then the failure-time distribution has increasing failure rate.

2.4.2 Increasing Lévy and pure jump degradation processes

Abdel-Hameed [3] extends the results above for the gamma degradation process, assuming that the degradation process is a non-homogeneous Lévy process. Let μ be the Radon-Nikodym derivatives of the Lévy measure ν . The main results are as follows:

- (a) If G has increasing failure rate, $\Lambda(t)$ is a convex function, and μ is PF₂, then F has increasing failure rate.
- (b) If G has increasing failure rate average, the function $\Lambda(t)/t$ is increasing in its argument, then F has increasing failure rate average.
- (c) If G is NBU, the function $\Lambda(t)$ is super-additive ($\Lambda(t+s) \leq \Lambda(t) + \Lambda(s)$), then F is NBU.

Dual results for the DFR, DFRA, and NWU classes are shown to hold.

Abdel-Hameed [4] extends the above results to include the case where the degradation process is an increasing pure jump process. He finds conditions on the resistance distribution function and the function c in (1) that insure that the distribution of the failure time ρ is IFR, IFRA, NBU, DFR, etc.

2.4.3 Brownian motion like degradation processes

Durham and Padget [11], Park and Padget [12] study the case where the degradation process is a functional of Brownian motion. In the second paper the authors proposed to approximate the distribution function of the degradation process as follows:

$$P\{X_t \leq x\} \approx \Phi\left(\frac{x - \mu t}{\sigma\sqrt{t}}\right)$$

for every $x \geq 0$, where Φ is the standard normal distribution function.

2.5 Maintenance Policies of Devices Subject to Degradation

Assume that a device is subject to degradations and the degradation process is of the gamma type. Abdel-Hameed [2] study the case where the device is replaced at failure (corrective maintenance) or at when the deterioration level exceeds a predetermine level (preventive maintenance). The cost of corrective maintenance is fixed, while the cost of preventive maintenance depends on the deterioration level at the time when the maintenance is performed. He obtains an explicit formula for the long-run average cost per a unit of time. He also discuss a discrete version of the model, where the damage is observed at discrete points in time, and not continuously. Noortwijk [13] applies this result to maintenance of a cylinder on a swing bridge.

Abdel-Hameed [5] treats the optimal inspection policy of a device when the degradation process is an increasing pure jump Markov process, the degradation is monitored periodically. There is a penalty cost of observing the degradation level, and a cost of preventive maintenance as well as the cost of corrective maintenance. In this model, he considers two decision variables, the inspection interval and the preventive maintenance level. A failure is detected only by inspection. He finds the optimal maintenance policy that minimizes the long-run average cost per unit of time.

Abdel-Hameed and Nakhi [7] treat the maintenance policy for devices subject to degradation, when the degradation process is an increasing semi-Markov process. Specifically, let the degradation process (Z) be an increasing semi-Markov process with embedded Markov renewal process $(X, T) = (X_n, T_n; n \in N)$, where $X_n = Z(T_n)$. Let $Q = \{Q(x, A, t), x, t \in R_+, A \subset R_+\}$ be the semi-Markov kernel associated with (X, T) , that is, $Q(x, A, t) = \Pr\{X_{n+1} \in A, T_{n+1} - T_n \leq t | X_n = x\}$, and Markov renewal kernel $R = \{R(x, A, t), x, t \in R_+, A \subset R_+\}$, where $R(x, A, t) = \sum_{n=0}^{\infty} Q^{(n)}(x, A, t)$. The system has a resistance level (denoted by random variable Y), and the device fails once the degradation level crosses the resistance level. The resistance level and the degradation process are assumed to be independent. We denote the failure time by ρ . Let \hat{Z} be the degradation process, obtained by killing the process Z at the failure time, that is, $\hat{Z} = (Z_t, t < \rho)$. Define \hat{Q} and \hat{R} as the corresponding semi-Markov kernel and Markov renewal kernel, respectively. The system can be replaced before or at failure and is maintained continuously. The maintenance and non-failure costs are state dependent. They determine the optimal maintenance policy, using the total discounted as well as the long-run average cost per unit of time. Let $g : R_+ \rightarrow R$ be the function describing the maintenance rate. In the case, where the state space is countable, we define for degradation levels i, j in the state space

$$\begin{aligned} q(i, j) &= P\{X_{n+1} = j | X_n = i\}, \\ m(i) &= E_i(T), \\ \hat{q}(i, j) &= q(i, j) \frac{\bar{G}(j)}{\bar{G}(i)}; \\ \hat{h}(i) &= P_i\{T_1 = \rho\}. \end{aligned}$$

Assume that the costs of a preventative (corrective) maintenance are c_1 and c_2 , ($c_2 > c_1$), respectively, and define the matrix $Q = (q(i, j))$. The optimal replacement policy that minimizes the long-run average cost per a unit time can be summarized in the following algorithm. For more detailed explanations, the reader is referred to the reference above.

Algorithm. Assume that the degradation level at time zero is equal to i , normally taken equal to zero.

Step 1. let $j = i$.

Step 2. Compute the matrix \hat{R} , using the well-known relationship $\hat{R} = [I - \hat{Q}]^{-1}$, where I is the identity matrix of proper dimensions.

Step 3. For i, j let $\hat{r}(i, j) = \hat{R}(i, j) - \hat{R}(i, j - 1)$.

Step 4. Compute

$$b_j(k) = c_2 \left[\frac{m(k)\hat{h}(j)}{m(j)} - \hat{h}(k) \right] + m(k)(g(j) - g(k)),$$

for $k = i, \dots, j$.

Step 5. Compute

$$F(j) = \sum_i^j \hat{r}(i, j)b_j(k).$$

Step 6. If $F(j) \geq c_1$, then j is the optimum replacement level, otherwise $j = j+1$ and go to step 2.

Abdel-Hameed [8] considers the optimal maintenance policy for a system subject to degradation. The degradation level is only observed at successive inspection times. It follows that the degradation levels at inspections and the times of successive inspections form a Markov renewal process. Failure is detected only by inspection; at this point in time the system goes through a corrective maintenance. The system is also maintained when the degradation exceeds a predetermined level (preventive maintenance). He determines the optimal maintenance policy using both the total discounted and the long-run average cost criteria.

References

1. Abdel-Hameed, M. S.: A gamma wear process, *IEEE Transactions on Reliability*, **24**, 152–153 (1975).
2. Abdel-Hameed, M. S.: Optimal replacement policies for devices subject to a gamma wear process, In *The Theory and Applications of Reliability; with Emphasis on Bayesian and Nonparametric Methods* (Ed., C.P. Tsokos, I.N. Shimi), pp. 397–412, Academic Press, New York (1977).
3. Abdel-Hameed, M. S.: Life distribution of devices subject to a Lévy wear process, *Mathematics of Operations Research*, **9**, 606–614 (1984).
4. Abdel-Hameed, M. S.: Life distribution of devices subject to a pure jump damage process, *Journal of Applied Probability*, **21**, 816–825 (1984).
5. Abdel-Hameed, M. S.: Inspection and maintenance policies for devices subject to deterioration, *Advances in Applied Probability*, **19**, 917–931 (1987).

6. Abdel-Hameed, M. S. and Nakhi, Y.: Optimal Control of a finite dam using $P_{\lambda, \tau}^M$ policies and penalty cost: total discounted and long run average cost cases, *Journal of Applied Probability*, **28**, 888–898 (1990).
7. Abdel-Hameed, M. S. and Nakhi, Y.: Optimal replacement and maintenance of systems subject to semi-Markov damage, *Stochastic Processes and Their Applications*, **37**, 141–160 (1991).
8. Abdel-Hameed, M. S.: Optimal predictive maintenance policies for a deteriorating system: The total discounted and long-run average cost cases, *Communications in Statistics: Theory and Methods*, **33**, 735–745 (2004).
9. Barlow, R. and Proschan, F.: *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston, New York (1975).
10. Cinlar, E. and Jacod, J.: Representation of semimartingales Markov processes in terms of Wiener Processes and Poisson random measure, In *Seminars on Stochastic Processes* (Ed., E. Cinlar, K.L. Chung and R.K. Gettoor), pp. 159–242, Birkhauser, Boston (1981).
11. Durham, S. D. and Padgett, W. J.: A cumulative damage model for system failure with applications to carbon and fiber composites, *Technometrics*, **39**, 34–44 (1977).
12. Park, C. and Padgett, W. J.: Accelerated degradation models for failure based on geometric Brownian motion and gamma processes, *Lifetime Data Analysis*, **11**, 511–527 (2005).
13. van Noortwijk, J. M.: Optimal replacement decisions for structures under stochastic deterioration, In *Proceedings of the Eighth IFIP WG. 7.5 working conference on Reliability and Optimization of Structural Systems* (Ed., A. S. Nowak), pp. 273–280, University of Michigan, Michigan (1998).

Defect Initiation, Growth, and Failure – A General Statistical Model and Data Analyses

Wayne B. Nelson

Consultant, Schenectady, NY, wnconsult@aol.com

Abstract: This chapter describes a versatile new model for defect initiation and growth leading to specimen failure. This model readily yields (1) the distribution of time to defect initiation and (2) the distribution of time to failure (when a specified defect size is reached). The model can be readily fitted to defect size and failure data from lab tests or actual service, when defect size is observed once for each specimen. This can be done with software for fitting regression models to censored data. The model and fitting methods are illustrated with an application to dendrite growth on circuit boards. The model and data analyses are also suited to non-engineering applications, for example, initiation and growth of tumors.

Keywords and phrases: Defined failure, censored data, confidence limits, defect initiation, defect failure, defect growth, degradation model, initiation time distribution, maximum likelihood fitting, regression model, residuals

3.1 Introduction

Purpose. In some products, defects initiate and grow over time. Time to defect initiation is random, and defect growth over time is stochastic. In some applications, the product fails when the defect reaches a certain size. This chapter documents a new model and methods for fitting it to data on defect initiation, growth, and failure. These results have appeared only in a talk of Nelson [4].

Applications. Examples of applications of the author are as follows:

- In a test of automotive paint on metal, blisters initiate and grow over time.
- On a circuit board, dendrites initiate and grow between two parallel conductors. When a dendrite reaches from one conductor to the other, it causes a short and circuit malfunction.
- Cracks initiate and grow in a noncritical component of a jet engine. When a crack reaches a specified size, the component has a defined “failure” and should be replaced.

- Cracks initiate and grow in the container wall for a catalytic converter on cars. The container fails when a piece of wall falls out and beads of catalyst escape.

In such applications, one needs a model to understand, predict, and manage such defect initiation, growth, and failure. In particular, one seeks

- to estimate model parameters and functions of them;
- to estimate the cumulative distribution $G(t)$ of time t to defect initiation; and
- to estimate the cumulative distribution $H(t)$ of time t to reach a specified defect size (resulting in a defined or a catastrophic failure). $H(t)$ is used to predict future numbers of failures during (1) warranty, (2) design life, or (3) a specified calendar period. It is also used to determine when to use preventive replacement to avoid failures in service.

Overview. The contents of this chapter are as follows:

- Section 3.2 presents an application and data that motivated this work.
- Section 3.3 motivates and describes a basic model, which provides desired information. This new degradation model does not appear in Meeker and Escobar [1] or Nelson [5, Chapter 11]; they present only statistical models for product degradation which lack a defect-free initiation period.
- Section 3.4 describes how to fit the model to such data using computer packages that fit regression models to censored data.

In a future article, Nelson [7] extends and generalizes the basic model and data analyses here.

3.2 An Application

Purpose. This section presents a client's application with data on defect initiation and growth. It motivated development of the basic model and model fitting described here. Also, it is used to illustrate the model and fitting it to data.

Application. Hermetically sealed circuits had failures in service. My client wanted to estimate the distribution of time to circuit failure in order to determine a preventive replacement policy to avoid such failures in service. Failure analysis revealed that the circuits failed when a dendrite grew between two parallel copper conductors and produced a short. A dendrite is a copper filament that grows from one conductor to the other due to the electric field between them. To estimate the distribution of time to failure, a sample of circuits of differing ages was opened, and the size of each circuit's biggest dendrite was measured. In some circuits, no dendrite had initiated; so the observed size was zero. Here "size" was the length of the dendrite as a percentage of the distance from one conductor to the other. A dendrite cannot be repeatedly measured over time to monitor its growth, because breaking the hermetic seal alters dendrite growth. Thus we have a single measurement of dendrite size for each circuit. This simplifies the modeling and model fitting, as will be seen.

Table 3.1. Dendrite age and size data

Circuit	Age weeks	Dendrite size %
1	171	48
2	81	0
3	74	0
4	180	54
5	198	57
6	149	34
7	180	40
8	159	43
9	92	0
10	141	15
11	140	32
12	99	0

Data. Table 3.1 displays data on 12 circuits. For each circuit, we have its age when it was opened and the size of its dendrite as a percentage of the distance between the conductors. The four circuits with a size of 0 had not yet initiated a dendrite.

Data display. Figure 3.1 displays the data in a crossplot. The vertical axis corresponds to the size of a circuit’s dendrite as a percentage of the gap between the parallel conductors. The horizontal axis corresponds to its age when it was opened and its largest dendrite was measured. A circuit that has not initiated a dendrite has an X (size 0%) on the time axis.

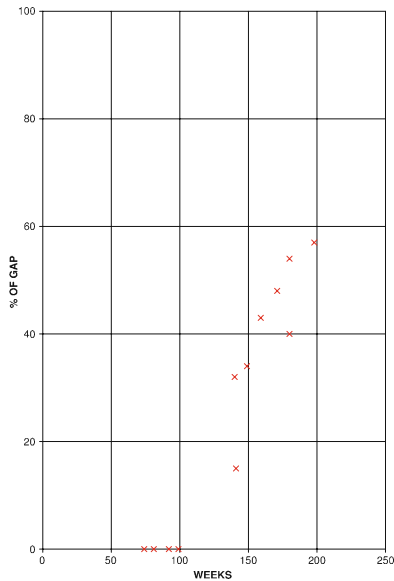


Figure 3.1. Display of dendrite size data

Desired information. The client wanted an estimate of the distribution of time to circuit failure, that is, when dendrite length reaches 100%. This distribution was used to determine when to preventively replace circuits to avoid failure in service.

3.3 A Population Model for Defect Initiation and Growth

Purpose. This section describes a population model for defect initiation and growth data. First, two applications are presented to motivate the model. Then the model and its properties are presented.

Paint application. Dr. Jon Martin of NIST observed blisters on small sheets of metal coated with an automotive paint. The specimens were subjected to high temperature and humidity to hasten initiation and growth of blisters, which were round. Each specimen had an initial blister-free period; then blisters initiated and grew. For specimen i , the diameter of its largest blister at time t was the defect size $Y_i(t)$. Each specimen's largest blister was repeatedly measured over time. Figure 3.2 depicts the size functions $Y_i(t)$ for 11 such specimens. Here blister size increases essentially linearly with time after initiation. Also, all specimen $Y_i(t)$ plots have the same slope; this should not be expected in other applications. Note that for an age t , there is a vertical population density $f_0(y; t)$ of defect size Y in Figure 3.2. The density should be imagined perpendicular to the plane of the figure. Such a distribution consists of (1) a continuous distribution for the sizes of observed defects at time t and of (2) a discrete spike of probability for the population fraction still with no defect. Because the $Y_i(t)$ are parallel here, a paint specimen that is at the P th percentile of the size distribution at time t is also at the P th percentile at a later time t' . Dr. Martin sought

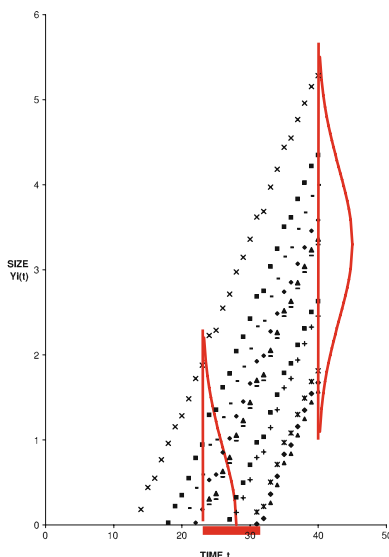


Figure 3.2. Blister size $Y_i(t)$ over time t

to model such blister growth and estimate model parameters. He was most interested in estimating the distribution of time when such blisters reach a specified diameter, a defined failure.

Dendrite application. For the circuits, it was anticipated that the dendrite growth, if it could have been continuously observed, was more like that depicted in Figure 3.3. This figure is speculative, because the dendrite size could not be repeatedly measured. But Figure 3.3 appears reasonable. Here the dendrite size $Y_i(t)$ curves of different units are neither linear nor parallel. A unit that is at the P th percentile of the size distribution at time t is not necessarily at the P th percentile at a later time t' . That is, the $Y_i(t)$ curves may cross. Nevertheless, at a time t , there is a vertical probability density of defect size that is part continuous and part discrete as shown in Figure 3.3. The mixed continuous-discrete densities in Figures 3.2 and 3.3 are similar, although the behavior of the $Y_i(t)$ curves differs in the two figures. In both applications, we regard defect size $Y_i(t)$ of population unit i as an unspecified stochastic process. In such applications, $Y_i(t) = 0$ up to the initiation time T_i and then increases monotonically. For example, cracks and dendrites do not grow shorter.

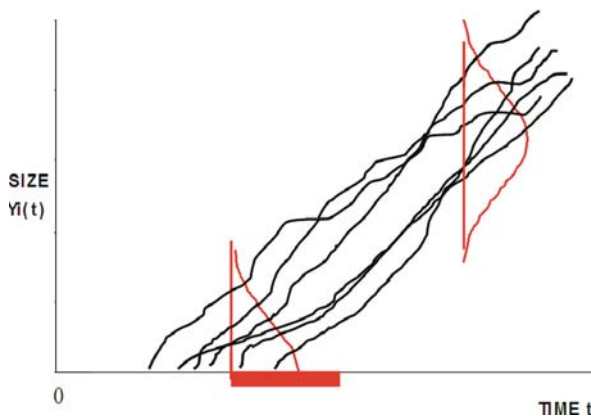


Figure 3.3. Dendrite size $Y_i(t)$ over time t

Size distribution. As shown in Figures 3.2 and 3.3, at any age t , there is a population probability density $f_0(y; t)$ for defect size Y . Such probability densities $f_0(y; t)$ should be regarded as perpendicular to the plane of a figure. Let $F(y; t)$, $-\infty < y < \infty$, $t > 0$, be a potential underlying cumulative distribution for size Y . The mixed continuous-discrete population cdf $F_0(y; t)$ is then modeled as

$$F_0(y; t) = \begin{cases} F(y; t) & \text{for } y > 0, \\ F(0; t) & \text{for } y = 0. \end{cases} \quad (3.1)$$

Figures 3.2 and 3.3 depict the probability density $f_0(y; t)$, and Figure 3.4 depicts the probability density $f(y; t)$. Thus, at time t , the area of the density $f(y; t)$ below 0 in Figure 3.4 equals the corresponding spike of probability $F_0(0)$ in Figures 3.2, 3.3, and 3.4.

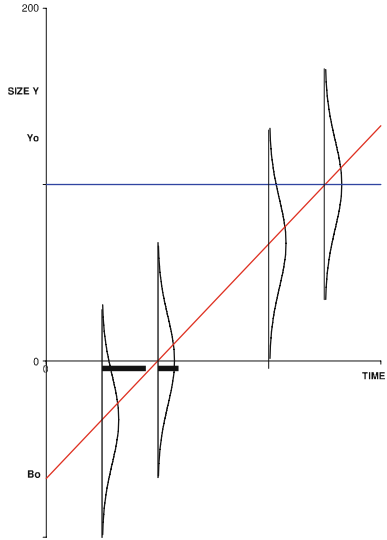


Figure 3.4. Basic model with areas below 0 for the population fraction not initiated

Relationships. The basic model for the underlying cumulative distribution $F(y; t)$ of size Y at age t has a location parameter $\mu(t)$, which depends on t , and a constant scale parameter σ . In practice regression functions $\mu(t)$ and $\sigma(t)$ are ideally based on physical theory for defect growth, such as Paris' [8] equation for crack growth in metal subjected to cyclic fatigue. For concreteness, we use

$$\mu(t) = \gamma_0 + \gamma_1 t. \quad (3.2)$$

The γ_j coefficients and σ are to be estimated from data. In this regression relationship, size Y is the dependent variable and time t is the independent variable. This linear relationship was thought to be reasonable for dendrite growth, because the growing end of a dendrite experiences a constant electric field which likely makes it grow at the same rate for any length. In applications, the $Y_i(t)$ increase; so $\mu(t)$ must increase. Thus γ_1 must be positive. γ_0 is the negative intercept of (3.2) in Figure 3.4.

Regression model. Also, for concreteness, the $F(y; t)$ we use is a normal cdf where $\mu(t)$ and σ are its mean and standard deviation. That is, we use the linear-normal regression model

$$F(y; t) = \Phi\{[y - \mu(t)]/\sigma\}, \quad (3.3)$$

where $\Phi\{\}$ is the standard normal cdf. This basic regression model is depicted in Figure 3.4 with its probability densities $f(y; t)$. For this model, the fraction of the size Y distribution below 0 at time t is

$$F(0; t) = \Phi\{[0 - \mu(t)]/\sigma\} = \Phi\{-(\gamma_0 + \gamma_1 t)/\sigma\}. \quad (3.4)$$

This is the lower tail area of $f(y; t)$ below 0 in Figure 3.4. Here $F(0; t)$ is interpreted as the fraction of the population that has not initiated a defect. $F(0; t)$ is the discrete spike of probability in Figures 3.2 and 3.3. In a future article, Nelson [7] generalizes

this basic model to other size distributions and other time functions of the parameters $\mu(t)$ and $\sigma(t)$.

Percentiles. For the underlying basic model (3.3), the F fractile $\eta_F(t)$ of the size distribution as a function of time is

$$\eta_F(t) = \mu(t) + z_F \sigma = \gamma_0 + \gamma_1 t + z_F \sigma, \quad (3.5)$$

where z_F is the standard normal F fractile. In Figure 3.4, such percentile relationships would be straight parallel lines.

Measurement error. Note that this basic model does not include measurement error for size. That is, measurement error is assumed negligible here. A more complicated model with measurement error appears in Nelson [7]. Also, note that the properties of the stochastic process $Y_i(t)$ such as autocorrelation are ignored in this model, since they are not needed and not estimable when there is just one size-age observation (Y_i, t_i) for sample unit i .

Initiation time distribution. The initiation time distribution is as follows. The basic regression model is depicted in Figure 3.5. There $F(0; t)$ is the population fraction that has not initiated a defect by age t . Consequently, the population fraction $G(t)$ that has initiated a defect by time t is

$$G(t) = 1 - F(0; t). \quad (3.6)$$

Then for the basic regression model, the cdf of time to defect initiation is

$$G(t) = 1 - \Phi\{[0 - \mu(t)]/\sigma\} = \Phi\{[t - (-\gamma_0/\gamma_1)]/(\sigma/\gamma_1)\}. \quad (3.7)$$

This distribution is normal with mean and standard deviation

$$\mu_G = -\gamma_0/\gamma_1 \quad \text{and} \quad \sigma_G = \sigma/\gamma_1. \quad (3.8)$$

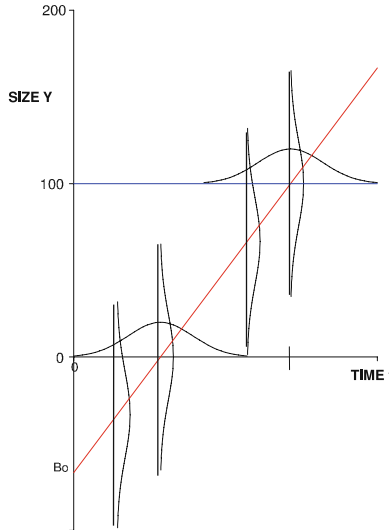


Figure 3.5. Basic model with initiation time and failure time distributions

Thus the F fractile of this distribution is

$$\tau_F = \mu_G + z_F \sigma_G = (-\gamma_0/\gamma_1) + z_F(\sigma/\gamma_1). \quad (3.9)$$

In practice, one must use data to get estimates of the model parameters γ_0, γ_1 , and σ and use them to estimate μ_G and σ_G .

Failure time distribution. The failure time distribution follows. Suppose y_0 is the defect size corresponding to failure. In Figure 3.5, $F(y_0; t)$ is interpreted as the population fraction that has not failed by age t ; this is the lower tail area of $f(y; t)$ below y_0 . Consequently, the population fraction failed $H(t)$ by time t is the population fraction with size greater than y_0 , namely,

$$H(t) = 1 - F(y_0; t); \quad (3.10)$$

this is the cdf of time to failure. Then, for the basic regression model,

$$H(t) = \Phi\{[t - ((y_0 - \gamma_0)/\gamma_1)]/(\sigma/\gamma_1)\}. \quad (3.11)$$

This time to failure distribution is normal with mean and standard deviation

$$\mu_H = (y_0 - \gamma_0)/\gamma_1 \quad \text{and} \quad \sigma_H = \sigma/\gamma_1. \quad (3.12)$$

Thus the F fractile of this distribution is

$$\tau_F = \mu_H + z_F \sigma_H = [(y_0 - \gamma_0)/\gamma_1] + z_F(\sigma/\gamma_1). \quad (3.13)$$

$H(t)$ is the distribution of time to circuit failure that the client wanted to know. In practice, one must get estimates of the model parameters γ_0, γ_1 , and σ from data and estimates of μ_H and σ_H from them.

Initiated and failed. With this model, the population fraction that has already initiated a defect at time $t = 0$ is

$$G(0) = \Phi\{(\gamma_0/\gamma_1)/(\sigma/\gamma_1)\} = \Phi\{\gamma_1/\sigma\}.$$

Similarly, the population fraction that has already failed at time $t = 0$ is

$$H(0) = \Phi\{[-(y_0 - \gamma_0)/\gamma_1]/(\sigma/\gamma_1)\} = \Phi\{-(y_0 - \gamma_0)/\sigma\}.$$

In practice, both these fractions should be negligible, as units are usually initially defect free.

3.4 Model Fitting by Computer

Purpose. This section describes how defect size data (Y_i, t_i) can be used to

- (1) estimate the distribution of time to defect initiation,
- (2) estimate the distribution of time to failure, namely, the time when a unit's defect size reaches a specified size y_0 , which may be a defined or catastrophic failure,

- (3) estimate the parameters of the basic model, which can be used to estimate distributions (1) and (2) more accurately.

Data. Data come from n units where some have a defect and others may not. The dendrite data in Table 3.1 are an example of such data. Then the (size,age) data on those with a defect can be expressed as $(Y_1, t_1), (Y_2, t_2), \dots, (Y_r, t_r)$, where r is the random number of units with a defect. Data on those without a defect are expressed as $(0, t_{r+1}), (0, t_{r+2}), \dots, (0, t_n)$.

Software and theory. Software for maximum likelihood (ML) fitting of regression models to censored data can easily be used to fit the basic regression model to defect size data with just one size observation Y_i per sample unit i . Dendrites were observed just once, as opening a hermetic circuit altered its dendrite growth. With one observation per unit, the autocorrelation structure of the $Y_i(t)$ process cannot and need not be modeled. Nelson [5, Chapter 5] describes such commercially available software, which all provide ML fitting of the basic model and more general ones given by Nelson [7]. Nelson [5, Chapter 5] presents the ML theory for fitting the basic model and Nelson's [7] more general models to such data.

Fitted model. Table 3.2 shows the ML estimates of the model parameters and their approximate confidence limits for the dendrite data, obtained with the Minitab [2] software by Dr Scott Kowalski. These were used to obtain parameter estimates for the normal distributions of time to initiation and to failure in Table 3.2; their confidence limits are not provided by most software. However, Nelson [5, 6] provides theory for such approximate limits based on the propagation of error method.

Table 3.2. ML estimates of parameters and confidence limits for dendrite data

Parameter	Estimate	95% Confidence limits	
		Lower	Upper
γ_0 (weeks)	-69.05	-95.39	-42.70
γ_1 (%/week)	0.6605	0.4971	0.8240
σ (%)	6.651	4.073	10.86
μ_G (weeks)	104.5	—	—
σ_G (weeks)	10.1	—	—
μ_H (weeks)	255.9	—	—
σ_H (weeks)	10.1	—	—

Assumptions. The ML fitting method depends on the usual assumptions, such as

- (1) The n observed units are statistically independent,
- (2) The observed Y_i values are statistically independent of their observation times t_i . Better said Y_i is a random observation from the continuous-discrete size distribution $F_0(y; t_i)$.

Observed and censored. For purposes of fitting the basic model above to such data, each unit with a defect is treated as an observed value (Y_i, t_i) . For each unit without a defect, its size is treated as censored on the left at 0, that is, as the value

$(0^-, t_i)$. Here the trailing superscript minus sign “ $-$ ” indicates that the unit’s size is censored on the left. Such fitting is done with the ML method by various software packages. They provide estimates of model parameters and functions of them, such as $G(t)$ and $H(t)$, and corresponding approximate confidence limits.

Initiation distribution. One could directly and less accurately estimate the distribution of time to defect initiation as follows. The data above provide quantal-response data on initiation times. Suppose that t_1, t_2, \dots, t_r are the ages of units with observed defect sizes greater than zero. The initiation times of those units occurred earlier. That is, those initiation times are left censored at the observed ages. Suppose that $t_{r+1}, t_{r+2}, \dots, t_n$ are the observed ages of units that have not yet initiated a defect. Those units will initiate a defect later. Thus those initiation times are right censored at the observed ages. Such left and right censored data are called quantal-response data. Nelson [5, Chapter 5], [6, Chapter 9] describes software that fits a distribution to such data. Of course, this analysis ignores the observed sizes Y_1, Y_2, \dots, Y_n , and consequently its estimates are less accurate than those from fitting the basic regression model, provided it is satisfactory.

Failure distribution. Similarly, one could directly and less accurately estimate the distribution of time to failure, using software described by Nelson [5, Chapter 5], [6, Chapter 9]. Then there are two situations:

- (1) The q failures are catastrophic when $Y_i(t)$ reaches y_0 , and their exactly observed times are t_1, t_2, \dots, t_q . The remaining times $t_{q+1}, t_{q+2}, \dots, t_n$ are right censored. For example, the dendrite failures are catastrophic.
- (2) The failures are defined and not catastrophic. Then the observed $Y_i(t)$ values may exceed y_0 . Cracks in jet engine components are of this type. Then ages t_i with Y_i values greater than y_0 are left censoring times for their failure times. Also, the remaining t_i ages are right censoring times for their failure times. These are quantal-response data.

Of course, for both situations, the corresponding estimates of the time to failure distribution ignore the information in the observed sizes Y_1, Y_2, \dots, Y_n , and consequently each such distribution estimate is less accurate than that from fitting the basic regression model, provided it is satisfactory.

Residuals. Such a fitted basic model yields censored residuals, which can be analyzed as usual to assess how well the model fits the data and to check for peculiar data, as described by Nelson [3], [5, Chapter 5]. Here a unit with an observed defect (Y_i, t_i) has the observed residual

$$r_i = Y_i - \mu^*(t_i) = Y_i - (\gamma_0^* + \gamma_1^* t_i). \quad (3.14)$$

Here γ_0^* and γ_1^* are the ML estimates of γ_0 and γ_1 . A unit that has not initiated a defect $(0, t_i)$ has the left censored residual

$$r_i^- = 0^- - \mu^*(t_i) = -(\gamma_0^* + \gamma_1^* t_i)^-. \quad (3.15)$$

Here the trailing superscript minus sign “ $-$ ” denotes a left censored residual. The n observed and left censored residuals can be plotted on normal probability paper and interpreted, as described by Nelson [3, 5, 6] and Meeker and Escobar [1].

Further work. This chapter is only a first step on modeling and analyzing data on defect initiation and growth. The following work would be useful:

- The basic model needs to be extended to more general models, that is, to other regression relationships for $\mu(t)$ and $\sigma(t)$ and other distributions for $F(y; t)$. Nelson [7] provides some useful extensions and ML methods for fitting such models to defect size data.
- Hypothesis tests need to be developed for evaluating goodness of fit of the basic and more general models.
- In practice, it may be possible to choose the ages t_1, t_2, \dots, t_n of the sample units. Then one seeks ages (a test plan) that in some sense optimizes an estimate of interest. Work on such test plans would provide more accurate estimates.

Acknowledgments. The author thanks Profs. Mikhail S. Nikulin, N. (Bala) Balakrishnan, and Nikolaos Limnios for their encouragement to contribute this work to this book and for their help during the review and publication process. Dr. Scott Kowalski of Minitab Inc. kindly provided the fit of the model to the data, using Minitab.

Concluding remarks. This basic model and the ML fitting method are useful and easy to use in practice. Suitable physical models (distributions and relationships $\mu(t)$ and $\sigma(t)$) need to be developed for specific applications.

References

1. Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*, Wiley, New York.
2. Minitab Inc. (2006). *Reliability Analysis II Training Manual*, Minitab 15 Statistical Software, Minitab Inc., www.minitab.com.
3. Nelson, W. (1973). Analysis of residuals from censored data, *Technometrics*, **15**, 697–715.
4. Nelson, W. (1995). Defect Initiation and Growth – A General Statistical Model and Data Analyses, talk at the Second Annual Spring Research Conf. on Statistics in Industry and Technology, Univ. of Waterloo.
5. Nelson, W.B. (2004a). *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*, paperback edition, 601pp, Wiley, New York.
6. Nelson, W.B. (2004b). *Applied Life Data Analysis*, paperback edition, 634pp., Wiley, New York.
7. Nelson, W.B. (2010). Extensions of a General Statistical Model and Data Analyses for Defect Initiation, Growth, and Failure, in preparation. Request from WNconsult@aol.com.
8. Paris, P.C., Gomez, M.P., and Anderson, W.E. (1961). A rational analytic theory of fatigue, *The Trend in Engineering*, **13**, 9–14.

Properties of Lifetime Estimators Based on Warranty Data Consisting only of Failures

Kazuyuki Suzuki, Watalu Yamamoto, Takashi Hara, and Md. Mesbahul Alam

Department of Systems Engineering, University of Electro-Communications, Tokyo, Japan

Abstract: Nowadays, many consumer durable goods, such as, automobiles, appliances, and photocopiers etc., are sold with manufacturer's warranty to insure product quality and reliability. Warranty claims contain a large amount of useful information about reliability of the products, such as, failure times, usage, and failure modes etc. For engineering purposes, usage is more relevant, and hence, modeling usage accumulation is of great interest in reliability analysis using warranty data. Such models are needed to manufacturers to evaluate reliability, predict warranty costs, and to assess design modification and customer satisfaction. Usually, warranty data consists of only failure information, and non-failure information is not obtainable which makes the reliability analysis difficult. The sales data is also important for reliability analysis as it contains time-in-service in calendar timescale for each non-failed product during the warranty plan. This chapter discusses maximum likelihood estimation of lifetime parameters using warranty data along with sales data and examines the precision of the estimators by the asymptotic variances obtained from Fisher Information Matrix. The practical consequence of this finding is that the proposed method produces estimators of the lifetime parameters with good precision for large sales amount.

Keywords and phrases: Weibull distribution, maximum likelihood estimation, lognormal distribution

4.1 Introduction

Professor W.Q. Meeker has been playing a lot of important roles in the research fields of reliability theory and reliability engineering. His contribution to the research in reliability is so vast, especially in accelerated lifetime testing, degradation analysis, and warranty data analysis. This chapter has overcome the difficulty of the estimation problem from warranty data consisting only of failures. It is so nice to present this content in the memorial book for Professor W.Q. Meeker.

Two main sources of data for reliability evaluation are *test data* and *field data*. The former is obtained through a test program in a laboratory. The latter is product performance data collected from the field. Here, *field* means the actual operating environment of a product in the hands of a customer. Field data provides more reliable information

about the lifetime distribution as compared to test data. As it captures actual usage profiles and the combined environmental exposures that are difficult to simulate in the laboratory, manufacturing companies use it to assess field reliability and make comparisons with engineering predictions, to provide information for product modifications, to assess the effect of design changes, to estimate and explain warranty costs, and to aid in the design of warranty plan. When a warranted product fails within the warranty plan and its user makes a claim to the manufacturer for repair or replacement of the product, the claim is known as *warranty claim*. The resulting warranty claims contain field performance data, obtained under actual operating conditions, which manufacturers use to track product lifetimes. As a result, warranty data can be used as a prime source of field reliability data. However, it can be automatically created and updated at no extra cost from warranty claims that manufacturers receive during warranty coverage.

Since it is difficult and costly to track the performance of products sold, the collection and analysis of warranty data thus become more valuable to manufacturing companies than before. However, a class of problems involving incomplete information and unclear nature of warranty data can be identified [14]. Such a database contains only failure information that failed during the warranty period, that is, it does not provide information of censored products (i.e., the products that do not fail during the warranty period). For example, warranty database cannot provide information on accumulated mileage of automobiles that do not fail within a warranty period. The sales data obtained from sales department is also valuable for reliability analysis in that the time-in-service in calendar timescale can be obtained for censored products.

Many factors contribute to product failures that result in warranty claims. One of the important factors is the age (the calendar time) of the product. The age-specific analysis using warranty data has engendered considerable interest in the literature [4–7, 11]. Modeling usage accumulation (e.g., mileage of automobile) is an integral part of reliability analysis since it is more relevant for engineering purposes. Because the usage-based lifetime distributions of the warranty failures are different from that of the censored products, modeling usage accumulation generally requires supplementary data from sources other than warranty database, such as usage accumulation information from customer surveys, follow-up studies, and periodic inspections and others.

Most of the previous researches [2, 3, 8, 10, 12, 13, 15] indicate that such additional data is important to know the censored usage time or usage time distribution. But collection of such information is very costly and sometimes impossible, and its unavailability makes the estimation of lifetime distribution difficult. That is why the reliability researchers have been looking for to have an effective way of estimation without using supplementary information for the last three decades. In this chapter an attempt has been made to deal with this problem and to discuss how to estimate the product lifetime distribution(s) by only using warranty data and sales data. Especially, this chapter deals with the case of two failure modes for lifetime distributions and shows the importance of utilizing the information of other failure modes to improve the precision of the estimators about the focused failure mode. The precision is examined by the asymptotic variances of the estimators obtained from Fisher Information Matrix.

4.2 A Model for Failure Data in Warranty Database

Automobile manufacturers can usually collect the data contained in Table 4.1 from warranty claim data and sales information, $N_t, t = 1, \dots, T$. The accumulated mileage at the time of failure is observed for each automobile that has failed during the warranty coverage as well as the associated failure mode, months in service (MIS), etc. Mileage is more preferable from the engineering aspects, and therefore, mileage-based lifetime estimation using warranty data is of great importance to manufacturer. As already mentioned in Section 4.1, there are large body of literatures on warranty data on calendar time. In this study, two failure modes on mileage are considered, and the properties of the proposed estimators are examined.

Table 4.1. Information available from warranty database

Sales month	Sales amount	Months in service (MIS), t					# failed	
		1	...	t	...	T	mode a	mode b
1	N_T			$r_T^{(a)}$	$r_T^{(b)}$
\vdots
$T - t + 1$	N_t		...				$r_t^{(a)}$	$r_t^{(b)}$
\vdots
T	N_1						$r_1^{(a)}$	$r_1^{(b)}$

$X^{(a)}$, $X^{(b)}$, and Y_t represent the independent, identically distributed set of random variables, with $X^{(a)}$ and $X^{(b)}$ the variables of interest “failure mode (a)” and “failure mode (b),” and Y_t censoring variable. We denote the distribution functions as $X^{(a)} \sim F_a(x^{(a)})$, $X^{(b)} \sim F_b(x^{(b)})$, and $Y_t \sim G_t(y)$, and the density functions as $f_a(\cdot)$, $f_b(\cdot)$, and $g_t(\cdot)$. For example, $X^{(a)}$ and $X^{(b)}$ might be the mileage to the first failure of the failure mode (a) and mode (b), and Y_t might be the total mileage during t months in service within the warranty period. Here t is measured by a calendar time (month).

In warranty database, we assume that observed failures are classified by sales month and months in service. Let $x_{ti}^{(a)}$ be the i th observed failure mileage of failure mode (a) at t MIS ($t = 1, 2, \dots, T$), $x_{tj}^{(b)}$ be the j th observed failure mileage of failure mode (b), and let the censored mileages be unknown. Further let $r_t^{(a)}$ and $r_t^{(b)}$ be the numbers of the observed failures corresponding to failure modes (a) and (b) of which MIS is “ t .” For simplicity of the argument, we assume that all products are sold at the same time in this chapter, i.e., there is a single sales month, and all products have months in service t . The extension to general cases as shown in Table 4.1 is straightforward.

Throughout this chapter, the following assumptions are made on the quantities introduced above:

1. $X^{(a)}$, $X^{(b)}$, and Y_t are independent; the distributional form of them is known.

2. The time scale of $X^{(a)}$, $X^{(b)}$, and Y_t is assumed to be operating time (e.g., mileage, number of copy volumes); whereas the observational period of the study is measured by calendar time (e.g., month, year)
3. The probability for the failure depends only on its operating time.
4. All failures during the warranty period will be reported to the manufacturer. If there is no failure, the owner will not report the mileage in that period. Consequently, “no record of failure” means there has been no failure.

In general, field data during warranty period consists of several failure modes, causes of failures, failed components, etc. For the formulation, we deal with them as “failure modes.” It is important to focus on the most critical failure mode of which lifetime is smallest among all failure modes. To find the critical failure mode, let $X^{(a)}$ describe the lifetime with the failure mode of interest and deal with the other failure modes as $X^{(b)}$ as a whole. For censored data of which mileage is not available from warranty claim records, we can utilize them as $\Pr(X^{(a)} > Y_t, X^{(b)} > Y_t)$ for the likelihood.

Rai and Singh [10] demonstrated that a lognormal distribution provides a good fit for mileage accumulation data with a linear trend between mileage at t MIS, Y_t versus MIS t . These data are drawn from vehicle recall data and can be assumed to be a random sample with respect to mileage and calendar time. They showed that $Y_t \sim \text{LN}(\mu + \log t, \sigma^2)$.

Philips and Sweeting [9] discussed usage-based estimation from only failure data assuming both the lifetime and censoring time variables are exponential. As shown by Suzuki [13], this case is solved analytically. We use the Weibull distribution, which is extensively used for modeling mileage due to wear out failure. $X^{(a)}$ and $X^{(b)}$ are the Weibull random variables with shape parameters m_a and m_b , and scale parameters η_a and η_b , respectively ($X^{(a)} \sim \text{Weibull}(m_a, \eta_a)$, $X^{(b)} \sim \text{Weibull}(m_b, \eta_b)$);

$$F_a(x^{(a)}) = 1 - \exp \left\{ - \left(\frac{x^{(a)}}{\eta_a} \right)^{m_a} \right\}, \quad x^{(a)} \geq 0,$$

$$F_b(x^{(b)}) = 1 - \exp \left\{ - \left(\frac{x^{(b)}}{\eta_b} \right)^{m_b} \right\}, \quad x^{(b)} \geq 0.$$

4.3 Maximum Likelihood Estimation

Maximum likelihood estimation is formulated for the lifetime model with Weibull distributions introduced in Section 4.2, assuming that the mileage distribution is lognormal.

The likelihood function can be obtained as

$$L(\theta) = \prod_{t=1}^T \left[\prod_{i=1}^{r_t^{(a)}} f_a(x_{ti}^{(a)}) \bar{F}_b(x_{ti}^{(a)}) \bar{G}_t(x_{ti}^{(a)}) \prod_{j=1}^{r_t^{(b)}} f_b(x_{tj}^{(b)}) \bar{F}_a(x_{tj}^{(b)}) \bar{G}_t(x_{tj}^{(b)}) \right. \\ \left. \times \left[\int_0^\infty g_t(\tau) \bar{F}_a(\tau) \bar{F}_b(\tau) d\tau \right]^{N_t - r_t^{(a)} - r_t^{(b)}} \right], \quad (4.1)$$

where $\theta = (m_a, \eta_a, m_b, \eta_b, \mu, \sigma)'$ is the vector of parameters, $r_t^{(a)}$ and $r_t^{(b)}$ are the number of failures of the failure modes (a) and (b) at t MIS, and $N_T - r_t^{(a)} - r_t^{(b)}$ is the number of censored units at t MIS. Now we have

$$\bar{G}_t(x_{ti}^{(a)}) = \bar{\Phi} \left(\frac{\log x_{ti}^{(a)} - \mu - \log t}{\sigma} \right),$$

where $\bar{\Phi}(\cdot)$ is the survival function of the $N(0, 1)$ distribution. $\bar{G}_t(x_{tj}^{(b)})$ is given in the similar way. Furthermore, we see that the censored part of the likelihood function (1) can be calculated as

$$\begin{aligned} \Psi_t(\theta) &= \int_0^\infty g_t(\tau) \bar{F}_a(\tau) \bar{F}_b(\tau) d\tau \\ &= \int_{-\infty}^\infty e^{-\eta_a m_a t^{m_a} e^{m_a(\mu + \sqrt{2}\sigma v)} - \eta_b m_b t^{m_b} e^{m_b(\mu + \sqrt{2}\sigma v)}} e^{-v^2} dv, \end{aligned} \quad (4.2)$$

where $\theta = (m_a, \eta_a, m_b, \eta_b, \mu, \sigma)'$, $v = (\log \tau - \mu - \log t)/\sqrt{2}\sigma$. Using the above notations in (4.1) and after some simplification, we obtain the log-likelihood function for a single sales month case with months in service t as

$$\begin{aligned} \log L(\theta) &= \left[r_t^{(a)} \log m_a - r_t^{(a)} m_a \log \eta_a + (m_a - 1) \sum_{i=1}^{r_t^{(a)}} \log x_{ti}^{(a)} - \sum_{i=1}^{r_t^{(a)}} \left(x_{ti}^{(a)} / \eta_a \right)^{m_a} \right. \\ &\quad - \sum_{i=1}^{r_t^{(a)}} \left(x_{ti}^{(a)} / \eta_b \right)^{m_b} + \sum_{i=1}^{r_t^{(a)}} \log \bar{\Phi}_t \left(\frac{\log x_{ti}^{(a)} - \mu - \log t}{\sigma} \right) \\ &\quad + r_t^{(b)} \log m_b - r_t^{(b)} m_b \log \eta_b + (m_b - 1) \sum_{j=1}^{r_t^{(b)}} \log x_{tj}^{(b)} - \sum_{j=1}^{r_t^{(b)}} \left(x_{tj}^{(b)} / \eta_b \right)^{m_b} \\ &\quad - \sum_{j=1}^{r_t^{(b)}} \left(x_{tj}^{(b)} / \eta_a \right)^{m_a} + \sum_{j=1}^{r_t^{(b)}} \log \bar{\Phi}_t \left(\frac{\log x_{tj}^{(b)} - \mu - \log t}{\sigma} \right) \\ &\quad \left. + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) \log \Psi_t(\theta) \right]. \end{aligned} \quad (4.3)$$

As $g_t(\tau) \bar{F}_a(\tau) \bar{F}_b(\tau) d\tau$ and $\partial g_t(\tau) \bar{F}_a(\tau) \bar{F}_b(\tau) d\tau / \partial \theta$, where $\theta = (m_a, \eta_a, m_b, \eta_b, \mu, \sigma)'$, are continuous, the integration and differentiation are interchangeable. With this property, $\Psi_t = \Psi_t(\theta)$ in (4.2) can be differentiated with respect to $m_a, \eta_a, m_b, \eta_b, \mu$, and σ , and hence, the gradients and Hessians of the log-likelihood function (4.3) with respect to the parameters can be obtained.

To find the maximum likelihood estimators of $m_a, \eta_a, m_b, \eta_b, \mu$, and σ , we need to maximize (4.3). However, as no closed form solutions can be obtained, we need to solve it by using an iteration method. To perform a Newton–Raphson iteration, we need to compute the Ψ_t -type integrals. Since Ψ_t in (4.2) cannot be evaluated analytically, the integral must be evaluated numerically. For given values of t and θ , Ψ_t can be numerically evaluated by applying the Gauss–Hermite quadrature method (a brief description is given in Appendix, for details see Abramowitz and Stegun [1], for example). Also, the improper integrals similar to the one in Ψ_t (all the first and second order derivatives of Ψ_t with respect to $m_a, \eta_a, m_b, \eta_b, \mu$, and σ are in the form of (4.2)) that are included

in the gradients and Hessians of log-likelihood function (4.3), can also be numerically computed using the Gauss–Hermite quadrature method. The performance of the estimators depends on how accurately one can compute the improper integrals included in log-likelihood function (4.3), its gradients and Hessians. Hence, for better accuracy, we used the Hermite integration order to be 25 in applying Gauss–Hermite quadrature method. The score functions can then be numerically solved by iteration to find the MLEs of $m_a, \eta_a, m_b, \eta_b, \mu$, and σ .

4.4 Properties of Maximum Likelihood Estimators

With assuming the product is automobile, the properties of the proposed estimators are examined in this section. First, the case where just one failure mode (a) exists is dealt with. Then we investigate the cases with two failure modes. Hereafter, calendar time is measured in years as is in the automobile industry, e.g., 12 months in service = 1 year in service.

4.4.1 Fisher information matrix

We assume that $\hat{\theta} = (\hat{m}_a, \hat{\eta}_a, \hat{m}_b, \hat{\eta}_b, \hat{\mu}, \hat{\sigma})'$ is the maximum likelihood estimator, which maximizes (4.3). Under the large-sample theory, MLE $\hat{\theta} = (\hat{m}_a, \hat{\eta}_a, \hat{m}_b, \hat{\eta}_b, \hat{\mu}, \hat{\sigma})'$ asymptotically follows a normal distribution with mean $\theta = (m_a, \eta_a, m_b, \eta_b, \mu, \sigma)'$ and variance-covariance matrix $I^{-1}(\theta)$, where $I(\theta)$ is the Fisher information matrix, defined as $I = -E[\frac{\partial^2 \log L}{\partial \theta \cdot \partial \theta'}]$, negative of the expected Hessian matrix of the log-likelihood function evaluated at the true parameter set, θ . For details of calculations refer to Appendix.

4.4.2 One failure mode case

We set up the parameter vector $\theta = (m_a, \eta_a, \mu, \sigma)'$ of $X^{(a)}$ and Y_t as shown in Table 4.2. $X^{(a)}$ of set 1 is increasing failure rate (IFR), set 2 is constant failure rate (CFR), and set 3 is decreasing failure rate (DFR). To become the same $E[X^{(a)}](= 100,000 \text{ km})$ for all sets, η_a is changed as shown in Table 4.2. Setting $t = 1$ and $N_1 = 1$, asymptotic variances of the unknown parameters are given in Table 4.3. Tables 4.4, 4.5 and 4.6 represent the changes in the asymptotic variances of $\hat{m}_a, \hat{\eta}_a, \hat{\mu}$, and $\hat{\sigma}$ with change in observational period $t = 1, 3$, and 5 for different shape parameters of $X^{(a)}$, $m_a = 2.0, 1.0$, and 0.7.

4.4.3 Two failure modes case

Now we investigate a case with two failure modes, which consists of a competing risks model from $X^{(a)}$, $X^{(b)}$, and operating time Y_t in months in service t . The parameter vector $\theta = (m_a, \eta_a, m_b, \eta_b, \mu, \sigma)'$ is set as shown in Table 4.7. Setting $t = 1$, $N_1 = 1$, Table 4.8 shows asymptotic variances of $\hat{\theta}$. It is clear that asymptotic variances of $\hat{m}_a, \hat{\eta}_a, \hat{\mu}$, and $\hat{\sigma}$ are drastically improved as compared with Table 4.3. This comes from the information of $X^{(b)}$. Therefore, in the analysis of field data, the utilization of other failure modes other than the focused failure mode is so important.

Table 4.2. Parameter setting for examining properties of estimators (one failure mode)

	Set 1	Set 2	Set 3
m_a	2.0	1.0	0.7
η_a	112,800	100,000	79,000
$E[X_a]$	100,000		
μ	9.0		
σ	0.65		
$E[Y_1]$	10,000		
Pattern of failure rate	IFR	CFR	DFR

Table 4.3. Asymptotic variances of estimators (one failure mode)

	Set 1($m_a = 2.0$)	Set 2($m_a = 1.0$)	Set 3 ($m_a = 0.7$)
Prob. of failure; p_a (%)	1.16	9.29	19.7
$AVar(\hat{m}_a)$	3,480	32.7	5.04
$AVar(\hat{\eta}_a)$	1.36×10^{14}	4.24×10^{12}	8.36×10^{11}
$AVar(\hat{\mu})$	4,290	118	37.9
$AVar(\hat{\sigma})$	180	18.9	11.2

Table 4.4. Changes in the asymptotic variances with change in observational period, $t = 1, 3, 5$ ($m_a = 2.0$)

t	1	3	5
p_a	1.16	8.82	19.7
$AVar(\hat{m}_a)$	3,480	214	60.1
$AVar(\hat{\eta}_a)$	1.36×10^{14}	4.87×10^{12}	9.93×10^{11}
$AVar(\hat{\mu})$	4,290	436	133
$AVar(\hat{\sigma})$	180	70.2	51.3

Table 4.5. Changes in the Asymptotic variances with change in observational period, $t = 1, 3, 5$ ($m_a = 1.0$)

t	1	3	5
p_a	9.29	24.4	36.2
$AVar(\hat{m}_a)$	32.7	11.8	7.22
$AVar(\hat{\eta}_a)$	4.24×10^{12}	8.18×10^{11}	3.53×10^{11}
$AVar(\hat{\mu})$	118	41.6	23.4
$AVar(\hat{\sigma})$	19.0	11.7	10.6

Table 4.6. Changes in the asymptotic variances with change in observational period,
 $t = 1, 3, 5$ ($m_a = 0.7$)

t	1	3	5
p_a	19.7	37.0	47.6
$AVar(\hat{m}_a)$	5.03	2.47	1.77
$AVar(\hat{\eta}_a)$	8.36×10^{11}	2.31×10^{11}	1.24×10^{11}
$AVar(\hat{\mu})$	37.9	17.7	11.8
$AVar(\hat{\sigma})$	11.2	7.95	7.20

Table 4.7. Parameter setting for examining properties of estimators (two failure modes)

	Set1	Set2	Set3
m_a	2.0	1.0	0.7
η_a	112,800	100,000	79,000
$E[X_a]$	100,000		
m_b	1.0		
η_b	100,000		
$E[X_b]$	100,000		
μ	9.0		
σ	0.65		
$E[Y_1]$	10,000		
Pattern of failure rate	IFR	CFR	DFR

Table 4.8. Asymptotic variances of parameters (two failure modes)

	Set1($m_a = 2.0$)	Set2($m_a = 1.0$)	Set3($m_a = 0.7$)
Prob. of failure; p_a (%)	1.00	8.68	18.8
Prob. of failure; p_b (%)	9.22	8.68	7.90
$AVar(\hat{m}_a)$	45.6	2.92	0.56
$AVar(\hat{\eta}_a)$	2.91×10^{11}	8.48×10^{10}	2.85×10^{10}
$AVar(\hat{m}_b)$	1.61×10^{-2}	2.92	3.62
$AVar(\hat{\eta}_b)$	1.74×10^8	8.48×10^{10}	1.10×10^{11}
$AVar(\hat{\mu})$	13.3	10.9	7.47
$AVar(\hat{\sigma})$	6.31	5.90	5.03

4.5 Effect of Sales Amount on Precision

Asymptotic variances in Tables 4.4, 4.5 and 4.6 show the precisions when the sales amount $N=1$. In general N , the variance becomes $1/N$. Figures 4.1 and 4.2 show the effect of N on $\sqrt{AVar(\hat{m}_a/m_a)}$ and $\sqrt{AVar(\hat{\eta}_a/\eta_a)}$ when $m_a = 2.0$, $\eta_a = 112,800(\text{km})$, $\mu = 9.0$, and $\sigma = 0.65$. 95% confidence interval of m_a is given by

$$m_a \pm 1.96\sqrt{AVar(\hat{m}_a)} = m_a \left(1 \pm 1.96\sqrt{AVar\left(\frac{\hat{m}_a}{m_a}\right)} \right).$$

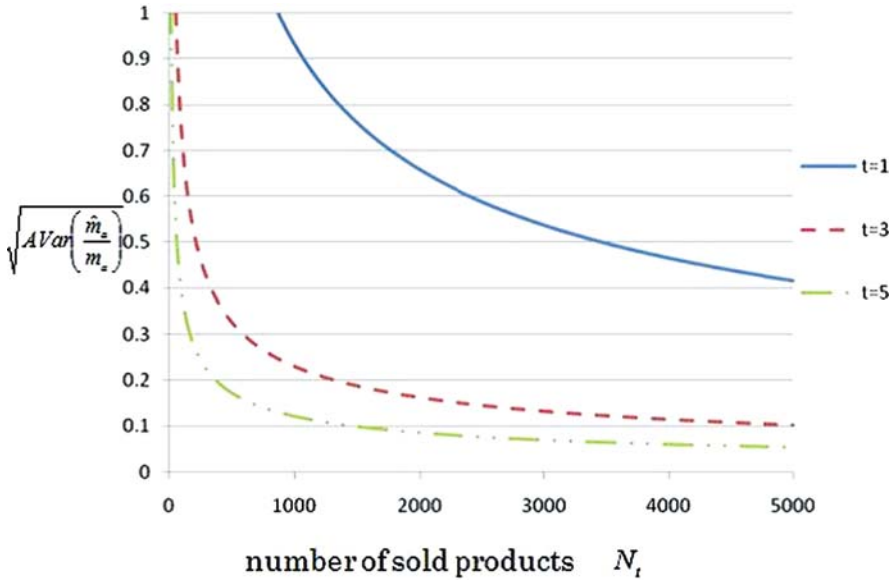


Figure 4.1. Effect of number of sold products N_t and observational period t on $\sqrt{AVar\left(\frac{\hat{m}_a}{m_a}\right)}$

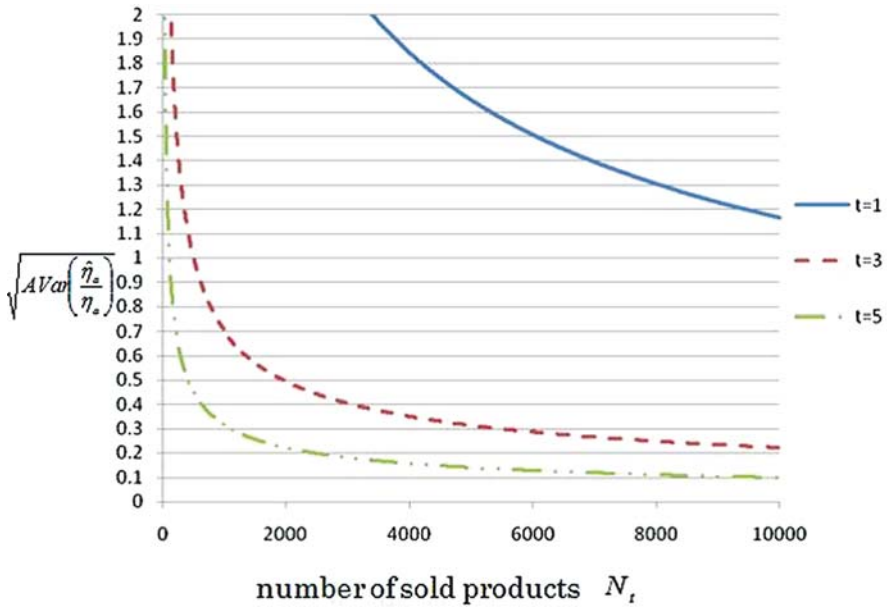


Figure 4.2. Effect of number of sold products N_t and observational period t on $\sqrt{AVar\left(\frac{\hat{\eta}_a}{\eta_a}\right)}$

Using this equation, the necessary number of products is evaluated. For example, to get $\sqrt{AVar(\hat{m}_a/m_a)} = 0.1$, N_t must be larger than 5,000 for $t = 3$, and 1,500 for $t = 5$ from Figure 4.1. In general, there exists more than two failure modes like in the discussion of Section 4.4.3, one gets higher precision than that of Figure 4.1. Besides this, the number of sold products is so large, e.g., 144,051 in the case of Toyota “Corolla” in 2008, that auto manufacturers can easily attain the good precision of the estimators.

4.6 Conclusion

Assuming all failures during the warranty period will be reported to the manufacturer and “no record” means there has been no failure, this chapter discussed the analysis of warranty data and examined the properties of the proposed estimators. If one utilizes the failure information about other failure modes than the focused failure mode, one can drastically improve the precision of the estimators in interest. Also, both the observational period and the number of sold products affect so much on the precision of the estimators. In generally, since the number of sales amounts of automobile is so large, one can estimate the lifetime parameters with sufficient precision.

Appendix

We know that integration and differentiation of a function are interchangeable provided that they exist and the function is continuous. As $g_t(\tau)\bar{F}_a(\tau)\bar{F}_b(\tau)d\tau$ and $\partial g_t(\tau)\bar{F}_a(\tau)\bar{F}_b(\tau)d\tau/\partial\theta$, where $\theta = (m_a, \eta_a, m_b, \eta_b, \mu, \sigma)'$, are continuous, the integration and differentiation are interchangeable. With this property, by differentiating $\Psi_t = \Psi_t(\theta)$ in (4.2) with respect to the parameter vector θ , the first and second order derivatives, and Gradients and Hessians of log-likelihood function (4.3) can be obtained.

A Gradients and Hessians of Log-Likelihood (4.3)

For the simplicity of the presentation of formulas, we denote $\frac{\log x_{ti}^{(a)} - \mu - \log t}{\sigma}$ as $z_{ti}^{(a)}$. $z_{tj}^{(b)}$ is defined similarly.

A.1 Gradients

$$\begin{aligned}
\frac{\partial \log L}{\partial m_a} &= \frac{r_t^{(a)}}{m_a} - r_t^{(a)} \log \eta_a + \sum_{i=1}^{r_t^{(a)}} \log x_{ti}^{(a)} - \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_a} \right)^{m_a} \log \left(\frac{x_{ti}^{(a)}}{\eta_a} \right) \\
&\quad - \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_a} \right)^{m_a} \log \left(\frac{x_{tj}^{(b)}}{\eta_a} \right) + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) \frac{\Psi_{m_a}}{\Psi}, \\
\frac{\partial \log L}{\partial m_b} &= \frac{r_t^{(b)}}{m_b} - r_t^{(b)} \log \eta_b + \sum_{j=1}^{r_t^{(b)}} \log x_{tj}^{(b)} - \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_b} \right)^{m_b} \log \left(\frac{x_{tj}^{(b)}}{\eta_b} \right) \\
&\quad - \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_b} \right)^{m_b} \log \left(\frac{x_{ti}^{(a)}}{\eta_b} \right) + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) \frac{\Psi_{m_b}}{\Psi}, \\
\frac{\partial \log L}{\partial \eta_a} &= -\frac{r_t^{(a)} m_a}{\eta_a} + \frac{m_a}{\eta_a} \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_a} \right)^{m_a} + \left(\frac{m_a}{\eta_a} \right) \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_a} \right)^{m_a} \\
&\quad + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) \frac{\Psi_{\eta_a}}{\Psi}, \\
\frac{\partial \log L}{\partial \eta_b} &= -\frac{r_t^{(b)} m_b}{\eta_b} + \frac{m_b}{\eta_b} \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_b} \right)^{m_b} + \left(\frac{m_b}{\eta_b} \right) \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_b} \right)^{m_b} \\
&\quad + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) \frac{\Psi_{\eta_b}}{\Psi}, \\
\frac{\partial \log L}{\partial \mu} &= \sum_{i=1}^{r_t^{(a)}} \frac{\xi(z_{ti}^{(a)})}{\sigma} + \sum_{j=1}^{r_t^{(b)}} \frac{\xi(z_{tj}^{(b)})}{\sigma} + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) \frac{\Psi_{\mu}}{\Psi}, \\
\frac{\partial \log L}{\partial \sigma} &= \sum_{i=1}^{r_t^{(a)}} \frac{\xi(z_{ti}^{(a)}) z_{ti}^{(a)}}{\sigma} + \sum_{j=1}^{r_t^{(b)}} \frac{\xi(z_{tj}^{(b)}) z_{tj}^{(b)}}{\sigma} + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) \frac{\Psi_{\sigma}}{\Psi}, \tag{4.4}
\end{aligned}$$

where $\Psi_{\theta} := \frac{\partial \Psi}{\partial \theta}$. Here, we omit the suffix gth in Ψ, Ψ_{\bullet} , and $\Psi_{\bullet\bullet}$ throughout the Appendix.

A.2 Hessians

$$\begin{aligned}
\frac{\partial^2 \log L}{\partial m_a^2} &= -\frac{r_t^{(a)}}{m_a^2} - \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_a} \right)^{m_a} \log^2 \left(\frac{x_{ti}^{(a)}}{\eta_a} \right) - \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_a} \right)^{m_a} \log^2 \left(\frac{x_{tj}^{(b)}}{\eta_a} \right) \\
&\quad + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [(\Psi_{m_a m_a} / \Psi) - (\Psi_{m_a} / \Psi)^2], \\
\frac{\partial^2 \log L}{\partial \eta_a^2} &= \frac{r_t^{(a)} m_a}{\eta_a^2} - \frac{m_a(m_a + 1)}{\eta_a^2} \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_a} \right)^{m_a} - \frac{m_a(m_a + 1)}{\eta_a^2} \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_a} \right)^{m_a} \\
&\quad + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [(\Psi_{\eta_a \eta_a} / \Psi) - (\Psi_{\eta_a} / \Psi)^2],
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \log L}{\partial m_b^2} &= -\frac{r_t^{(b)}}{m_b^2} - \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_b} \right)^{m_b} \log^2 \left(\frac{x_{tj}^{(b)}}{\eta_b} \right) - \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_b} \right)^{m_b} \log^2 \left(\frac{x_{ti}^{(a)}}{\eta_b} \right) \\
&\quad + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [(\Psi_{m_b m_b} / \Psi) - (\Psi_{m_b} / \Psi)^2], \\
\frac{\partial^2 \log L}{\partial \eta_b^2} &= \frac{r_t^{(b)} m_b}{\eta_b^2} - \frac{m_b(m_b + 1)}{\eta_b^2} \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_b} \right)^{m_b} - \frac{m_b(m_b + 1)}{\eta_b^2} \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_b} \right)^{m_b} \\
&\quad + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [(\Psi_{\eta_b \eta_b} / \Psi) - (\Psi_{\eta_b} / \Psi)^2], \\
\frac{\partial^2 \log L}{\partial \mu^2} &= -\sum_{i=1}^{r_t^{(a)}} \frac{A(z_{ti}^{(a)})}{\sigma^2} - \sum_{j=1}^{r_t^{(b)}} \frac{A(z_{tj}^{(b)})}{\sigma^2} + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [(\Psi_{\mu\mu} / \Psi) - (\Psi_{\mu} / \Psi)^2] \\
\frac{\partial^2 \log L}{\partial \sigma^2} &= -\sum_{i=1}^{r_t^{(a)}} \frac{C(z_{ti}^{(a)})}{\sigma^2} - \sum_{j=1}^{r_t^{(b)}} \frac{C(z_{tj}^{(b)})}{\sigma^2} + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [(\Psi_{\sigma\sigma} / \Psi) - (\Psi_{\sigma} / \Psi)^2], \\
\frac{\partial^2 \log L}{\partial m_a \partial \eta_a} &= -\frac{r_t^{(a)}}{\eta_a} + \frac{1}{\eta_a} \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_a} \right)^{m_a} + \frac{m_a}{\eta_a} \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_a} \right)^{m_a} \log \left(\frac{x_{ti}^{(a)}}{\eta_a} \right) \\
&\quad + \frac{1}{\eta_a} \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_a} \right)^{m_a} + \frac{m_a}{\eta_a} \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_a} \right)^{m_a} \log \left(\frac{x_{tj}^{(b)}}{\eta_a} \right) \\
&\quad + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{m_a \eta_a} / \Psi - \Psi_{m_a} \Psi_{\eta_a} / \Psi^2], \\
\frac{\partial^2 \log L}{\partial m_a \partial m_b} &= \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{m_a m_b} / \Psi - \Psi_{m_a} \Psi_{m_b} / \Psi^2], \\
\frac{\partial^2 \log L}{\partial m_a \partial \eta_b} &= \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{m_a \eta_b} / \Psi - \Psi_{m_a} \Psi_{\eta_b} / \Psi^2], \\
\frac{\partial^2 \log L}{\partial m_a \partial \mu} &= \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{m_a \mu} / \Psi - \Psi_{m_a} \Psi_{\mu} / \Psi^2], \\
\frac{\partial^2 \log L}{\partial m_a \partial \sigma} &= \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{m_a \sigma} / \Psi - \Psi_{m_a} \Psi_{\sigma} / \Psi^2], \\
\frac{\partial^2 \log L}{\partial \eta_a \partial m_b} &= \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{\eta_a m_b} / \Psi - \Psi_{\eta_a} \Psi_{m_b} / \Psi^2], \\
\frac{\partial^2 \log L}{\partial \eta_a \partial \eta_b} &= \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{\eta_a \eta_b} / \Psi - \Psi_{\eta_a} \Psi_{\eta_b} / \Psi^2], \\
\frac{\partial^2 \log L}{\partial \eta_a \partial \mu} &= \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{\eta_a \mu} / \Psi - \Psi_{\eta_a} \Psi_{\mu} / \Psi^2], \\
\frac{\partial^2 \log L}{\partial \eta_a \partial \sigma} &= \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{\eta_a \sigma} / \Psi - \Psi_{\eta_a} \Psi_{\sigma} / \Psi^2],
\end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \log L}{\partial m_b \partial \eta_b} = & -\frac{r_t^{(b)}}{\eta_b} + \frac{1}{\eta_b} \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_b} \right)^{m_b} + \frac{m_b}{\eta_b} \sum_{j=1}^{r_t^{(b)}} \left(\frac{x_{tj}^{(b)}}{\eta_b} \right)^{m_b} \log \left(\frac{x_{tj}^{(b)}}{\eta_b} \right) + \frac{1}{\eta_b} \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_b} \right)^{m_b} \\ & + \frac{m_b}{\eta_b} \sum_{i=1}^{r_t^{(a)}} \left(\frac{x_{ti}^{(a)}}{\eta_b} \right)^{m_b} \log \left(\frac{x_{ti}^{(a)}}{\eta_b} \right) + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{m_b \eta_b} / \Psi - \Psi_{m_b} \Psi_{\eta_b} / \Psi^2], \end{aligned}$$

$$\frac{\partial^2 \log L}{\partial m_b \partial \mu} = \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{m_b \mu} / \Psi - \Psi_{m_b} \Psi_{\mu} / \Psi^2],$$

$$\frac{\partial^2 \log L}{\partial m_b \partial \sigma} = \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{m_b \sigma} / \Psi - \Psi_{m_b} \Psi_{\sigma} / \Psi^2],$$

$$\frac{\partial^2 \log L}{\partial \eta_b \partial \mu} = \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{\eta_b \mu} / \Psi - \Psi_{\eta_b} \Psi_{\mu} / \Psi^2],$$

$$\frac{\partial^2 \log L}{\partial \eta_b \partial \sigma} = \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{\eta_b \sigma} / \Psi - \Psi_{\eta_b} \Psi_{\sigma} / \Psi^2],$$

$$\frac{\partial^2 \log L}{\partial \mu \partial \sigma} = -\sum_{i=1}^{r_t^{(a)}} \frac{B(z_{ti}^{(a)})}{\sigma^2} - \sum_{j=1}^{r_t^{(b)}} \frac{B(z_{tj}^{(b)})}{\sigma^2} + \left(N_t - r_t^{(a)} - r_t^{(b)} \right) [\Psi_{\mu \sigma} / \Psi - \Psi_{\mu} \Psi_{\sigma} / \Psi^2],$$

where

$$\Psi_{\theta_1 \theta_2} := \frac{\partial^2 \Psi}{\partial \theta_1 \partial \theta_2} \quad (4.6)$$

$$z_t^{(a)} := \left(\log x_t^{(a)} - \mu - \log t \right) / \sigma$$

$$z_t^{(b)} := \left(\log x_t^{(b)} - \mu - \log t \right) / \sigma$$

$$\xi(z) := \phi(z) / \bar{\Phi}(z)$$

$$A(z) := \xi(z) [\xi(z) - z]$$

$$B(z) := \xi(z) + A(z) \cdot z$$

$$C(z) := z[\xi(z) + B(z)]. \quad (4.7)$$

Here, $\phi(\cdot), \bar{\Phi}(\cdot)$ represent the pdf and survival function of the standard normal distribution.

B Expectation of the Elements of Hessians

$$\begin{aligned} E \left[\frac{\partial^2 \log L}{\partial m_a^2} \right] = & -\frac{\rho_t^{(a)}}{m_a^2} - \frac{1}{\eta_a^{m_a}} (E_{a3} - 2 \log \eta_a \cdot E_{a2} + \log^2 \eta_a \cdot E_{a1}) \\ & - \frac{1}{\eta_a^{m_a}} (E_{b3} - 2 \log \eta_a \cdot E_{b2} + \log^2 \eta_a \cdot E_{b1}) \\ & + \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [(\Psi_{m_a m_a} / \Psi) - (\Psi_{m_a} / \Psi)^2], \end{aligned}$$

$$\begin{aligned}
E \left[\frac{\partial^2 \log L}{\partial \eta_a^2} \right] &= \frac{\rho_t^{(a)} m_a}{\eta_a^2} - \frac{m_a(m_a + 1)}{\eta_a^{m_a+2}} E_{a1} - \frac{m_a(m_a + 1)}{\eta_a^{m_a+2}} E_{b1} \\
&\quad + \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [(\Psi_{\eta_a \eta_a} / \Psi) - (\Psi_{\eta_a} / \Psi)^2], \\
E \left[\frac{\partial^2 \log L}{\partial m_b^2} \right] &= -\frac{\rho_t^{(b)}}{m_b^2} - \frac{1}{\eta_b^{m_b}} (E_{b3} - 2 \log \eta_b \cdot E_{b2} + \log^2 \eta_b \cdot E_{b1}) \\
&\quad - \frac{1}{\eta_b^{m_b}} (E_{a3} - 2 \log \eta_b \cdot E_{a2} + \log^2 \eta_b \cdot E_{a1}) \\
&\quad + \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [(\Psi_{m_b m_b} / \Psi) - (\Psi_{m_b} / \Psi)^2], \\
E \left[\frac{\partial^2 \log L}{\partial \eta_b^2} \right] &= \frac{\rho_t^{(b)} m_b}{\eta_b^2} - \frac{m_b(m_b + 1)}{\eta_b^{m_b+2}} E_{b1} - \frac{m_b(m_b + 1)}{\eta_b^{m_b+2}} E_{a1} \\
&\quad + \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [(\Psi_{\eta_b \eta_b} / \Psi) - (\Psi_{\eta_b} / \Psi)^2], \\
E \left[\frac{\partial^2 \log L}{\partial \mu^2} \right] &= -\frac{E_{a4}}{\sigma^2} - \frac{E_{b4}}{\sigma^2} + \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [(\Psi_{\mu \mu} / \Psi) - (\Psi_{\mu} / \Psi)^2], \\
E \left[\frac{\partial^2 \log L}{\partial \sigma^2} \right] &= -\frac{E_{a6}}{\sigma^2} - \frac{E_{b6}}{\sigma^2} + \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [(\Psi_{\sigma \sigma} / \Psi) - (\Psi_{\sigma} / \Psi)^2], \\
E \left[\frac{\partial^2 \log L}{\partial m_a \partial m_b} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{m_a m_b} / \Psi - \Psi_{m_a} \Psi_{m_b} / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial m_a \partial \eta_a} \right] &= -\frac{\rho_t^{(a)}}{\eta_a} + \frac{1}{\eta_a^{m_a+1}} \{m_a E_{a2} + (1 - m_a \log \eta_a) \cdot E_{a1}\} \\
&\quad + \frac{1}{\eta_a^{m_a+1}} \{m_a E_{b2} + (1 - m_a \log \eta_a) \cdot E_{b1}\} \\
&\quad + \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{m_a \eta_a} / \Psi - \Psi_{m_a} \Psi_{\eta_a} / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial m_a \partial \eta_b} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{m_a \eta_b} / \Psi - \Psi_{m_a} \Phi_{\eta_b} / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial m_a \partial \mu} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{m_a \mu} / \Psi - \Psi_{m_a} \Psi_{\mu} / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial m_a \partial \sigma} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{m_a \sigma} / \Psi - \Psi_{m_a} \Psi_{\sigma} / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial \eta_a \partial m_b} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{\eta_a m_b} / \Psi - \Psi_{\eta_a} \Phi_{m_b} / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial \eta_a \partial \eta_b} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{\eta_a \eta_b} / \Psi - \Psi_{\eta_a} \Psi_{\eta_b} / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial \eta_a \partial \mu} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{\eta_a \mu} / \Psi - \Psi_{\eta_a} \Psi_{\mu} / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial \eta_a \partial \sigma} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{\eta_a \sigma} / \Psi - \Psi_{\eta_a} \Psi_{\sigma} / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial m_b \partial \eta_b} \right] &= -\frac{\rho_t^{(b)}}{\eta_b} + \frac{1}{\eta_b^{m_b+1}} \{m_b E_{b2} + (1 - m_b \log \eta_b) \cdot E_{b1}\} \\
&\quad + \frac{1}{\eta_b^{m_b+1}} \{m_b E_{a2} + (1 - m_b \log \eta_b) \cdot E_{a1}\} \\
&\quad + \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{m_b \eta_b} / \Psi - \Psi_{m_b} \Phi_{\eta_b} / \Psi^2],
\end{aligned}$$

$$\begin{aligned}
E \left[\frac{\partial^2 \log L}{\partial m_b \partial \mu} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{m_b \mu} / \Psi - \Psi_{m_b} \Psi_\mu / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial m_b \partial \sigma} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{m_b \sigma} / \Psi - \Psi_{m_b} \Psi_\sigma / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial \eta_b \partial \mu} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{\eta_b \mu} / \Psi - \Psi_{\eta_b} \Psi_\mu / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial \eta_b \partial \sigma} \right] &= \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{\eta_b \sigma} / \Psi - \Psi_{\eta_b} \Psi_\sigma / \Psi^2], \\
E \left[\frac{\partial^2 \log L}{\partial \mu \partial \sigma} \right] &= -\frac{E_{a5}}{\sigma^2} - \frac{E_{b5}}{\sigma^2} + \left(N_t - \rho_t^{(a)} - \rho_t^{(b)} \right) [\Psi_{\mu \sigma} / \Psi - \Psi_\mu \Psi_\sigma / \Psi^2],
\end{aligned}$$

where $E_{a\cdot}$ and $E_{b\cdot}$ are defined as follows:

$$\begin{aligned}
E_{a1} &= E[x_t^{(a)m_a} | X_t^{(a)} < Y_t, X_t^{(a)} < x_t^{(b)}] \\
&= \rho_t^{(a)} \cdot \int_0^\infty x_t^{m_a} \cdot \frac{f_a(x_t^{(a)}) \bar{F}_b(x_t^{(a)}) \bar{G}(x_t^{(a)})}{p_t^{(a)}} dx_t^{(a)}, \\
E_{a2} &= E[x_t^{(a)m_a} \log x_t^{(a)} | X_t^{(a)} < Y_t, X_t^{(a)} < x_t^{(b)}] \\
&= \rho_t^{(a)} \cdot \int_0^\infty x_t^{m_a} \log x_t^{(a)} \cdot \frac{f_a(x_t^{(a)}) \bar{F}_b(x_t^{(a)}) \bar{G}(x_t^{(a)})}{p_t^{(a)}} dx_t^{(a)}, \\
E_{a3} &= E[x_t^{(a)m_a} \log^2 x_t^{(a)} | X_t^{(a)} < Y_t, X_t^{(a)} < x_t^{(b)}] \\
&= \rho_t^{(a)} \cdot \int_0^\infty x_t^{m_a} \log^2 x_t^{(a)} \cdot \frac{f_a(x_t^{(a)}) \bar{F}_b(x_t^{(a)}) \bar{G}(x_t^{(a)})}{p_t^{(a)}} dx_t^{(a)}, \\
E_{a4} &= E[A(z_t^{(a)}) | X_t^{(a)} < Y_t, X_t^{(a)} < x_t^{(b)}] \\
&= \rho_t^{(a)} \cdot \int_0^\infty A(z_t^{(a)}) \cdot \frac{f_a(x_t^{(a)}) \bar{F}_b(x_t^{(a)}) \bar{G}(x_t^{(a)})}{p_t^{(a)}} dx_t^{(a)}, \\
E_{a5} &= E[B(z_t^{(a)}) | X_t^{(a)} < Y_t, X_t^{(a)} < x_t^{(b)}] \\
&= \rho_t^{(a)} \cdot \int_0^\infty B(z_t^{(a)}) \cdot \frac{f_a(x_t^{(a)}) \bar{F}_b(x_t^{(a)}) \bar{G}(x_t^{(a)})}{p_t^{(a)}} dx_t^{(a)}, \\
E_{a6} &= E[C(z_t^{(a)}) | X_t^{(a)} < Y_t, X_t^{(a)} < x_t^{(b)}] \\
&= \rho_t^{(a)} \cdot \int_0^\infty C(z_t^{(a)}) \cdot \frac{f_a(x_t^{(a)}) \bar{F}_b(x_t^{(a)}) \bar{G}(x_t^{(a)})}{p_t^{(a)}} dx_t^{(a)}, \\
E_{b1} &= E[x_t^{(b)m_b} | x_t^{(b)} < Y_t, x_t^{(b)} < X_t^{(a)}] \\
&= \rho_t^{(b)} \cdot \int_0^\infty x_t^{m_b} \cdot \frac{f_b(x_t^{(b)}) \bar{F}_a(x_t^{(b)}) \bar{G}(x_t^{(b)})}{p_t^{(b)}} dx_t^{(b)},
\end{aligned}$$

$$\begin{aligned}
E_{b2} &= E[x_t^{(b)m_b} \log x_t^{(b)} | x_t^{(b)} < Y_t, x_t^{(b)} < X_t^{(a)}] \\
&= \rho_t^{(b)} \cdot \int_0^\infty x_t^{(b)m_b} \log x_t^{(b)} \cdot \frac{f_b(x_t^{(b)}) \bar{F}_a(x_t^{(b)}) \bar{G}(x_t^{(b)})}{p_t^{(b)}} dx_t^{(b)}, \\
E_{b3} &= E[x_t^{(b)m_b} \log^2 x_t^{(b)} | x_t^{(b)} < Y_t, x_t^{(b)} < X_t^{(a)}] \\
&= \rho_t^{(b)} \cdot \int_0^\infty x_t^{(b)m_b} \log^2 x_t^{(b)} \cdot \frac{f_b(x_t^{(b)}) \bar{F}_a(x_t^{(b)}) \bar{G}(x_t^{(b)})}{p_t^{(b)}} dx_t^{(b)}, \\
E_{b4} &= E[A(z_t^{(b)}) | x_t^{(b)} < Y_t, x_t^{(b)} < X_t^{(a)}] \\
&= \rho_t^{(b)} \cdot \int_0^\infty A(z_t^{(b)}) \cdot \frac{f_b(x_t^{(b)}) \bar{F}_a(x_t^{(b)}) \bar{G}(x_t^{(b)})}{p_t^{(b)}} dx_t^{(b)}, \\
E_{b5} &= E[B(z_t^{(b)}) | x_t^{(b)} < Y_t, x_t^{(b)} < X_t^{(a)}] \\
&= \rho_t^{(b)} \cdot \int_0^\infty B(z_t^{(b)}) \cdot \frac{f_b(x_t^{(b)}) \bar{F}_a(x_t^{(b)}) \bar{G}(x_t^{(b)})}{p_t^{(b)}} dx_t^{(b)}, \\
E_{b6} &= E[C(z_t^{(b)}) | x_t^{(b)} < Y_t, x_t^{(b)} < X_t^{(a)}] = \rho_t^{(b)} \cdot \int_0^\infty C(z_t^{(b)}) \cdot \\
&\quad \frac{f_b(x_t^{(b)}) \bar{F}_a(x_t^{(b)}) \bar{G}(x_t^{(b)})}{p_t^{(b)}} dx_t^{(b)},
\end{aligned}$$

and

$$\begin{aligned}
\rho_t^{(a)} &:= p_t^{(a)} \cdot N_t, \\
\rho_t^{(b)} &:= p_t^{(b)} \cdot N_t,
\end{aligned}$$

where $p_t^{(a)}$ and $p_t^{(b)}$ represent probabilities of failure modes (a) and (b), respectively;

$$\begin{aligned}
p_t^{(a)} &:= \int_0^\infty f_a(x) \bar{F}_b(x) \bar{G}_t(x) dx, \\
p_t^{(b)} &:= \int_0^\infty f_b(x) \bar{F}_a(x) \bar{G}_t(x) dx,
\end{aligned}$$

C Gauss–Hermite Quadrature

The numerical integration, also called quadrature, is the study of how the numerical value of an integral can be found. The basic idea for the Gaussian quadrature rule is

$$\int_a^b w(v) dv = \int_a^b p(v) h(v) dv \cong \sum_{j=1}^q h(v_j), \quad (4.8)$$

where $p(v)$ is a weight function and q is a Hermite integration order. Only $h(v)$ needs to be a polynomial or close to polynomial, so that the weight function can be singular. The

weights and nodes of the rule depend on the particular choice of the weight function. If the weight function, $p(v) = \exp(-v^2)$, $-\infty < v < \infty$, (4.8) is called Gauss–Hermite quadrature. By definition the Gauss–Hermite representation of $\Psi_t(\theta)$ in (4.2) can be given by

$$\Psi_t(\theta) \cong \sum_{j=1}^q \frac{p_j}{\sqrt{\pi}} e^{-\eta_a - m_a t^{m_a} e^{m_a(\mu + \sqrt{2}\sigma v_j)} - \eta_b - m_b t^{m_b} e^{m_b(\mu + \sqrt{2}\sigma v_j)}}. \quad (4.9)$$

The weights, p_j , and abscissas, v_j , are tabulated in Table 25.10 of Abramowitz and Stegun [1] for $q \leq 20$. In this study we use $q = 25$ for better accuracy. For this purpose, we use a program to generate p_j , and v_j .

References

1. Abramowitz, M. and Stegun, I. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9 edition. New York: Dover.
2. Davis, T. (1999). A simple method for estimating the joint failure time and failure mileage distribution from automobile warranty data, *Ford Technical Journal*, **2**.
3. Hu, S.J., Lawless, J.F., and Suzuki, K. (1998). Non-parametric estimation of a lifetime distribution when censoring times are missing, *Technometrics*, **40**, 3–13.
4. Kalbfleisch, J.D. and Lawless, J.F. (1996). Statistical analysis of warranty claims data, In: *Product Warranty Handbook*, W.R. Blischke and D.N.P. Murthy (eds.), pp. 231–259, New York: Marcel Dekker.
5. Kalbfleisch, J.D., Lawless, J.F., and Robinson, J.A., (1991). Methods for the analysis and prediction of warranty claims, *Technometrics*, **33**, 273–285.
6. Karim, M.R., Yamamoto, W., and Suzuki, K. (2001). Statistical analysis of marginal count failure data, *Lifetime Data Analysis*, **7**, 173–186.
7. Lawless, J.F. (1998). Statistical analysis of product warranty data, *International Statistical Review*, **66**, 41–60.
8. Lawless, J.F., Hu, X.J., and Cao, J. (1995). Methods for the estimation of failure distributions and rates from auto-mobile warranty data, *Lifetime Data Analysis*, **1**, 227–240.
9. Philips, M.J. and Sweeting, T.J. (2001). Estimation from censored data with incomplete information, *Lifetime Data Analysis*, **7**, 279–288.
10. Rai, B. and Singh, N. (2006). Customer-rush near warranty expiration limit and non-parametric hazard rate estimation from known mileage accumulation rates, *IEEE Transactions on Reliability*, **55**, 480–489.
11. Robinson, J.A. and McDonald, G.C. (1991). Issues related to field reliability and warranty data, In: *Data Quality Control: Theory and Pragmatics*, G.E. Liepins and V.R.R. Uppuluri (eds.), New York: Marcel Dekker.
12. Suzuki, K. (1985a). Nonparametric estimation of lifetime distribution from a record of failures and follow-ups, *Journal of the American Statistical Association*, **80**, 68–72.
13. Suzuki, K. (1985b). Estimation of lifetime parameters from incomplete field data, *Technometrics*, **27**, 263–271.
14. Suzuki, K., Karim, M.R., and Wang, L. (2001). Statistical analysis of reliability warranty data, In: *Handbook of Statistics: Advances in Reliability*, **20**, N. Balakrishnan and C.R. Rao (eds.), pp. 585–609, Amsterdam: Elsevier Science.
15. Suzuki, K. (2004). Analysis of reliability lifetime data, (in Japanese), *Hinshitsu (Journal of Japanese Society for Quality Control)*, **34**, 157–165.

Part II

Shock Models

Shock Models

Allan Gut¹ and Jürg Hüsler²

¹ Department of Mathematics, Uppsala University, Uppsala, Sweden

² Department of Statistics, University of Bern, Bern, Switzerland

Abstract: The standard assumptions in shock models are that the failure (of a system) is related either to the cumulative effect of a (large) number of shocks or that failure is caused by a shock that exceeds a certain critical level. An extension is to consider a mixture, that is, a system breaks down either because of a large shock or as a result of many smaller ones, whichever appears first. In this chapter we survey our results on this problem as well as on a further generalization in which a shock can partly harm the system which implies a lower critical boundary for the following shocks to be fatal. For the cumulative model we also deal with the case in which only the sum of the most recent shocks implies a system failure. In addition we consider the combination of both models with some link functions and briefly discuss an extension of shock models which are based on a Markovian model.

Keywords and phrases: Shock, intershock time, cumulative shock, extreme shock, mixed shock, varying critical load, stopped random walk, renewal theory, Markovian shock model, first passage times, moments, convergence, exact distribution, limit distribution

5.1 Introduction

Shock models are systems that at random times are subject to shocks of random magnitudes. One distinguishes between two major types; cumulative shock models and extreme shock models. Systems governed by the former kind breakdown when the cumulative shock magnitude exceeds some given level, whereas systems modeled by the latter kind breakdown as soon as an individual shock exceeds some given level. For some background and examples, see [21, 23, 2, 3, 6]. Another shock model which is based on a run of k critical shocks was introduced in [20] with $k \geq 1$. If $k = 1$ their model reduces to the extreme shock model.

The aim of this chapter is to present an overview of our joint efforts in the area. We confine ourselves to results related to laws of large numbers and the central and other limit theorems, and refer the reader to the original sources [10, 13, 11, 14] for

further results, such as convergence of moments and laws of the iterated logarithm. Some related applications are discussed in [8, 9].

5.2 Cumulative Shock Models

The general setup in cumulative shock models is a family $\{(X_k, Y_k), k \geq 0\}$ of i.i.d. two-dimensional random vectors, where X_k represents the magnitude of the k th shock and Y_k represents the time between the $(k-1)$ st and the k th shock in model I, and the time between the k th and the $(k+1)$ st shock in model II. The main object of interest is the lifetime/failure time of the system.

In model I we have $X_0 = Y_0 = 0$, and set $T_n = \sum_{k=1}^n Y_k$ and $S_n = \sum_{k=1}^n X_k$, $n \geq 1$, and define the first passage time process $\{\tau(t), t \geq 0\}$ by

$$\tau(t) = \min\{n : S_n > t\}. \quad (5.1)$$

In this model $\tau(t)$ is the number of shocks until failure. The failure time is described by the random variable $T_{\tau(t)}$.

An first observation is that $\tau(t)$ is the first passage time for the random walk $\{S_n, n \geq 1\}$. Therefore, assuming that $EX_1 > 0$, which typically would be the case since shocks in general are nonnegative anyway, “everything” is known for the first passage times as well as for the failure times; see [15, 12]. The following theorem (see [10] for details) therefore follows immediately from the theory of (two-dimensional) stopped random walks in [12], Section 4.2.

Theorem 1. (i) If $\mu_x = EX_1 > 0$ and $\mu_y = EY_1 < \infty$ exist, then

$$\frac{T(t)}{t} \xrightarrow{\text{a.s.}} \frac{\mu_y}{\mu_x} \quad \text{as } t \rightarrow \infty.$$

(ii) If, in addition, $\sigma_x^2 = \text{Var}X_1$ and $\sigma_y^2 = \text{Var}Y_1$ are finite, and $\gamma^2 = \text{Var}(\mu_x Y_1 - \mu_y X_1) > 0$, then

$$\frac{T(t) - \frac{\mu_y}{\mu_x} t}{\sqrt{\mu_x^{-3} \gamma^2 t}} \xrightarrow{d} N(0, 1) \quad \text{as } t \rightarrow \infty.$$

Remark 1. Since the random variables $\{Y_k, k \geq 1\}$ represent time points it is implicitly assumed that $Y_k \geq 0$ a.s. for all k . The results, however, frequently remain true without this assumption.

5.3 Extreme Shock Models

The setup in extreme shock models is the same as before, but in contrast to the cumulative shock model, the failure of the system is caused by one large or extreme shock

which surpasses some given threshold t . The lifetime of the system is, again, $T_{\tau(t)}$ where, however,

$$\tau(t) = \min\{n : X_n > t\}. \quad (5.2)$$

The stopping time in this case thus equals the number of shocks until failure, or, equivalently, the number of the fatal shock. This model was dealt with in [13].

In this setting the stopping times behave much different from the previous ones. The essential new problem is that there is no law of large numbers available for the stopping times. However, the failure time $T_{\tau(t)}$ is still a *stopped random walk*.

An immediate observation is that $\tau(t)$ is geometric with mean $1/p_t$ and variance q_t/p_t^2 , where $p_t = P(X_1 > t)$ and $q_t = 1 - p_t = P(X_1 \leq t)$. A standard computation using, e.g., characteristic functions, therefore shows that

$$p_t \tau(t) \xrightarrow{d} \text{Exp}(1) \quad \text{as } t \rightarrow x_F \quad (5.3)$$

under the natural assumption that

$$p_t \rightarrow 0 \quad \text{as } t \rightarrow x_F, \quad (5.4)$$

where $x_F (\leq \infty)$ denotes the upper endpoint of the distribution F . This fact, together with an application of the law of large numbers for random sums, yields the following result which is [13], Theorem 1.

Theorem 2. *Suppose that (5.4) holds and that $\mu_y = EY_1 < \infty$. Then*

$$p_t T_{\tau(t)} = (1 - F_X(t)) T_{\tau(t)} \xrightarrow{d} \text{Exp}(\mu_y) \quad \text{as } t \rightarrow x_F,$$

or, equivalently,

$$\frac{T_{\tau(t)}}{E T_{\tau(t)}} \xrightarrow{d} \text{Exp}(1) \quad \text{as } t \rightarrow x_F.$$

This has been proved earlier by [21], see their Corollary 1.A.5. However, our proof is simpler and more straightforward.

For details and generalizations to more general failures we refer to our paper [13], where also moment convergence is established.

5.4 Mixed Shock Models

In this model the system breaks down when the cumulative shocks reach “some high” level or when a single “large” shock appears whichever comes first. The number of shocks at the time of failure thus equals $\min\{\nu(t), \tau(t)\}$. The same mixed kind of first passage times, but relative to certain damage processes and more general Markov processes have been considered in [18, 19], however, with a somewhat different focus.

In order to obtain a nontrivial result (that is, in order to avoid that one of the stopping times dominates the other one), the levels to attain must be chosen so that the normalizations in the theorems are of the same order of magnitude. Moreover, with

respect to the cumulative model we must assume that $EX_1 > 0$, and, since high level in that model means “as $t \rightarrow \infty$ ”, it follows that we must also have $x_F = +\infty$.

In order for $\nu(t)$ and $\tau(t)$ to be of the same order of magnitude asymptotically, we let λ_t denote the θ/t -quantile of the distribution F of X_1 for some $\theta > 0$. With

$$\tau_\lambda(t) = \min\{n : X_n > \lambda_t\}, \quad t \geq 0, \quad (5.5)$$

the number of shocks until failure equals

$$\kappa(t) = \min\{\nu(t), \tau_\lambda(t)\} \xrightarrow{a.s.} +\infty \quad \text{as } t \rightarrow \infty. \quad (5.6)$$

The following theorem emerges.

Theorem 3. *If $\mu_x = EX_1 > 0$ and $\mu_y = EY_1 < \infty$ exist, then*

$$\frac{\kappa(t)}{t} \xrightarrow{d} Z \quad \text{as } t \rightarrow \infty,$$

where

$$f_Z(y) = \theta e^{-\theta y}, \quad 0 < y < 1/\mu_x \quad \text{and} \quad P(Z = 1/\mu_x) = e^{-\theta/\mu_x},$$

or, equivalently,

$$F_Z(y) = \begin{cases} 1 - e^{-\theta y}, & \text{for } 0 < y < 1/\mu_x, \\ 1, & \text{for } y \geq 1/\mu_x. \end{cases}$$

Moreover,

$$\begin{aligned} \frac{T_{\kappa(t)}}{t} &\xrightarrow{d} \mu_y Z \quad \text{as } t \rightarrow \infty, \\ \frac{S_{\kappa(t)}}{t} &\xrightarrow{d} \mu_x Z \quad \text{as } t \rightarrow \infty, \\ \frac{X_{\kappa(t)}}{t} &\xrightarrow{p} 0 \quad \text{and} \quad \frac{\max_{1 \leq k \leq \kappa(t)} X_k}{t} \xrightarrow{p} 0 \quad \text{as } t \rightarrow \infty. \end{aligned}$$

Remark 2. Note that $X_{\kappa(t)}$ is the last shock (which may or may not have caused failure), whereas $\max_{1 \leq k \leq \kappa(t)} X_k$ is the largest shock so far at the time of failure. In the event that failure is caused by an extreme shock these quantities coincide.

5.4.1 An auxiliary result

The essential ingredient in the proof is (the first half of) the following more general result.

Proposition 1. *Suppose that $\{U_t, t \geq 0\}$ and $\{V_t, t \geq 0\}$ are families of random variables, such that*

$$U_t \xrightarrow{p} a \quad \text{and} \quad V_t \xrightarrow{d} V \quad \text{as } t \rightarrow \infty,$$

for some finite constant, a , and a random variable V . Then

$$P(\min\{U_t, V_t\} > y) \rightarrow \begin{cases} P(V > y), & \text{for } y < a, \\ 0, & \text{for } y > a, \end{cases} \quad \text{as } t \rightarrow \infty,$$

and

$$P(\max\{U_t, V_t\} \leq y) \rightarrow \begin{cases} 0, & \text{for } y < a, \\ P(V \leq y), & \text{for } y > a, \end{cases} \quad \text{as } t \rightarrow \infty.$$

For details of the proof and additional results we refer to [11].

5.4.2 Some examples

Example 1. Boxing. A somewhat drastic example is boxing. In a fight a knockout may be caused either by a series of small (moderate) punches or by a real big one.

Example 2. Rainfall. On 17, August 1997, the city of Uppsala experienced extremely heavy rain for about 1 hour, and the basement in the home of Allan Gut was flooded. A year later it was raining fairly heavily on and off during a few days, which led to the basement being flooded again. The first instance obviously corresponds to an extreme shock causing failure, the second one to the cumulative situation. A more general example of the same kind is flooding in rivers or dams.

Example 3. Fatigue, tenacity. A material, for example a rope or a wire, can break either because of the cumulative effect of “normal” loads after a certain time period or because of a sudden (very) big load. A more common example is a coat hanger.

Example 4. Environmental damage. A factory may leak poisonous waste products into a river. After some time the vegetation and the fish in the river are dead, because of the cumulative effect. Or, they might be killed because of some catastrophe in the factory that, instantaneously, pours a huge amount of waste into the river.

Example 5. Radioactivity. A variation of the previous example is the atomic power station, which normally emits a daily amount of radioactivity, which after some (large?) time may cause a higher rate of cancer in the nearby population. Or, in case of a sudden meltdown.

5.5 More Realistic Model

A more realistic approach in the cumulative case would be to say that “minor shocks” have no effect in the long run. Instead of considering the total accumulation of shocks one might therefore prefer to consider only the last k_n shocks at the n th step, where k_n is “conveniently” chosen, that is, one considers only a window. Failure then occurs as soon as the total shock load in the window exceeds some given level.

In the extreme case one would like to include the fact that objects subject to shocks may be weakened by minor or “moderately sized” shocks, that they can be exhausted, which suggests some kind of “discount” of earlier shocks, for example that the level to be surpassed by an individual shock is reduced after the appearance of a moderately sized shock.

The mixed model would of course be to mix these extensions.

As an illustration we use the boxing example above. In the cumulative case the generalization amounts to considering the accumulated punches a boxer receives during, say, 1 minute, rather than during the whole fight so far. In the extreme case the second variation amounts to saying that the required power of a single knockout punch is reduced for each semi-terrible punch the fighter receives. This is, admittedly, not a very pleasant example, but it is illustrative.

Before moving into the more general models we collect some facts about the asymptotics of delayed sums that will be needed in the sequel.

5.5.1 Delayed sums

Let X, X_1, X_2, \dots be i.i.d. random variables (not necessarily positive), $\mu_x = E X < \infty$, and let $\{S_n, n \geq 1\}$ denote their partial sums. (Backward) *delayed sums* or *lag sums* are defined as

$$S_{k,n} = \sum_{j=n-k+1}^n X_j, \quad 1 \leq k \leq n.$$

The corresponding arithmetic means or *moving averages* are $Y_{k,n} = S_{k,n}/k$.

In the following assume that $k = k_n$ is nondecreasing and satisfies

$$k_n = o(n) \quad \text{and} \quad k_n \rightarrow \infty \quad \text{as} \quad n \rightarrow \infty. \quad (5.7)$$

The *weak* law of large numbers and the central limit theorem,

$$Y_{k_n,n} \xrightarrow{p} \mu_x \quad \text{and} \quad \frac{S_{k_n,n} - k_n \mu_x}{\sigma_x \sqrt{k_n}} \xrightarrow{d} N(0, 1) \quad \text{as} \quad n \rightarrow \infty,$$

respectively (the latter provided $\sigma_x^2 = \text{Var } X < \infty$) are immediate consequences of stationarity and the fact that they are purely distributional results.

However, in order to obtain strong laws, the growth of k_n plays the important role. The result in the case $k = k_n \sim \theta n$ for some $0 < \theta < 1$ is immediate, since

$$\frac{S_{k_n,n}}{k_n} = \frac{S_n}{n} \cdot \frac{n}{k_n} - \frac{S_{n-k_n}}{n-k_n} \cdot \frac{n-k_n}{k_n} \xrightarrow{a.s.} \mu_x \frac{1}{\theta} - \mu_x \left(\frac{1}{\theta} - 1\right) = \mu_x \quad \text{as} \quad n \rightarrow \infty.$$

The following strong law for the more interesting window size

$$k_n = [k(n)] \quad \text{where} \quad k(x) = cx^\gamma, \quad (0 < \gamma < 1, c > 0), \quad (5.8)$$

is, independently, due to [7], Theorem 1, and [16], Theorem 3.1. Here $[x]$ denotes the largest integer $k \leq x$. For extensions see [22].

Theorem 4. *We have*

$$\max_{k_n \leq j \leq n} \left| \frac{S_{j,n}}{j} - \mu_x \right| \xrightarrow{a.s.} 0 \quad \text{as} \quad n \rightarrow \infty,$$

if and only if $E|X|^{1/\gamma} < \infty$.

It follows in particular that

$$Y_{k_n,n} = \frac{S_{k_n,n}}{k_n} \xrightarrow{a.s.} \mu_x \quad \text{as} \quad n \rightarrow \infty, \quad (5.9)$$

if $E|X|^{1/\gamma} < \infty$, and it is not difficult to see that the moment assumption is also necessary. Furthermore $\nu(t)/t^{1/\gamma} \rightarrow (c\mu_x)^{-1/\gamma}$ a.s. (see [14]).

5.5.2 A generalized cumulative shock model

This model is based on the assumption that only some suitably final portion of the summands is of interest in every step, which calls for the family of stopping times

$$\nu(t) = \min\{n : S_{k_n, n} = \sum_{j=n-k_n+1}^n X_j > t\}, \quad t \geq 0. \quad (5.10)$$

Even though we do not consider all shocks in the fatal shock, $T_{\nu(t)} = \sum_{j=1}^{\nu(t)} Y_j$ is the lifetime of the system, while $T_{k_{\nu(t)}, \nu(t)} = \sum_{j=\nu(t)-k_{\nu(t)}+1}^{\nu(t)} Y_j$ is the duration of the fatal window.

Referring to the discussion in Section 5.5.1, we let $k(x) = cx^\gamma$, where $0 < \gamma < 1$ and c is some positive constant, and define the window size as $k_n = [k(n)]$ as in (5.8).

Since $S_{k_n, n} \xrightarrow{a.s.} \infty$ and $k_n \rightarrow \infty$ as $n \rightarrow \infty$, it follows that $\nu(t) \xrightarrow{a.s.} \infty$ and $k_{\nu(t)} \xrightarrow{a.s.} \infty$ as $t \rightarrow \infty$, that is, stopping occurs almost surely within finite time, in fact, one can show that all moments of the stopping time are finite ([14], Theorem 4.1).

The strong law and the central limit theorem for the fatal window and the number of shocks until failure are obtained with the aid of techniques from [12].

Theorem 5. *Suppose that $\mu_x < \infty$. Then for k_n satisfying (5.8)*

$$\begin{aligned} \frac{k_{\nu(t)}}{t} &\xrightarrow{a.s.} \frac{1}{\mu_x} \quad \text{as } t \rightarrow \infty, \\ \frac{\nu(t)}{t^{1/\gamma}} &\xrightarrow{a.s.} \frac{1}{(c\mu_x)^{1/\gamma}} \quad \text{as } t \rightarrow \infty. \end{aligned}$$

If, in addition, $\sigma_x^2 < \infty$, then

$$\frac{k_{\nu(t)} - t/\mu_x}{\sigma_x \mu_x^{-3/2} \sqrt{t}} \xrightarrow{d} N(0, 1) \quad \text{as } t \rightarrow \infty,$$

and

$$\frac{\gamma c^{1/\gamma} \mu_x^{(2+\gamma)/2\gamma}}{\sigma_x} \cdot \frac{\nu(t) - (\frac{t}{c\mu_x})^{1/\gamma}}{t^{(2-\gamma)/2\gamma}} \xrightarrow{d} N(0, 1) \quad \text{as } t \rightarrow \infty.$$

Next, here are the strong laws for the size $S_{k_{\nu(t)}, \nu(t)}$ of the fatal window and the stopped sum $S_{\nu(t)}$.

Theorem 6. *We have*

$$\begin{aligned} \frac{S_{k_{\nu(t)}, \nu(t)}}{t} &\xrightarrow{a.s.} 1 \quad \text{as } t \rightarrow \infty, \\ \frac{S_{\nu(t)}}{t^{1/\gamma}} &\xrightarrow{a.s.} \frac{\mu_x^{1-(1/\gamma)}}{c^{1/\gamma}} \quad \text{as } t \rightarrow \infty. \end{aligned}$$

The analogous results for the duration of the nonfatal window $T_{k_{\nu(t)}, \nu(t)}$ and the total time until failure $T_{\nu(t)}$ are as follows.

Theorem 7. *We have*

$$\begin{aligned}\frac{T_{k_{\nu(t)}, \nu(t)}}{t} &\xrightarrow{a.s.} \frac{\mu_y}{\mu_x} \quad \text{as } t \rightarrow \infty, \\ \frac{T_{\nu(t)}}{t^{1/\gamma}} &\xrightarrow{a.s.} \frac{\mu_y}{(c\mu_x)^{1/\gamma}} \quad \text{as } t \rightarrow \infty.\end{aligned}$$

If, in addition, σ_x^2 and σ_y^2 are finite, and that

$$\sigma^2 = \text{Var}(\mu_x Y - \mu_y X) > 0.$$

Then

$$\frac{T_{k_{\nu(t)}, \nu(t)} - \frac{\mu_y t}{\mu_x}}{\sigma \sqrt{\mu_x^{-3} t}} \xrightarrow{d} N(0, 1) \quad \text{as } t \rightarrow \infty,$$

and

$$\frac{T_{\nu(t)}/\nu(t) - \mu_y}{\sigma_y} (t/c\mu_x)^{1/2\gamma} \xrightarrow{d} N(0, 1) \quad \text{as } t \rightarrow \infty.$$

The proof of the theorem consists of an appropriate blend of [13, 12].

5.5.3 A generalized extreme shock model

The usual extreme shock model considers failure caused by one large shock. However, many systems are even harmed by some large but nonfatal shocks which influence the load a system can take. Then a large, less extreme shock after some nonfatal shocks might be sufficient in order to cause failure. More precisely, for fixed t a shock X_i can harm the system if it is larger than a boundary value $\beta = \beta_t$, say. It is nonfatal as long as $X_i < t$. If a first nonfatal large shock occurs with value in $[\beta_t, t]$, the crucial load limit of the system is no more t but a value $\alpha_t(1) \in [\beta_t, t]$, say. If a second nonfatal shock occurs with value in $[\beta_t, \alpha_t(1)]$, then the crucial load limit for the following shocks is lowered again. This value is denoted by $\alpha_t(2)$, and so on. Hence, for each t , we use a monotone sequence of boundary values $\alpha_t(k)$, $k \geq 0$, with

$$t = \alpha_t(0) \geq \alpha_t(1) \geq \alpha_t(2) \geq \dots \geq \beta_t,$$

and define the stopping by

$$\tau(t) = \min\{n : X_n \geq \alpha_t(L_t(n-1))\} \quad (5.11)$$

with the number of shocks larger than β_t until time n

$$L_t(n) = \sum_{i=1}^n 1(X_i \geq \beta_t) \quad \text{and} \quad L_t(0) = 0.$$

If the levels $\alpha_t(k)$ are all the same, then, by replacing X_i by $(X_i - \beta_t)_+$ we can apply the results of Section 5.3. If, however, the levels $\alpha_t(k)$ depend on the random shocks or better on the impacts of a shock above the level β_t , we might define the impacts $g(X_i - \beta_t)$ with $g(x) = 0$, for negative x , and the levels $\alpha_t(k) = \alpha_t(0) - \sum_{i \leq k} g(X_i - \beta_t)$. Then failure occurs if the impact of a large shock $g(X_n - \beta_t) > \alpha_t(n-1)$. But this is again the cumulative shock model with suitably replaced X_i 's.

The derivation of the distribution of $\tau(t)$ for nonrandom $\alpha_t(k)$ is rather straightforward using the independence of the X_i 's. The event $\{\tau(t) > m\}$, $m \geq 1$, is equal to the union of the disjoint events that there are exactly j nonfatal shocks up to index m (i.e., $L_t(m) = j$, with $0 \leq j \leq m$). The k th nonfatal large shock takes a value in $[\beta_t, \alpha_t(k-1)]$ with $1 \leq k \leq j$. The remaining shocks X_i are all smaller than β_t . The case $j = 0$ is obviously simple. Using $\prod_{k=0}^{-1} = 1$, this gives

$$P\{\tau(t) > m\} = \sum_{j=0}^m \binom{m}{j} F^{m-j}(\beta_t) \prod_{k=0}^{j-1} [F(\alpha_t(k)) - F(\beta_t)]. \quad (5.12)$$

The setup allows that $\alpha_t(k) = t$ for every k , which means that the nonfatal shocks do not harm the load of the system. It is also possible to choose $\beta_t = t$ implying $\alpha_t(k) = t$, which means the same again. This is the situation which was considered in [13].

The sequence $\{\alpha_t(k)\}$ could be parametrized in different ways, e.g., as a linear decrease $\alpha_t(k) = t(1 - \alpha k) \vee \beta_t$ related to the beginning growth of cracks in material or some functional form $\alpha_t(k) = t(1 - \alpha k^\delta) \vee \beta_t$.

The asymptotic behaviour depends on the behaviour of the distribution F and on the boundaries $\alpha_t(k), \beta_t$. We shall basically deal with two different types of $\alpha_t(k), \beta_t$. Since for fixed t , $\{\alpha_t(k)\}$ is monotone decreasing and bounded by β_t , there exists

$$\lim_{k \rightarrow \infty} \alpha_t(k) = \alpha_t(\infty).$$

The two cases of interest are

- (i) $\lim_{t \rightarrow \infty} \bar{F}(\alpha_t(\infty))/\bar{F}(\beta_t) = c > 0$,
- (ii) $\lim_{t \rightarrow \infty} \bar{F}(\alpha_t(\infty))/\bar{F}(\beta_t) = 0$, where $\bar{F}(x) = 1 - F(x)$.

In the first case, if a shock exceeds β_t , the chance of exceeding even one of the $\alpha_t(k)$'s is not (asymptotically) negligible. In the other case, the boundary β_t does not play a significant role for extreme shocks exceeding the $\alpha_t(k)$ levels. In the first case the mean number of exceedances of β_t before the fatal shock is finite, in the second case it is infinite, as will be shown.

For the first case we assume that

$$\bar{F}(\alpha_t(k))/\bar{F}(\beta_t) \rightarrow c_k \quad \text{for each } k. \quad (5.13)$$

Note that the monotonicity of $\{\alpha_t(k)\}$ implies that the c_k 's are increasing in $[0, 1]$. Here mainly the values c_k are important, not the load limits $\alpha_t(k)$. The constant c_k is the proportion of the probability that a shock exceeds the k th load limit, if the shock is a nonfatal one exceeding β_t .

Theorem 8. *Assume that (5.13) holds with $\bar{F}(\beta_t) \rightarrow 0$. Then*

$$P(\bar{F}(\beta_t)\tau(t) > z) \rightarrow \sum_{j \geq 0} e^{-z} \frac{z^j}{j!} \prod_{k=0}^{j-1} (1 - c_k) \quad \text{as } t \rightarrow \infty.$$

The proof follows from the asymptotic analysis of the exact distribution in (5.12). The details are given in [14].

Example 6. (i) If, as mentioned, all c_k 's are equal to 1, the levels $\alpha_t(k)$ are asymptotically as extreme as β_t . This is similar to the case $t = \alpha_t(k) = \beta_t$ for all t . Hence the simple result for the level t and no influence of nonfatal but large shocks derived in [13] are included too.

(ii) If $c_k = 0$ for every $k \leq K$, and $c_{K+1} = 1$, then $\bar{F}(\beta_t)\tau(t)$ is asymptotically $\Gamma(K+1, 1)$ -distributed. This occurs, for example if we assume that $\beta_t = \beta t$ and that $\alpha_t(k) = (t(1 - \alpha k) \vee \beta t)$ with $\bar{F}(x) \sim \exp(-x^{-\rho})$. Then $c_k = 0$ for $k \leq [(1 - \beta)/\alpha]$ and 1, otherwise. If \bar{F} is regularly varying with parameter ρ , we have $c_k = ((1 - \alpha k)/\beta)^{-\rho}$ for $k \leq [(1 - \beta)/\alpha]$ and the limiting distribution is

$$\sum_{j=0}^{[(1-\beta)/\alpha]} e^{-z} \frac{z^j}{j!} \prod_{k=0}^{j-1} \left[1 - \left(\frac{1 - \alpha k}{\beta} \right)^{-\rho} \right].$$

This distribution is a particular mixed Poisson distribution. Hence, for a practical safety discussion, it is rather easy to derive the asymptotic probabilities of a failure after n nonfatal shocks exceeding β_t numerically for given α , β and ρ .

The case $c_k = 0$ for all k implies that $\tau(t)$ is not properly normalized with $\bar{F}(\beta_t)$ in Theorem 8. We then apply the normalizations with respect to the fatal levels $\alpha_t(\infty)$, that is, with $\bar{F}(\alpha_t(\infty))$. Hence, the relations between the fatal levels $\alpha_t(k)$ are important.

Theorem 9. *Assume that*

$$\bar{F}(\alpha_t(k))/\bar{F}(\alpha_t(\infty)) \rightarrow a_k \quad \text{for every } k, \quad (5.14)$$

and that $\bar{F}(\alpha_t(\infty)) \rightarrow 0$ as $t \rightarrow \infty$ (note that the a_k 's are increasing in k , with values in $[0, 1]$). If $a_k \rightarrow 1$ as $k \rightarrow \infty$, then

$$P(\bar{F}(\alpha_t(\infty))\tau(t) > z) \rightarrow e^{-z} \quad \text{as } t \rightarrow \infty.$$

The proof is based on introducing two boundary schemes to have an upper and a lower bound for the considered distribution. These boundary schemes are $\alpha_t^-(k) = \alpha_t(\infty)$ for all k and $\alpha_t^+(k) = \alpha_t(k) \vee \alpha_t(K)$ for some K large. This implies that the corresponding stopping times $\tau^+(t)$ and $\tau^-(t)$ approximate $\tau(t)$: $\tau^-(t) \leq \tau(t) \leq \tau^+(t)$. It is then shown in [14] that the upper and lower stopping times have the same asymptotic distribution.

The asymptotic behaviour of the failure time $T_{\tau(t)}$ can now easily be derived.

Theorem 10. *Assume that either the conditions of Theorem 8 or 9 hold. Let μ_y exists. Then*

$$p_t T_{\tau(t)} \xrightarrow{d} \mu_y Z, \quad \text{as } t \rightarrow \infty,$$

where Z is the random variable with distribution given in Theorem 8 or 9, respectively.

5.5.4 A generalized mixed shock model

The two shock models can be mixed together as in Section 5.4. Namely, consider once again the mixture of the two shock models as given there with $\kappa(t) = \min\{\nu(t), \tau(t)\}$,

where $\nu(t)$ and $\tau(t)$ denote the corresponding stopping times of the two models considered, and define, in addition, $\kappa^*(t) = \max\{\nu(t), \tau(t)\}$. We assume that μ_x exists, that $k(n)$ is given by (5.8) with $0 < \gamma < 1$, and that $\beta(t)$ or $\alpha_t(\infty)$ are the upper $\theta t^{-1/\gamma}$ -quantiles of F in Theorems 8 or 9, respectively, which mean that, for some $\theta > 0$,

$$P(X_1 > u_t) = \theta t^{-1/\gamma} \quad (5.15)$$

for $u_t = \beta(t)$ and $\alpha_t(\infty)$, respectively. If the boundary u_t does not satisfy (5.15), then either $\tau(t)$ or $\nu(t)$ is dominating. More precisely, if $1/\bar{F}(u_t) = o(t^{1/\gamma})$, then $\kappa(t)$ is asymptotically equal to $\tau(t)$ and $\kappa^*(t)$ to $\nu(t)$. In the other case $t^{1/\gamma} = o(1/\bar{F}(u_t))$, and then $\kappa^*(t)$ is asymptotically equal to $\tau(t)$ and $\kappa(t)$ to $\nu(t)$. Given this setup the following result emerges.

Theorem 11. *Assume that the conditions of Theorem 8 or 9 with (5.15) and (5.8) hold. Then*

$$\kappa(t)/t^{1/\gamma} \xrightarrow{d} \min\{(1/c\mu_x)^{1/\gamma}, Z/\theta\},$$

and

$$\kappa^*(t)/t^{1/\gamma} \xrightarrow{d} \max\{(1/c\mu_x)^{1/\gamma}, Z/\theta\}.$$

Just as for the proof of Theorem 3, a key tool is Proposition 1.

5.5.5 Extreme shock models with varying threshold

In Section 5.5.3 we introduced a varying threshold which is decreasing because of partial damage by a large nonfatal shock. It is also quite sensible that in some systems the critical threshold might increase in particular in the beginning or during a run-in period. In the same way we can derive theoretical results for this even more realistic model.

We assume that the run-in period during which the critical load can increase and the structure is strengthened ends with the first damage or crack. After such an event the system can only be weakened. Such a pattern can be modeled as follows.

The loads $X_i, i \geq 1$ are an i.i.d. sequence of r.v.'s with distribution F . A shock or stroke X_i is strengthening the material if $X_i \in [\gamma, \beta)$. At the beginning the material supports a maximal load α , the critical threshold. After a strengthening stroke, the maximal load becomes larger, say $\alpha_1 = \alpha + b_1$ with $b_1 > 0$. This boundary increases with each strengthening stroke, inducing boundaries $\alpha_j = \alpha + b_j, j \geq 1$ with $b_j \uparrow$. After the first harmful stroke larger than β , but smaller than the critical level at this time point, the load boundary decreases, because of possible cracks or some weakening of the material. If it has reached the level α_k (because of k strengthening strokes prior to the first harmful stroke), the critical level then becomes $\alpha_k - c_1$ and decreases further by the next harmful, nonfatal strokes to $\alpha_k - c_2, \alpha_k - c_3, \dots$, with $c_j \uparrow (\geq 0)$. This is shown in Figure 5.1. There might be an upper load limit α^* for the α_k , as well as a lower load limit α_* for the $\alpha_k - c_l$. We set that $\alpha_* \geq \beta$. It is convenient to set $b_0 = c_0 = 0$.

Let $N_-(n)$ denote the number of weakening shocks $X_i, i < n$, before the n th shock

$$N_-(n) = \sum_{i < n} 1(X_i \in [\beta, \alpha + b_{N_+(i)} - c_{N_-(i)}),$$

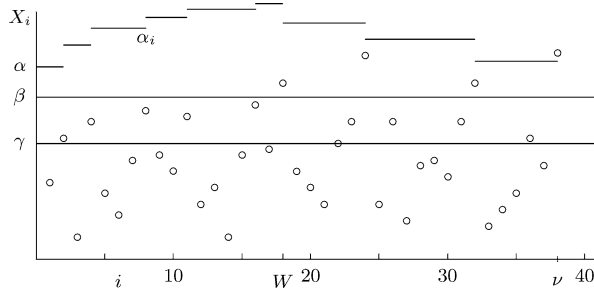


Figure 5.1. Realization of a sequence of shocks with strengthening and weakening load limits, depending on the values X_i . Here $\nu = 38$, $W = 18$, $N_+(\nu) = 5$, $N_-(\nu) = 3$

with $N_-(0) = 0$ and, similarly, $N_+(n)$ the number of strengthening strokes X_i , $i < n$, before the n th shock and before the first weakening or fatal shock

$$N_+(n) = \sum_{i < n} 1(X_i \in [\gamma, \beta), N_-(i) = 0),$$

with $N_+(0) = 0$. The critical boundary for the i th shock X_i is $\alpha_i = \alpha + b_{N_+(i)} - c_{N_-(i)}$.

In addition, let W be the index of the first harmful nonfatal or fatal shock, i.e., $X_i \geq \beta$, if such a shock occurs

$$W = \inf\{i : X_i \geq \beta\} \leq \infty.$$

If the set is empty, we set $W = \infty$.

Summarizing, the shock X_i has no impact if $X_i \leq \gamma$; it induces a strengthening of the material if $i < W$ and $X_i \in [\gamma, \beta)$; and it is fatal, if $X_i \geq \alpha_i = \alpha + b_{N_+(i)} - c_{N_-(i)}$. Note also that $N_+(k) = N_+(W)$ for all $k \geq W$ ($< \infty$).

We analyse the distribution of the number ν of shocks until the first fatal shock:

$$\nu = \min\{i : X_i \geq \alpha + b_{N_+(i)} - c_{N_-(i)}\},$$

with $b_0 = 0$ and $c_0 = 0$. The time until the fatal shock thus is T_ν . Its distribution depends on the distribution of ν and the distribution G of the interarrival times Y_i .

5.5.6 The exact distribution of ν

The distribution of ν can be derived in this more general model as in the simpler model dealt with in Section 5.5.3, see details in [9]. To derive $P(\nu > m)$ we have now to condition on the other random variables. If $N_-(m+1) = 0$, then $W > m$ and simply $P(\nu > m, N_-(m+1) = 0) = F^m(\beta)$. In the following we also use the notation $\alpha_{k,l} = \alpha + b_k - c_l$ for any $k, l \geq 0$.

If $l > 0$ with $k < j$ and $m \geq j + l - 1$, we can derive the joint distribution

$$\begin{aligned} P(\nu > m, N_+(m) = k, N_-(m+1) = l, W = j) &= \binom{j-1}{k} F^{j-1-k}(\gamma) \\ &\times [\bar{F}(\gamma) - \bar{F}(\beta)]^k \binom{m-j}{l-1} F^{m-j-l+1}(\beta) \prod_{h=1}^l [\bar{F}(\beta) - \bar{F}(\alpha_{k,h-1})] \end{aligned} \quad (5.16)$$

or for $m \geq j + l$ and $k < j$

$$P(\nu = m, N_+(m) = k, N_-(m) = l, W = j) = \binom{j-1}{k} F^{j-1-k}(\gamma) \\ \times [\bar{F}(\gamma) - \bar{F}(\beta)]^k \binom{m-j-1}{l-1} F^{m-j-l}(\beta) \prod_{h=1}^l [\bar{F}(\beta) - \bar{F}(\alpha_{k,h-1})] \bar{F}(\alpha_{k,l}). \quad (5.17)$$

If $k \geq j$ or $m \leq j + l - 1$, the latter probabilities are 0.

By summing the appropriate terms, we get the exact univariate and multivariate distributions for $\nu, N_+(m), N_-(l)$ and W , as well as for $N_+(\nu), N_-(\nu)$ which, however, have a simple form only in some special cases. Only such results are mentioned in the following.

Theorem 12. *The joint distribution of $N_+(\nu), N_-(\nu)$ with $l = 0$ is*

$$P(N_+(\nu) = k, N_-(\nu) = 0) = [1 - \bar{F}(\beta)/\bar{F}(\gamma)]^k [\bar{F}(\alpha_{k,0})/\bar{F}(\gamma)]. \quad (5.18)$$

The distribution of ν is given by $P(\nu > m) = P(\nu > m, N_-(\nu) > 0) + P(\nu > m, N_-(\nu) = 0)$ with $P(\nu > m, N_-(\nu) = 0)$ given above and

$$P(\nu > m, N_-(\nu) > 0) \\ = \sum_{k \geq 0, l > 0, j \geq 1} P(\nu > m, N_+(m) = k, N_-(m+1) = l, W = j) \\ = \sum_{k \geq 0, l > 0, j \geq 1} \binom{j-1}{k} F^{j-1-k}(\gamma) [\bar{F}(\gamma) - \bar{F}(\beta)]^k \\ \times \binom{m-j}{l-1} F^{m-j-l+1}(\beta) \prod_{h=1}^l [\bar{F}(\beta) - \bar{F}(\alpha_{k,h-1})]. \quad (5.19)$$

5.5.7 The asymptotic distribution

We consider the asymptotic behaviour of these random variables by letting the parameters $\alpha = \alpha(t), \beta = \beta(t)$ and $\gamma = \gamma(t)$ tend to $x_F \leq \infty$ as $t \rightarrow \infty$. We assume that x_F is a continuity point of F . For the limit distributions, certain additional restrictions will be imposed also on the b_k 's and c_k 's, being also dependent on t . Hence $\nu = \nu(t)$ will tend to infinity in general, depending on the underlying distribution F .

For the asymptotic behaviour let $\alpha(t) \rightarrow \infty, \beta(t) \rightarrow \infty$ and $\gamma(t) \rightarrow \infty$ as $t \rightarrow \infty$. The asymptotic behaviour depends also on the assumptions of the sequences $\{b_k\}$ and $\{c_k\}$. To simplify the notation, we do not indicate the dependence on t in the following, e.g., we write ν or b_k instead of $\nu(t)$ and $b_k(t)$, respectively. Again, $\beta(t)$ is such that the condition $\bar{F}(\beta(t)) \rightarrow 0$ as $t \rightarrow \infty$ holds. Furthermore, assume that

$$\lim_t \frac{\bar{F}(\beta(t))}{\bar{F}(\gamma(t))} = g \in [0, 1], \quad (5.20)$$

and

$$\lim_t \frac{\bar{F}(\alpha(t) + b_k(t) - c_l(t))}{\bar{F}(\beta(t))} = \lim_t \frac{\bar{F}(\alpha_{k,l}(t))}{\bar{F}(\beta(t))} = a_{k,l} \in [0, 1]. \quad (5.21)$$

The assumed monotonicity of the sequences $\{b_k\}$ and $\{c_k\}$ implies that the sequence $\{a_{k,l}\}$ is monotone, i.e., monotone decreasing in k with l fixed, and monotone increasing in l with k fixed.

We consider only the interesting cases with $g, a_{k,l} \in (0, 1)$. In general we approximate

$$\prod_{h=1}^l [\bar{F}(\beta) - \bar{F}(\alpha + b_k - c_{h-1})] \sim \bar{F}^l(\beta) \prod_{h=1}^l (1 - a_{k,h})$$

for each $l \geq 1$. Combining the exact distributions and the asymptotic conditions we derive the following result.

Theorem 13. *If (5.20) and (5.21) hold with $g, a_{k,l} \in (0, 1)$, then for any $k \geq 0$ and $l \geq 0$*

$$P(N_+(\nu) = k, N_-(\nu) = l) \rightarrow g(1 - g)^k \prod_{h=0}^{l-1} (1 - a_{k,h}) a_{k,l} \quad \text{as } t \rightarrow \infty.$$

For the particularly simple case that $a_{k,h} = a$ for all k and h , this limit distribution is the product of two geometric distributions

$$P(N_+(\nu) = k, N_-(\nu) = l) \rightarrow g(1 - g)^k (1 - a)^l a \quad \text{as } t \rightarrow \infty.$$

Because of the particular additional assumptions, the number of strengthening strokes does not have an influence on the number of weakening strokes asymptotically, which implies the asymptotic independence of $N_+(\nu)$ and $N_-(\nu)$ as $t \rightarrow \infty$.

The asymptotic investigation of the trivariate distribution of $\nu, N_+(\nu)$ and $N_-(\nu) = 0$ gives our next result (for details see [9]).

Theorem 14. *Suppose that (5.20) and (5.21) hold with $a_{k,l} \in (0, 1]$ for each k, l .*

(i) *For $k \geq 0 = l$ we have*

$$\lim_{t \rightarrow \infty} P(\nu \geq z/\bar{F}(\beta), N_+(\nu) = k, N_-(\nu) = 0) = \int_{z/g}^{\infty} v^k e^{-v} dv g(1 - g)^k a_{k,0}/k!.$$

(ii) *For $l \geq 1$ and $k \geq 0$,*

$$\begin{aligned} \lim_{t \rightarrow \infty} P(\nu \geq z/\bar{F}(\beta), N_+(\nu) = k, N_-(\nu) = l) \\ = \int_z^{\infty} \int_0^1 y^k (1 - y)^{l-1} \exp\{-yu(g^{-1} - 1)\} dy \exp\{-u\} u^{k+l} du \\ \times ((1 - g)/g)^k \prod_{h=0}^{l-1} (1 - a_{k,h}) a_{k,l} / (k!(l - 1)!) \quad \text{as } t \rightarrow \infty. \end{aligned}$$

If we set $z = 0$, the integrals of both statements can be determined explicitly, which reestablish Theorem 13.

Other limit distributions are determined by summing also the appropriate terms

$$P(\nu = m, N_+(m) = k, N_-(\nu) = l, W = j),$$

where the resulting sum can be simplified only under additional assumptions. We mention a last case which generalizes Theorem 5.1 in [14]. Consider the case that the impact of the strengthening strokes is asymptotically negligible, i.e., when $a_{k,h} = a_h$, for all k, h . Then the limit distributions of ν and T_ν have a simple form. The proof is given in [9].

Theorem 15. *If (5.20) and (5.21) hold with $a_{k,h} = a_h \in [0, 1]$ for each k and $h \geq 1$, then for $z_\beta = z/\bar{F}(\beta)$ with $z > 0$*

$$P(\nu > z_\beta) \rightarrow \sum_{l=0}^{\infty} \frac{z^l}{l!} e^{-z} \prod_{h=0}^{l-1} (1 - a_h) = 1 - H(z),$$

and

$$P(T_\nu > z_\beta) \rightarrow 1 - H(z/\mu_y),$$

as $t \rightarrow \infty$, where $\mu_y = E(Y_1) < \infty$.

Note also that the dependence between X_k and Y_k has no influence on the limit distribution of T_ν .

5.6 An Extension to Markovian Shock Models

An abbreviated description of the setup in Markov renewal theory (for further details we refer to [5, 4]) consists of measurable spaces (S, \mathcal{S}) , (Y, \mathcal{Y}) with countably generated σ -fields and a transition kernel $\mathbb{P} : S \times (\mathcal{S} \otimes \mathcal{B} \otimes \mathcal{Y}) \rightarrow [0, 1]$, where \mathcal{B} is the Borel- σ -algebra on \mathbb{R} . Further, $\{(M_n, X_n, Y_n), n \geq 0\}$ is an associated Markov chain defined on a probability space (Ω, \mathcal{A}, P) , with state space $S \times \mathbb{R} \times \mathcal{Y}$, i.e.

$$\begin{aligned} P(M_{n+1} \in A, X_{n+1} \in B, Y_{n+1} \in C \mid M_n, X_n, Y_n) \\ = \mathbb{P}(M_n, A \times B \times C) \quad \text{a.s.} \end{aligned}$$

for all $n \geq 0$ and $A \in \mathcal{S}, B \in \mathcal{B}, C \in \mathcal{Y}$. Thus $(M_{n+1}, X_{n+1}, Y_{n+1})$ depends on the past only through M_n . This produces a Markov chain.

Markov renewal theory deals with certain asymptotic properties of a *Markov random walk* $\{(M_n, S_n), n \geq 0\}$, where $S_n = X_0 + X_1 + \dots + X_n$ for $n \geq 0$ and related processes. An important assumption is *Harris recurrence*, which induces a regenerative structure on the full sequence $\{(M_n, X_n, Y_n), n \geq 0\}$ that generates a strictly increasing sequence of *regeneration times* σ_n , which split the process under consideration into one-dependent cycles (blocks) $C_n =: (M_k, X_k, Y_k)_{\sigma_n \leq k < \sigma_{n+1}}$. The cycles are stationary except for the first two; the full sequence is stationary under the stationary cycle initial distribution η .

Some results for stopped random walks are extended to the Markov renewal setup where the random walk is driven by a Harris recurrent Markov chain.

We conclude this chapter by sketching briefly how the results in Sections 5.2–5.4 can be extended to this more general model.

5.6.1 Cumulative shock models

With the aid of the extension of the strong law of large numbers and the central theorem of [15] (also [12], Section 4.2) which were instrumental for Theorem 1, one finds immediately that this very result remains true also in this more general setting. See also [6] for some related work.

5.6.2 Extreme shock models

We modify the arguments in [1] as follows. Let A_n , $n \geq 1$, denote the largest shock within the respective cycles, viz.

$$A_n = \max_{\sigma_n \leq k \leq \sigma_{n+1}-1} \{X_k\},$$

and define

$$\tau(t) = \min\{n : X_n > t\} \quad \text{and} \quad \tau_A(t) = \min\{n : A_n > t\},$$

that is, the stopping time for the individual shocks and for the cycles. Next, we replace p_t in Section 5.3 with

$$p_t = P(A_n > t) \quad \text{and} \quad p_t^{(A)} = P(A_n > t \mid A_{n-1} \leq t).$$

For the stopping time $\tau_A(t)$ we now have $P(\tau_A(t) = 1) = (1 - p_t)p_t^{(A)}$ and

$$P(\tau_A(t) = n) = (1 - p_t)(1 - p_t^{(A)})^{n-2} p_t^{(A)}, \quad n \geq 2,$$

and conclude as in Section 5.3 that

$$p_t \tau_A(t) \xrightarrow{d} \text{Exp}(1) \quad \text{as } t \rightarrow \infty, \tag{5.22}$$

assuming that $p_t \rightarrow 0$, which implies that

$$\frac{p_t^{(A)}}{p_t} \rightarrow 1 \quad \text{as } t \rightarrow \infty.$$

Now, the regeneration times $\{\sigma_k, k \geq 1\}$ form a delayed renewal process, so that from renewal theory for m -dependent random variables (see [17]) we know, in particular, that $\tau_A(t) \rightarrow +\infty$ as $t \rightarrow \infty$.

Assuming that the mean cycle length $\mu_\sigma = E_\eta \sigma_1 < \infty$ (where E_η denotes expectation with respect to the stationary cycle initial distribution η), the law of large numbers for randomly indexed sums (cf. [12], Theorem 1.2.3) then yields

$$\frac{\sigma_{\tau_A(t)}}{\tau_A(t)} \xrightarrow{a.s.} E \sigma = \mu_\sigma \quad \text{as } t \rightarrow \infty, \tag{5.23}$$

and for the last cycle that

$$\frac{\sigma_{\tau_A(t)} - \sigma_{\tau_A(t)-1}}{\tau_A(t)} \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty. \tag{5.24}$$

Combining the above with the sandwich inequality

$$\sigma_{\tau_A(t)-1} < \tau(t) \leq \sigma_{\tau_A(t)}, \quad (5.25)$$

we finally conclude that

$$p_t \tau(t) \xrightarrow{d} \text{Exp}(\mu_\sigma) \quad \text{as } t \rightarrow \infty. \quad (5.26)$$

We have thus obtained the analog of (5.3) (under the assumption of finite expected cycle lengths).

As for the failure time, which is a stopped sum (as in the cumulative case), we let $B_n = \sum_{j=\sigma_n}^{\sigma_{n+1}-1} Y_j$, $n \geq 1$, denote the durations of the cycles. Assuming that $\mu_B = E_\eta B_1 < \infty$, an application of the law of large numbers for m -dependent sequences then tells us that

$$\frac{\sum_{k=1}^n B_k}{n} \xrightarrow{a.s.} \mu_B \quad \text{as } n \rightarrow \infty, \quad (5.27)$$

which, in view of [12], Theorem 1.2.3 (again) permits us to conclude that

$$\frac{\sum_{k=1}^{\tau(t)-1} B_k}{\tau(t)-1} \xrightarrow{a.s.} \mu_B \quad \text{and that} \quad \frac{\sum_{k=1}^{\tau(t)} B_k}{\tau(t)} \xrightarrow{a.s.} \mu_B \quad \text{as } t \rightarrow \infty. \quad (5.28)$$

Combining (5.28) with (5.26) finally establishes the analog of Theorem 2, namely

$$p_t T_{\tau(t)} \xrightarrow{d} \text{Exp}(\mu_B \mu_\sigma) \quad \text{as } t \rightarrow \infty, \quad (5.29)$$

by copying the arguments in [1], p. 375.

Remark 3. In the i.i.d. case the cycle lengths are equal to 1, the cycles reduce to single random variables, and the results in this subsection to those in Section 5.5.3.

5.6.3 Mixed shock models

An inspection of Section 5.4 tells us that the only thing that matters there is Proposition 1. The analog of Theorem 3 in the Markovian setting may therefore be obtained from the results in the cumulative and extreme cases, together with Proposition 4.1.

References

1. Alsmeyer, G. and Gut, A. (1999). Limit theorems for stopped functionals of Markov renewal processes. *Annals Institute of the Statistical Mathematics* **51**, 369–382.
2. Anderson, K.K. (1987). Limit theorems for general shock models with infinite mean intershock times. *Journal of Applied Probability* **24**, 449–456.
3. Anderson, K.K. (1988). A note on cumulative shock models. *Journal of Applied Probability* **25**, 220–223.
4. Asmussen, S. (2003). *Applied Probability and Queues*, 2nd ed. Springer-Verlag, New York.

5. Athreya, K.B. and Ney, P. (1978). A new approach to the limit theory of recurrent Markov chains. *Transactions of the American Mathematical Society* **245**, 493–501.
6. Boshuizen, F.A. and Gouweleeuw, J.M. (1993). General optimal stopping theorems for semi-Markov processes. *Advances in Applied Probability* **24**, 825–846.
7. Chow, Y.S. (1973). Delayed sums and Borel summability of independent, identically distributed random variables. *Bulletin of the Institute of Mathematics, Academia Sinica* **1**, 207–220.
8. Cirillo, P. and Hüsler, J. (2010a). An urn approach to generalized extreme shock models. *Statistics and Probability Letters* **79**, 969–976.
9. Cirillo, P. and Hüsler, J. (2010b). Shock models and firms' default: parametric and nonparametric models. *Submitted*.
10. Gut, A. (1990). Cumulative shock models. *Advances in Applied Probability* **22**, 504–507.
11. Gut, A. (2001). Mixed shock models. *Bernoulli* **7**, 541–555.
12. Gut, A. (2009). *Stopped Random Walks*, 2nd ed. Springer-Verlag, New York.
13. Gut, A. and Hüsler, J. (1999). Extreme shock models. *Extremes* **2**, 293–305.
14. Gut, A. and Hüsler, J. (2005). Realistic variation of shock models. *Statistics and Probability Letters* **74**, 187–204.
15. Gut, A. and Janson, S. (1983). The limiting behaviour of certain stopped sums and some applications. *Scandinavian Journal of Statistics* **10**, 281–292.
16. Hanson, D.L. and Russo, R.P. (1981). On the law of large numbers. *Annals of Probability* **9**, 513–519.
17. Janson, S. (1983). Renewal theory for m -dependent variables. *Annals of Probability* **11**, 558–568.
18. Li, H. and Shaked, M. (1995). On the first passage times for Markov processes with monotone convex transition kernels. *Stochastic Processes and their Applications* **58**, 205–216.
19. Li, H. and Shaked, M. (1997). Ageing first-passage times of Markov processes: A matrix approach. *Journal of Applied Probability* **34**, 1–13.
20. Mallor, F. and Omey, E. (2001). Shocks, runs and random sums. *Journal of Applied Probability* **38**, 438–448.
21. Shanthikumar, J.G. and Sumita, U. (1983). General shock models associated with correlated renewal sequences. *Journal of Applied Probability* **20**, 600–614.
22. Steinebach, J. (1986). Some remarks on strong approximations and the Hanson-Russo law of large numbers. *Statistics and Decisions* **4**, 237–249.
23. Sumita, U. and Shanthikumar, J.G. (1985). A class of correlated cumulative shock models. *Advances in Applied Probability* **17**, 347–366.

Parametric Shock Models

Waltraud Kahle and Heide Wendt

Otto-von-Guericke-University, Magdeburg, Germany

Abstract: For analyzing reliability of technical systems it is often important to investigate degradation processes. In this chapter we describe a degradation process (Z_t) which is assumed to be generated by a position-dependent marking of a doubly stochastic Poisson process. Some characteristics of this process are described. For general and special cases it is calculated as the probability that the damage process does not cross a certain threshold before a time t . For some parametric intensity kernels of the corresponding marked point process it is determined as maximum likelihood estimations. Censored observations are taken into account. Furthermore, the large sample case is considered. Moment estimations are found and compared with maximum likelihood estimations.

Keywords and phrases: Marked point process, doubly stochastic Poisson process, position-dependent marking, shock model, parametric models, maximum likelihood estimation, moment estimation, asymptotic properties

6.1 Introduction

In connection with the investigation of the reliability of systems it is often necessary to consider the development of degradation processes at elements of such systems. These include processes of wear, corrosion, and crack-growth. This chapter deals with a degradation process (Z_t) whose paths are monotone increasing step functions. For modeling it, we use marked point processes $\Phi = ((T_n, X_n))_{n \geq 1}$, presented in detail, e.g., in [14] or [3]. The cumulative process (Z_t) is assumed to be generated by a position-dependent marking of a doubly stochastic Poisson process (T_n) . The doubly stochastic Poisson process was introduced by Cox [6]. Cramer [7] and Albrecht [1] applied it in risk theory, and Grandell [10] gave a detailed discussion of these processes and their impact on risk theory. Further applications of the doubly stochastic Poisson process (T_n) may be found in reliability theory, medicine, and queueing theory ([5, 3, 11, 13]).

Here we will consider a shock model where each degradation from a shock at the random time T_n is described by the scalar positive random variable X_n ($n = 1, 2, \dots$).

There is a stream of papers concerning shock models. Applications were given – among others – by Sobczyk [17] and Esary, Marshall, and Proshan [8]. Further papers dealing with shock models are [9, 16, 4, 19]. In this chapter we assume cumulated intensities of (T_n) given as a random multiple Y of a deterministic function η . Frequently, these stochastic intensities are more realistic than the deterministic intensity of a Poisson process. Each realization of the random variable Y leads to an other individual intensity of (T_n) which could be a result of individual environmental conditions or of the individual frailty of the item.

Our aim is to describe suitable models for degradation accumulation, to calculate the corresponding cumulative degradation at time t , to consider a first passage time, and to determine its distribution.

Further, estimators for the parameters of the deterministic part, the distribution of Y , and the distribution of the marks X_n ($n \geq 1$) using maximum likelihood theory are determined. We compare the estimators in the deterministic term when a doubly stochastic Poisson process is assumed, with those when a nonhomogeneous Poisson process is assumed. Furthermore, the problem of estimating the parameters by means of a censored sample is considered. Finally, moment estimators are considered and compared with maximum likelihood estimators. The paper is a summary and extension of [12], [20], and [21].

6.2 Modeling Degradation by Marked Point Processes

Let $[\Omega, \mathcal{F}, P]$ be a fixed probability space and let $\{\mathcal{F}_t\}$ be a filtration in \mathcal{F} . The random variable T_n ($n \geq 1$) is the time of the n th shock. We suppose

$$T_n < T_{n+1} \quad \text{if } T_n < \infty \quad \text{and} \quad T_n = T_{n+1} = \infty \quad \text{otherwise}.$$

The size of the n th increment of the cumulative degradation process $(Z(t))_{t \geq 0}$ is given by a nonnegative random variable X_n ($n \geq 1$). Thus,

$$Z(t) = \sum_{n=1}^{\infty} I(T_n \leq t) \cdot X_n$$

describes the total amount of degradation at time t . The sequence $\Phi = ((T_n, X_n))$ is called a *marked point process*, and $\Phi(t)$ is defined as the random variable representing the number of events occurred up to time t . Frequently, it is of interest to discuss the first passage problem that the process (Z_t) exceeds a pre-specified constant threshold level $h > 0$ for the first time. It is also possible to regard a (random) state of first X_0 at time $T_0 = 0$. Then the corresponding first passage time Z^h is given as

$$Z^h = \inf\{t : \sum_{n=0}^{\infty} I(T_n \leq t) \cdot X_n \geq h\}. \quad (6.1)$$

Let us mention Z^h coincide with some T_m for $m \in \mathbf{N}$. Typical courses of realizations of the processes $(\Phi(t))$ and $(Z(t))$ as well as of the variable Z^h are sketched in Figure 6.1.

We consider the filtration $\{\mathcal{F}_t\} = \{\mathcal{F}_t^\Phi\} \vee \sigma(Y)$, where $\{\mathcal{F}_t^\Phi\}$ describes the internal history of Φ , and Y (generating \mathcal{F}_0) is a nonnegative random variable with finite

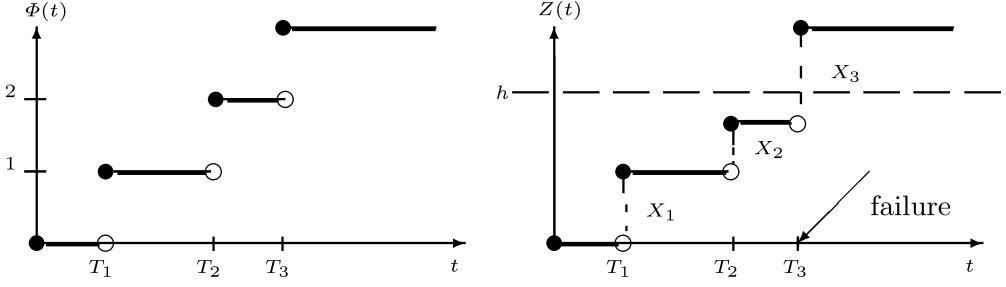


Figure 6.1. A realization of $\Phi(t)$ and $Z(t)$

expectation. The cumulated (P, \mathcal{F}_t) -stochastic intensity $\bar{\nu}(t)$ of (T_n) is assumed to be given by $\bar{\nu}(t) = Y \cdot \eta(t)$, where $\eta(t)$ is a deterministic function with derivative $\xi(t) \geq 0$. Hence, given the outcome $Y = y$ the random variable $\Phi(t)$ is Poisson distributed with mean $y \cdot \eta(t)$. And the unconditional distribution of $\Phi(t)$ is given by

$$P(\Phi(t) = k) = E \left[\frac{[Y\eta(t)]^k}{k!} \exp(-Y\eta(t)) \right], \quad k = 0, 1, \dots \quad (6.2)$$

The sequence (T_n) is called a (P, \mathcal{F}_t) -doubly stochastic Poisson process. Both the deterministic function η and the distribution of Y can be specified. For instance, if $\eta(t) = t$ then (T_n) is called a *mixed Poisson process*, and if $P(Y = y_0) = 1$ the sequence (T_n) represents a *nonhomogeneous Poisson process* with the deterministic intensity function $y_0 \cdot \xi(t)$.

The following types belong to the most common models for η :

1. Weibull type: $\eta(t) = t^{\alpha+1}$ ($\alpha > -1$)
2. log-linear type: $\eta(t) = t \cdot e^{\alpha t^\gamma}$ ($\alpha \geq 0, \gamma \geq 0$)
3. logistic type: $\eta(t) = t \cdot [1 + \ln(1 + \alpha t^\gamma)]$ ($\alpha \geq 0, \gamma > -1$).

The random variable Y is a frailty variable. The notion of frailty provides a convenient way to introduce random effects, association, and unobserved heterogeneity into models for survival data. In its simplest form, a frailty is an unobserved random proportionality factor that modifies the hazard function of an individual or of related individuals. Y can be specified too:

Example 1. Y is rectangular distributed in $[a, b]$ with $0 \leq a < b$. Then we get from (6.2) for $\eta(t) > 0$ by partial integration and with the convention $0^0 = 1$

$$p_k(t) = \frac{1}{\eta(t)(b-a)} \sum_{u=0}^k \left[\frac{[a\eta(t)]^u}{u!} e^{-a\eta(t)} - \frac{[b\eta(t)]^u}{u!} e^{-b\eta(t)} \right]. \quad (6.3)$$

◁

Example 2. Let $Y - y_0$ be gamma distributed with parameters $c > 0$ and $b > 0$ and pdf

$$f_Y(y) = I(y \geq y_0) \frac{c^b}{\Gamma(b)} (y - y_0)^{b-1} e^{-c(y-y_0)}.$$

Using

$$\int_{y_0}^{\infty} \frac{[c + \eta(t)]^{k-u+b}}{\Gamma(k-u+b)} (y - y_0)^{k-u+b-1} e^{-[c+\eta(t)] \cdot (y-y_0)} dy = 1,$$

we get

$$p_k(t) = \sum_{u=0}^k \frac{\Gamma(k-u+b)}{\Gamma(b)\Gamma(k+1)} \binom{k}{u} \left[\frac{c}{c+\eta(t)} \right]^b \left[\frac{\eta(t)}{c+\eta(t)} \right]^k \times \\ \times (y_0 [c + \eta(t)])^u e^{-y_0 \eta(t)}, \quad (6.4)$$

if we use the notations

$$q(t) := \frac{c}{c + \eta(t)} = \frac{c\eta^{-1}(t)}{c\eta^{-1}(t) + 1},$$

and

$$y_0[c + \eta(t)] = y_0\eta(t)[c\eta^{-1}(t) + 1],$$

then, we get

$$p_k(t) = \sum_{u=0}^k \left(\frac{[y_0\eta(t)]^u}{u!} e^{-y_0\eta(t)} \right) \times \\ \times \left(\frac{\Gamma(k-u+b)}{\Gamma(b)\Gamma(k-u+1)} q(t)^b (1-q(t))^{k-u} \right). \quad (6.5)$$

Hence, $p_k(t)$ are the probabilities of the Delaport distribution with parameters $\frac{c}{\eta(t)}$, b , and $y_0 \cdot \eta(t)$. From (6.5) the random number of shocks $\Phi(t)$ can be interpreted as a sum of two independent random variables $W_1(t)$ and $W_2(t)$, where $W_1(t)$ is Poisson distributed with expectation $y_0\eta(t)$ and $W_2(t)$ is negative binomial distributed with parameters $q(t) \in (0, 1)$ and $b > 0$. In the special case of $y_0 = 0$ $\Phi(t) = W_2(t)$ is negative binomial distributed and if Y is exponential distributed ($b = 1$) we get the geometrical distribution for $W_2(t)$ in $\Phi(t) = W_1(t) + W_2(t)$. \triangleleft

Example 3. Let Y be inverse Gaussian distributed with pdf

$$f_Y(y) = I(y \geq 0) \sqrt{\frac{\beta}{2\pi y^3}} \exp\left(-\frac{1}{2} \frac{\beta(y - \mu)^2}{\mu^2 y}\right).$$

From (6.2) we get

$$p_k(t) = \int_0^{\infty} \frac{(y\eta(t))^k}{k!} \sqrt{\frac{\beta}{2\pi y^3}} e^{-\frac{\beta[y\sqrt{1+2\eta(t)\mu^2/\beta}-\mu]^2}{2\mu^2 y}} dy \times \\ \times e^{-\frac{\beta}{\mu}(\sqrt{1+2\eta(t)\mu^2/\beta}-1)} \\ = e^{-\frac{\beta}{\mu}(\sqrt{1+2\eta(t)\mu^2/\beta}-1)} \frac{\eta(t)^k}{k! \sqrt{(1+2\eta(t)\mu^2/\beta)^k}} \cdot E[W^k].$$

The moments of order k of the inverse Gaussian distribution are given by

$$E[W^k] = \mu^k \sum_{u=0}^{k-1} \frac{(k-1+u)!}{(k-1-u)! u!} \left[\frac{\mu}{2\beta \sqrt{1+2\eta(t)\mu^2/\beta}} \right]^u.$$

Finally we get

$$p_k(t) = \exp\left(-\frac{\beta}{\mu}(\sqrt{1+2\eta(t)\mu^2/\beta}-1)\right) \left[\frac{\mu\eta(t)}{\sqrt{1+2\eta(t)\mu^2/\beta}}\right]^k \frac{1}{k!} \times \\ \times \sum_{u=0}^{k-1} \frac{(k-1+u)!}{(k-1-u)!u!} \left[\frac{\mu}{2\beta\sqrt{1+2\eta(t)\mu^2/\beta}}\right]^u. \quad (6.6)$$

◁

Let us now consider a marking of the sequence (T_n) [14]. At every time point T_n a shock causes a random degradation. We describe the degradation increment at T_n by the mark X_n . Let $(\mathbf{R}^+, \mathcal{B}^+)$ be the space of marks and let G be a stochastic kernel from $(\mathbf{R}^+, \mathcal{B}^+)$ to $(\mathbf{R}^+, \mathcal{B}^+)$. Then $\Phi = (T_n, X_n)$ is said to be a *position-dependent G -marking* of (T_n) if X_1, X_2, \dots are conditionally independent given (T_n) where for all $B \in \mathcal{B}^+$ and $n \in \mathbf{N}$ the following relation holds true:

$$P(X_n \in B | (T_n)) = G(T_n, B) \quad P - a.s. \quad \text{on } \{T_n < \infty\}. \quad (6.7)$$

Moreover, we assume that each mark X_n and Y are conditionally independent given (T_n) , i.e. $P(X_n \in B | (T_n), Y) = P(X_n \in B | (T_n))$. With a position-dependent marking it is possible to describe a degradation model where the increment at the n th shock depends on the time T_n of the n th shock.

Example 4. Let $t_0 \geq 0$ be a fixed time and let $(U_n), (V_n)$ be two sequences of i.i.d random variables with cdf F_U and F_V , respectively. The sequence of degradation increments (X_n) is defined by

$$X_n := I(T_n \leq t_0) U_n + I(T_n > t_0) V_n,$$

and $G(t, [0, x])$ is given by

$$G(t, [0, x]) = I(t \leq t_0) F_U(x) + I(t > t_0) F_V(x). \quad (6.8)$$

This means that at time t_0 the distribution of degradation increments is changing. For $t_0 = 0$ we get the independent marking. ◁

Example 5. Let (U_n) be a sequence of nonnegative i.i.d random variables with the density f_U and let $\delta \in \mathbf{R}$. We assume that the sequence (U_n) is independent of (T_n) . The sequence (X_n) is defined by $X_n = U_n \cdot e^{\delta T_n}$. That means we get degradation increments which tend to be increasing ($\delta > 0$) or decreasing ($\delta < 0$). The stochastic kernel G is given by

$$G(t, B) = \int_B f_U(x \cdot e^{-\delta t}) \cdot e^{-\delta t} dx, \quad B \in \mathcal{B}^+, \quad t \geq 0.$$

For $\delta = 0$ we have the special case of independent marking. In this case G defines a probability measure which is independent of the time t . ◁

6.3 Characteristics of the Degradation Process

In Section 6.2 we introduced a position-dependent marking of a (P, \mathcal{F}_t) -doubly stochastic Poisson process where the filtration $\{\mathcal{F}_t\}$ is given by $\mathcal{F}_t = \mathcal{F}_t^\Phi \vee \sigma(Y)$. Then by [14] $\Phi = (T_n, X_n)$ has the (P, \mathcal{F}_t) -stochastic intensity kernel

$$\lambda(t, B) = Y \xi(t) G(t, B), \quad B \in \mathcal{B}^+,$$

where

$$\lambda(t+, B) = \lim_{s \rightarrow 0+} \frac{1}{s} \cdot E \left[\sum_{n=1}^{\infty} I(t < T_n \leq t + s, X_n \in B) \middle| \mathcal{F}_t \right], \quad B \in \mathcal{B}^+.$$

But usually, Y cannot be observed such that the actual available information does not coincide with the filtration $\{\mathcal{F}_t\}$. In some cases as for applying the maximum likelihood theory we have to consider the intensity kernels with respect to the observed filtrations. For example, the (P, \mathcal{F}_t^Φ) -stochastic intensity kernel of Φ has the representation

$$\tilde{\lambda}(t, B) = E[\lambda(t, B) | \mathcal{F}_{t-}^\Phi] = \xi(t) G(t, B) E[Y | \mathcal{F}_{t-}^\Phi], \quad B \in \mathcal{B}^+,$$

with

$$E[Y | \mathcal{F}_{t-}^\Phi] = \frac{\int_0^\infty y^{\Phi(t-)+1} e^{-y \eta(t)} F_Y(dy)}{\int_0^\infty y^{\Phi(t-)} e^{-y \eta(t)} F_Y(dy)},$$

where F_Y denotes the distribution function of Y .

Example 6. Let $Y - y_0$ be a gamma distributed random variable (see Example 2). Then we find with $0^0 = 1$

$$E[Y | \mathcal{F}_{t-}^\Phi] = \frac{\Phi(t-) + 1}{c + \eta(t)} \cdot \frac{\sum_{u=0}^{\Phi(t-)+1} \frac{\Gamma(\Phi(t-)+1-u+b)}{u! (\Phi(t-)+1-u)!} [y_0 (c + \eta(t))]^u}{\sum_{u=0}^{\Phi(t-)} \frac{\Gamma(\Phi(t-)-u+b)}{u! (\Phi(t-)-u)!} [y_0 (c + \eta(t))]^u}.$$

◁

We want to present some characteristics of the counting process $(\Phi(t))$, the sequence of marks (X_n) , the degradation process (Z_t) , as well as the first passage time Z^h .

6.3.1 The counting process $(\Phi(t))$

Frequently, the expected number of shocks up to any time t and other moments of $\Phi(t)$ are of interest. According to (6.2) we can express the k th ordinary and central moments of $\Phi(t)$ as linear combinations of moments of Y multiplied by powers of the deterministic function $\eta(t)$.

Actually, let $S(k, u)$ denote the Stirling numbers of second kind where $S(k, u)$ can be recursively determined [2]

$$S(k, u) = S(k-1, u-1) + u \cdot S(k-1, u), \quad 1 \leq u \leq k, \quad (S(0, 0) := 1, S(0, u) = 0).$$

We make use of

$$n^k = \sum_{u=1}^k S(k, u) n(n-1) \cdots (n-u+1)$$

and we consider the factorial moments of a Poisson distributed random variable with mean $y\eta(t)$. Some elementary calculations yield

$$\begin{aligned} E[\Phi(t)^k] &= \int_0^\infty \sum_{n=0}^\infty n^k \frac{[y\eta(t)]^n}{n!} e^{-y\eta(t)} dF_Y(y) \\ &= \sum_{u=1}^k S(k, u) \cdot E[Y^u] \cdot \eta(t)^u. \end{aligned} \quad (6.9)$$

In particular, we find

$$\begin{aligned} E[\Phi(t)] &= E[Y] \eta(t), \\ \mu_2(\Phi(t)) &= \text{Var}(\Phi(t)) = E[Y] \eta(t) + D^2(Y) \eta(t)^2, \\ \mu_3(\Phi(t)) &= E[Y] \eta(t) + 3 D^2(Y) \eta(t)^2 + \mu_3(Y) \eta(t)^3, \\ \mu_4(\Phi(t)) &= E[Y] \eta(t) + (4 D^2(Y) + 3 E[Y^2]) \eta(t)^2 + \\ &\quad + (6 E[Y^3] - 12 E[Y^2] E[Y] + 6 E[Y]^3) \eta(t)^3 + \mu_4(Y) \eta(t)^4, \end{aligned} \quad (6.10)$$

where $\mu_k(\cdot)$ denotes the k th central moment of a random variable. For example, these results can be applied by the parameter estimation using the moments method which will be described in Section 6.6.

6.3.2 The sequence (X_n) and the degradation process (Z_t)

In the considered case of a position-dependent marking the size of the n th degradation increment depends on the time T_n of the n th shock. Using (6.7) the distribution of the n th mark X_n is given by

$$P(X_n \in B) = E[G(T_n, B)] = \int_0^\infty G(t, B) f_{T_n}(t) dt, \quad (6.11)$$

where the density f_{T_n} of T_n has the representation

$$f_{T_n}(t) = \frac{d}{dt} P(\Phi(t) \geq n) = n \cdot \frac{\xi(t)}{\eta(t)} \cdot P(\Phi(t) = n), \quad n \geq 1.$$

As we can see from (6.11) it is often difficult or strenuous to determine the distribution and the expectation of X_n , respectively. Frequently, we are more interested in the expected cumulative degradation at any time t . For the reason of more simple notations we deal with continuous marks X_n . Let $g(T_n, x)$ denote the conditional density of X_n where $g(t, x)$ is given by $g(t, x) = \frac{\partial}{\partial x} G(t, [0, x])$.

Theorem 1. *Let $\Phi = ((T_n, X_n))$ be a position-dependent marking of a doubly stochastic Poisson process with the stochastic intensity kernel*

$$\lambda(t, B) = Y \cdot \xi(t) G(t, B), \quad B \in \mathcal{B}^+.$$

Then the expectation of the cumulative degradation at time t is given as

$$E[Z(t)] = E[Y] \int_0^t \xi(s) \cdot \int_0^\infty x \cdot g(s, x) dx ds, \quad t \in \mathbf{R}^+. \quad (6.12)$$

Sketch of the proof: By the definition of an intensity it may be shown

$$E \left[\sum_{n=1}^{\Phi(t)} \psi(T_n) \right] = E[Y] \int_0^t \psi(s) \xi(s) ds, \quad t \in \mathbf{R}^+,$$

for any measurable function ψ . Further, we can express the expectation $E[Z(t)]$ as

$$E[Z(t)] = E \left[\sum_{n=1}^{\Phi(t)} \int_0^\infty x g(T_n, x) dx \right],$$

such that (6.12) immediately follows. \square

Let us mention that in the special case of an independent marking we get the well-known result $E[Z(t)] = E[\Phi(t)] \cdot E[X_1]$. Generally, the expectation $E[Z(t)]$ depends on both the shape of the deterministic functions η , ξ , and the kernel G . Figure 6.2 shows the mean cumulative degradation process $(E[Z_t])_{t \geq 0}$ for the special position-dependent marking $X_n = U_n e^{\delta T_n}$ from Example 5. In this case the following relation holds true:

$$E[Z(t)] = E[Y] \cdot E[U_1] \int_0^t \xi(s) \cdot e^{\delta s} ds.$$

Here we considered the function $\xi(t) = (\alpha + 1)t^\alpha$ from Weibull type for different α . We selected the values $\delta = -0.2$ (decreasing degradation increments), $\delta = 0$ (independent marking), and $\delta = 0.2$ (increasing degradation increments).

6.3.3 The cumulative degradation at time t

Another problem is to find the distribution of the cumulative degradation $Z(t)$ at time t . It is easy to see that the distribution can be calculated by

$$P(Z(t) \leq u) = I(0 \leq u)P(\Phi(t) = 0) + \sum_{n=1}^{\infty} P(X_1 + \cdots + X_n \leq u, \Phi(t) = n).$$

In the case of independent marking we get the more simple equation

$$P(Z(t) \leq u) = I(0 \leq u)P(\Phi(t) = 0) + \sum_{n=1}^{\infty} G^{*(n)}(u) \cdot P(\Phi(t) = n), \quad (6.13)$$

where G is the distribution function of any mark and $G^{*(n)}$ represents the n th convolution of G . The calculation of this distribution is difficult even in the case of independent marking. Methods for calculating the degradation distribution for special distributions of marks can be found in [15, 11]. Explicit expressions for this distribution are known mostly for simple models.

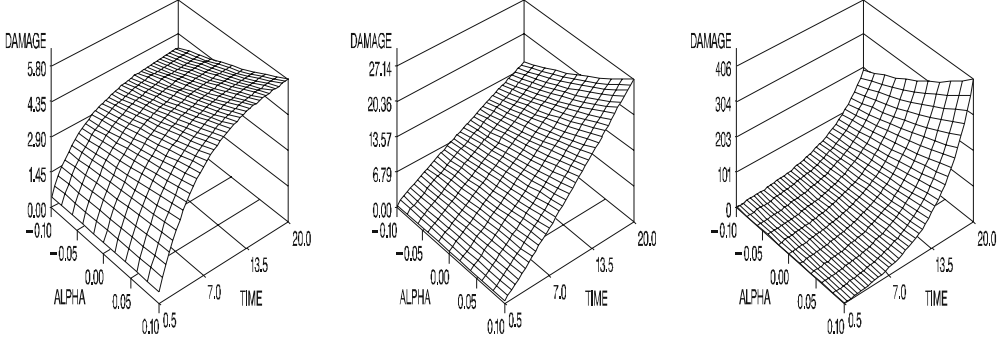


Figure 6.2. Mean cumulative degradation process for different markings of (T_n)

Example 7. We consider a nonhomogeneous Poisson process with intensity function $\lambda(t)$ and an independent marking. Let the marks be logarithmic $L(p)$ -distributed: $P(X_n = k) = \frac{1}{|\ln(1-p)|} \frac{p^k}{k}$ for $k = 1, 2, \dots$ and $p \in (0, 1)$. Then we get for the moment generating function $\mathcal{M}_X(s) := E[e^{s \cdot X}]$

$$\mathcal{M}_{Z(t)}(s) = e^{\lambda(t) \left[\frac{\ln(1-pe^s)}{\ln(1-p)} - 1 \right]} = \left[\frac{1-p}{1 - (1 - [1-p])e^s} \right]^{\lambda(t)/|\ln(1-p)|}, \quad (s < -\ln(p)).$$

It follows that $Z(t)$ is negative binomial distributed with the parameters $\frac{\lambda(t)}{|\ln(1-p)|}$ and $1-p$. \triangleleft

Now, we consider two nontrivial examples in which we also get a representation consisting of a finite number of terms. Again, we assume an independent G marking of $\bar{\Phi} = (T_n)$. Let the compensator of $\bar{\Phi}$ be given by $\bar{\nu}(t) = Y \eta(t)$ with an Erlang distributed Y , that is, $Y \sim \Gamma(c, b)$ with $b \in \mathbb{N}$.

Example 8. Let X_n be geometrically distributed: $P(X_n = j) = (1-p)p^{j-1}$ for $j = 1, 2, \dots$. Under the assumption about $\bar{\Phi}$ we get from (6.4)

$$P(\Phi(t) = n) = \binom{n+b-1}{b-1} \left[\frac{c}{c+\eta(t)} \right]^b \left[\frac{\eta(t)}{c+\eta(t)} \right]^n.$$

Let $[\cdot]$ be the integer part of a number. Then for $h \geq 0$ it follows from

$$P(Z(t) \leq h) = I(0 \leq h)P(\Phi(t) = 0) + \sum_{n=1}^{\infty} \sum_{j=n}^{[h]} (1-p)^n p^{j-n} \binom{j-1}{n-1} P(\Phi(t) = n)$$

that

$$P(Z(t) \leq h) = \left[\frac{c}{c+\eta(t)} \right]^b \left\{ 1 + \sum_{j=1}^{[h]} \sum_{n=1}^j \binom{n+b-1}{b-1} \binom{j-1}{n-1} \left[\frac{\eta(t) \cdot (1-p)}{c+\eta(t)} \right]^n p^{j-n} \right\}.$$

If $\lfloor h \rfloor$ is large, then an alternative representation is more comfortable. It can be shown that

$$F_{Z(t)}(h) = 1 - \left[\frac{c}{c + \eta(t)} \right]^b \left[\frac{pc + \eta(t)}{c + \eta(t)} \right]^{\lfloor h \rfloor} \cdot \sum_{q=0}^{b-1} \binom{b}{q+1} \left(\frac{\eta(t)}{c} \right)^{q+1} \sum_{u=0}^q \binom{\lfloor h \rfloor}{u} \left(\frac{c(1-p)}{pc + \eta(t)} \right)^u.$$

Then we get for $b = 1$, that is, for a geometrically distributed number of shocks $\Phi(t)$,

$$P(Z(t) \leq h) = 1 - \frac{\eta(t)}{c + \eta(t)} \cdot \left[\frac{p \cdot c + \eta(t)}{c + \eta(t)} \right]^{\lfloor h \rfloor} \quad (h \geq 0).$$

◁

A similar form of the distribution function we get in the following case:

Example 9. Let the marks X_n be exponentially distributed with parameter $p > 0$, so that $\sum_{n=1}^k X_n$ is Erlang distributed with parameters p and k . Then we get for $h \geq 0$

$$F_{Z(t)}(h) = \left[\frac{c}{c + \eta(t)} \right]^b \left(1 + \frac{\eta(t)}{c + \eta(t)} \frac{p}{(b-1)!} \int_0^h e^{-px} \sum_{n=0}^{\infty} \frac{(n+b)!}{(n+1)! n!} \left[\frac{\eta(t) p x}{c + \eta(t)} \right]^n dx \right).$$

Since

$$\sum_{n=0}^{\infty} \frac{(n+b)!}{(n+1)! n!} \cdot A^n = \sum_{q=0}^{b-1} \frac{(b-1)!}{q!} \binom{b}{q+1} \cdot A^q \cdot e^A \quad (A \in \mathbf{R}), \quad (6.14)$$

it follows that

$$F_{Z(t)}(h) = 1 - \left[\frac{c}{c + \eta(t)} \right]^b \exp \left(- \frac{p c h}{c + \eta(t)} \right) \sum_{q=0}^{b-1} \binom{b}{q+1} \left[\frac{\eta(t)}{c} \right]^{q+1} \sum_{u=0}^q \frac{1}{u!} \left[\frac{p c h}{c + \eta(t)} \right]^u.$$

Is $b = 1$, that is, Y is exponential distributed, then

$$P(Z(t) \leq h) = 1 - \frac{\eta(t)}{c + \eta(t)} \cdot \exp \left(- \frac{c p h}{c + \eta(t)} \right) \quad (h \geq 0).$$

◁

In some cases it is possible to express the infinite sum in (6.13) by integrals which can be calculated numerically. We will demonstrate it for a uniformly distributed frailty variable Y .

Example 10. We consider an independent marking of $\bar{\Phi}$, where every mark X_n is exponentially distributed with parameter p . Further, let Y be uniformly distributed on $[a, b]$ with $0 \leq a < b$. Then from (6.3) we get for $\eta(t) > 0$ and $h \geq 0$

$$F_{Z(t)}(h) = \frac{1}{(b-a)\eta(t)} \left\{ [e^{-a\eta(t)} - e^{-b\eta(t)}](1+hp) + \int_0^h p e^{-px} \sum_{n=0}^{\infty} \sum_{u=1}^{n+1} \left[\frac{[a\eta(t)]^u}{u!} e^{-a\eta(t)} - \frac{[b\eta(t)]^u}{u!} e^{-b\eta(t)} \right] \frac{[px]^n}{n!} dx \right\}. \quad (6.15)$$

Now we consider two parameters A, B and a function $r(y) = r(y; A, B)$ which depends on A, B , such that

$$r(y) = \sum_{n=0}^{\infty} \frac{R(n)}{n!} y^n \quad \text{with} \quad R(n) := B^n \sum_{u=1}^{n+1} \frac{A^u}{u!} \quad \text{and} \quad r(0) = A.$$

Our aim is to get another representation of $r(1)$. Since

$$r(y) = R(0) + \sum_{n=0}^{\infty} \frac{R(n+1)}{(n+1)!} y^{n+1}, \quad \text{where} \quad R(n+1) = B \left[R(n) + \frac{A^{n+2}}{(n+2)!} B^n \right],$$

we get

$$r(y) = A + B \sum_{n=0}^{\infty} \frac{R(n)}{(n+1)!} y^{n+1} + A \sum_{n=0}^{\infty} \frac{[AB]^n}{(n+1)!} \frac{y^{n+1}}{(n+2)!}.$$

It follows that

$$r(y) = A + B \cdot \int_0^y r(\tilde{y}) d\tilde{y} + A \cdot \sum_{n=1}^{\infty} \frac{[AB y]^n}{n! (n+1)!}. \quad (6.16)$$

Let I_1 be the modified Bessel function of first kind. Then we get using the properties of Bessel functions

$$\begin{aligned} S_{A,B}(y) &:= \sum_{n=1}^{\infty} \frac{[AB y]^n}{n! (n+1)!} = \frac{1}{\sqrt{AB y}} I_1(2 \cdot \sqrt{AB y}) - 1 \\ &= \frac{1}{\sqrt{AB y} \pi} \int_0^{\pi} \sinh(2 \cdot \sqrt{AB y} \cdot \sin(s)) \sin(s) ds - 1 \end{aligned}$$

with $S_{A,B}(0) = 0$. From (6.16) we get

$$r'(y) - B \cdot r(y) = A \cdot S'_{A,B}(y) \quad \text{where} \quad r(0) = A.$$

A unique solution $r(y)$ of this linear inhomogeneous differential equation of order 1 is

$$r(y) = A \cdot e^{B \cdot y} + A \cdot S_{A,B}(y) + AB e^{B y} \int_0^y e^{-B \cdot \tilde{y}} S_{A,B}(\tilde{y}) d\tilde{y}.$$

For $y = 1$ and with $B = px$, $A = a\eta(t)$, or $A = b\eta(t)$, respectively, we get from (6.15)

$$F_{Z(t)}(h) = \frac{I(0 \leq h)}{(b-a)\eta(t)} \left\{ [e^{-a\eta(t)} e^{-b\eta(t)}](1+h)p + p \int_0^h (\mathcal{K}(a, \eta(t), px) - \mathcal{K}(b, \eta(t), px)) dx \right\},$$

with

$$\mathcal{K}(\vartheta, \eta(t), px) = \vartheta \cdot \eta(t) e^{\vartheta \eta(t)} \quad \text{for } x = 0 \text{ or } \vartheta \eta(t) = 0,$$

and

$$\mathcal{K}(\vartheta, \eta(t), px) = \vartheta \eta(t) \cdot e^{-\vartheta \eta(t)} \left[e^{-px} \frac{I_1(2 \cdot \sqrt{\vartheta \eta(t) px})}{\sqrt{\vartheta \eta(t) px}} + px \int_0^1 e^{-px \cdot y} \frac{I_1(2 \cdot \sqrt{\vartheta \eta(t) pxy})}{\sqrt{\eta(t) pxy}} dy \right],$$

otherwise. The calculation of values of \mathcal{K} does not causes problems if there is some mathematical software. \triangleleft

Example 11. In case of an independent marking with exponential distributed marks X_n with parameter p and in case of an Erlang distributed random variable Y (see Example 2 with $b \in \mathbb{N}$ and $y_0 = 0$) it holds

$$P(Z(t) \leq u) = 1 - \left[\frac{c}{c + \eta(t)} \right]^b \exp \left(-\frac{pcu}{c + \eta(t)} \right) \cdot \sum_{q=0}^{b-1} \binom{b}{q+1} \left[\frac{\eta(t)}{c} \right]^{q+1} \sum_{l=0}^q \frac{1}{l!} \left[\frac{pcu}{c + \eta(t)} \right]^l.$$

\triangleleft

6.3.4 The first passage time Z^h

We defined the first passage time Z^h for any state of first X_0 at time $T_0 := 0$ and a random level $h - X_0$, respectively, i.e.,

$$Z^h = \inf \left\{ t : \sum_{n=0}^{\infty} I(T_n \leq t) \cdot X_n \geq h \right\} = \inf \left\{ t : Z(t) \geq h - X_0 \right\}.$$

If X_0 is deterministic and if it is possible to calculate the degradation distribution then with that the distribution of the first passage time is also given. Here we use the relation

$$P(Z^h > t) = P(Z(t) < h - X_0). \quad (6.17)$$

Example 12. Let us again consider the assumptions of example 11. We choose the fixed parameters $c = 4$, $b = 2$, and $h - x_0 = 100$. Figure 6.3 shows the corresponding densities of the first passage time Z^h for the different deterministic functions

$$\eta(t) = t^{\alpha+1}, \quad \eta(t) = t \exp(0.03t^{\alpha+0.2}), \quad \text{and} \quad \eta(t) = t [1 + \ln(1 + 1.85t^\alpha)].$$

The parameter α is varying where the number of shocks up to time t tends to be increasing with increasing amounts of α . \triangleleft

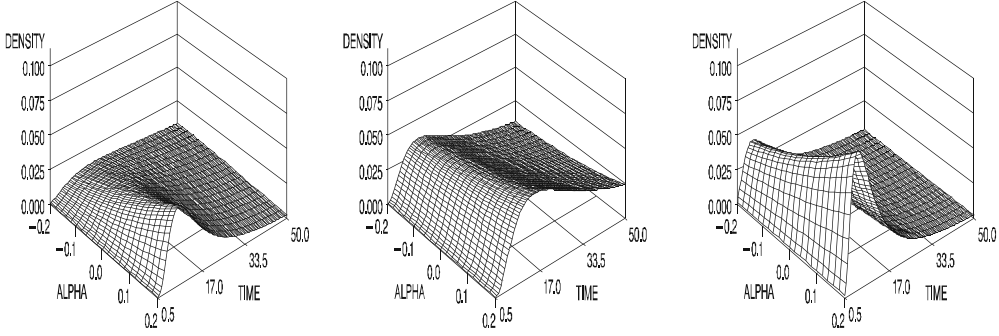


Figure 6.3. Density of the first passage time Z^h for different η

Now we want to consider a random state of first X_0 . In case of independent marking we get

$$P(Z^h > t) = \sum_{n=0}^{\infty} P(X_0 + \dots + X_n \leq h) \cdot P(\Phi(t) = n). \quad (6.18)$$

We have to determine the distribution of $X_0 + \dots + X_n$ where in general the distribution of X_0 does not coincide with that of any mark X_i , $i \geq 1$.

Example 13. In addition to the assumptions of Example 11 let X_0 be an exponential distributed random variable with parameter $p_0 \neq p$. Then we find (with $\sum_{r=0}^{-1} \cdot = 0$)

$$P(X_0 + \dots + X_n \leq h) = \frac{p_0 p^n}{(p - p_0)^n} \left\{ \frac{1 - e^{-hp_0}}{p_0} - \sum_{r=0}^{n-1} \frac{(p - p_0)^r}{p^{r+1}} \left[1 - \sum_{s=0}^r \frac{(hp)^s}{s!} e^{-hp} \right] \right\},$$

and after strenuous computations (see [19])

$$P(Z^h > t) = 1 - \left\{ e^{-p_0 h} \cdot \left[\frac{c(p - p_0)}{c(p - p_0) - p_0 \eta(t)} \right]^b - b e^{-h p \frac{c}{c + \eta(t)}} \right. \quad (6.19) \\ \cdot \left(\sum_{s=0}^{b-1} \frac{(b-1)!}{(b-1-s)!} \sum_{k=0}^b \binom{b}{k} (-1)^k \binom{k+s}{k} \sum_{v=1}^{k+s} \frac{(-1)^v}{v!} \left[\frac{\eta(t)}{c + \eta(t)} \right]^v [h \cdot p]^v \right. \\ \left. \left. \cdot \left[\left[\frac{1}{(p - p_0)h} \right]^{s+1} \left[\frac{c(p - p_0)}{c(p - p_0) - p_0 \eta(t)} \right]^b - \left[\frac{1}{ph} \right]^{s+1} \right] \right) \right\}.$$

This representation contains a finite number of terms. In the special case of an exponential distributed random variable Y with parameter c we get

$$P(Z^h > t) = 1 - \frac{c(p - p_0) e^{-p_0 \cdot h} - p_0 \eta(t) \exp\left(-\frac{c p h}{c + \eta(t)}\right)}{c(p - p_0) - p_0 \eta(t)},$$

$$\text{if } \frac{p}{p_0} \neq \frac{c + \eta(t)}{c} \quad \text{and}$$

$$P(Z^h > t) = 1 - \exp(-p_0 \cdot h) \left[1 + p_0 \left(1 - \frac{p_0}{p} \right) h \right],$$

$$\text{if } \frac{p}{p_0} = \frac{c + \eta(t)}{c}.$$

If p_0 tends to p it yields $P(Z^h > t) = 1 - \exp\left(-\frac{p c h}{c + \eta(t)}\right)$. \triangleleft

Example 14. Again, let us consider an independent marking of the sequence (T_n) where the marks are exponentially distributed with parameter p . We assume an exponential distributed variable Y with parameter c and an uniform distributed state of first X_0 with the density $f(x) = I(\vartheta \leq x \leq \beta) \frac{1}{\beta - \vartheta}$, $\vartheta > 0$.

Using the notations $a \wedge b = \min(a, b)$, $a \vee b = \max(a, b)$, and $0^0 = 1$ we find for $h \geq \vartheta$

$$\begin{aligned} P(X_0 + X_1 + \dots + X_n \leq h) &= \frac{(h \wedge \beta) - \vartheta}{\beta - \vartheta} - \frac{1}{(\beta - \vartheta) \cdot p} \\ &\cdot \sum_{v=0}^{n-1} (n-v) \cdot \left\{ \frac{(p[(h-\beta) \vee 0])^v}{v!} e^{-p[(h-\beta) \vee 0]} - \frac{(p[h-\vartheta])^v}{v!} e^{-p[h-\vartheta]} \right\}. \end{aligned}$$

Finally, some arithmetic yields

$$P(Z^h > t) = \frac{(h \wedge \beta) - \vartheta}{\beta - \vartheta} - \frac{\eta(t)}{(\beta - \vartheta) p c} \cdot \left[e^{-p[(h-\beta) \vee 0] \cdot \frac{c}{c + \eta(t)}} - e^{-p[h-\vartheta] \cdot \frac{c}{c + \eta(t)}} \right].$$

\triangleleft

In Figure 6.4 the influence of the state of first X_0 on the distribution of the first passage time Z^h is represented. We considered Examples 13 and 14 where we choose an exponential distributed random variable Y with parameter $c = 4$, a degradation level $h = 25$, a function η of logistic type, and $\vartheta = 0$. The parameters p_0 of the exponential distribution and β of the uniform distribution of the state of first X_0 are varying.

6.4 Maximum Likelihood Estimations

For practical applications it is necessary to estimate the parameters of all considered distributions. That can be done by likelihood theory or by the method of moments. In this section we consider maximum likelihood estimators.

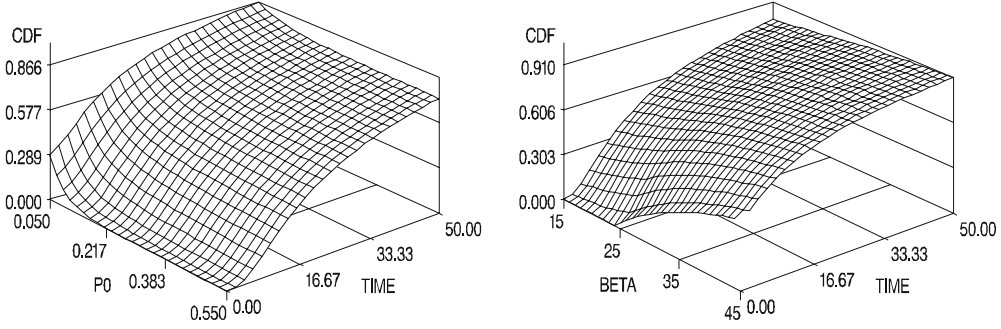


Figure 6.4. Distribution function of Z^h for different X_0

We assume the existence of a family of probability measures $P_\theta = (P_\theta, \theta \in \Theta)$ on (Ω, \mathcal{F}) . Let θ be given as $\theta = (\theta^Y, \theta^T, \theta^X) \in \mathbf{R}^p$ with $p = u + v + w$. Here, $\theta^Y \in \mathbf{R}^u$ is a parameter of the distribution function F_Y of Y , $\theta^T \in \mathbf{R}^v$ denotes a parameter of the deterministic terms η and ξ , respectively, and $\theta^X \in \mathbf{R}^w$ represents a parameter of the kernel G .

Remark 1. For the reason of more simple notations we deal with discrete random variables X_n . Here we use the notation

$$P_\theta(X_n = x | (T_n)) = G(T_n, \{x\}; \theta^X).$$

In the continuous case we have to replace the terms $G(T_n, \{x\}; \theta^X)$ by conditional densities of the marks X_n , $n \in \mathbf{N}$. \square

Under the assumptions of Section 6.2 we get the following $(P_\theta, \mathcal{F}_t)$ -stochastic intensity kernel of the marked point process $\Phi = ((T_n, X_n))$:

$$\lambda(t, B; \theta) = Y \cdot \xi(t; \theta^T) \cdot G(t, B; \theta^X), \quad B \in \mathcal{B}^+.$$

We can have different levels of information observing the degradation process:

1. All random variables, Y , (T_n) , and (X_n) are observable. Then the likelihood function is a product of three densities and the parameters can be estimated independently for each random variable. This case is not very realistic.
2. More interesting is the assumption that we can observe each time point of a shock and each increment of degradation but cannot observe the random variable Y . This is a more realistic assumption because Y is a variable which describes the individual shock intensity for each item, a frailty variable.
3. In many situations it might be possible that a failure is the result of an degradation process but we cannot observe the underlying degradation. By the maximum likelihood method it is possible to estimate all parameters in the model because the distribution of the first passage time contains all these parameters.

Let the observed history be given by $\{\mathcal{F}_t^\Phi\}$, that is, the frailty variable Y is not observable. In this case the σ -field $\sigma(Y)$ is unknown. Considering the $\{\mathcal{F}_t^\Phi\}$ -likelihood process

$(L(t; \theta))_{t \geq 0}$ we have to determine the $(P_\theta, \mathcal{F}_t^\Phi)$ -stochastic intensity kernel $\tilde{\lambda}$. According to Anderson et al. (93) and Last & Brandt (95) it follows that $\tilde{\lambda}$ is given as

$$\begin{aligned} \tilde{\lambda}(t, B; \theta) &= E_\theta[\lambda(t, B; \theta) | \mathcal{F}_{t-}^\Phi] = \xi(t; \theta^T) G(t, B; \theta^X) E_\theta[Y | \mathcal{F}_{t-}^\Phi] \\ &= \xi(t; \theta^T) G(t, B; \theta^X) \frac{\int_0^\infty y^{\Phi(t-)+1} e^{-y \eta(t; \theta^T)} F_Y(dy; \theta^Y)}{\int_0^\infty y^{\Phi(t-)} e^{-y \eta(t; \theta^T)} F_Y(dy; \theta^Y)} \end{aligned} \quad (6.20)$$

for $B \in \mathcal{B}^+$.

Now we are able to present the likelihood function $L(t; \theta)$ based on the sequence (T_n, X_n) up to time t . In Last & Brandt (95) it is shown that (the essential part of) $L(t; \theta)$ is given as

$$L(t; \theta) = \left[\prod_{n=1}^{\Phi(t)} \tilde{\lambda}(T_n, X_n; \theta) \right] \exp \left(\int_0^t \tilde{\lambda}(s, \mathbf{R}^+; \theta) ds \right).$$

Here $\tilde{\lambda}(t, \mathbf{R}^+; \theta) = \xi(t; \theta^T) E_\theta[Y | \mathcal{F}_t^\Phi]$ denotes the $(P_\theta, \mathcal{F}_t^\Phi)$ -intensity function of the sequence (T_n) . We use that

$$\Phi(s-) = n-1 \quad \text{on } (T_{n-1}, T_n] \quad \text{and} \quad \Phi(s-) = \Phi(t) \quad \text{on } (T_{\Phi(t)}, t].$$

Furthermore, it holds for any $k = 0, 1, \dots$

$$\begin{aligned} \frac{d}{ds} \ln \left(\int_0^\infty y^k e^{-y \eta(s; \theta^T)} F_Y(dy; \theta^Y) \right) \\ = -\xi(s; \theta^T) \frac{\int_0^\infty y^{k+1} e^{-y \eta(s; \theta^T)} F_Y(dy; \theta^Y)}{\int_0^\infty y^k e^{-y \eta(s; \theta^T)} F_Y(dy; \theta^Y)}. \end{aligned}$$

Hence, by elementary calculations we get the following log-likelihood function:

$$\begin{aligned} \ln L(t; \theta) &= \sum_{n=1}^{\Phi(t)} \left(\ln \xi(T_n; \theta^T) + \ln G(T_n, \{X_n\}; \theta^X) \right) + \\ &\quad + \ln \left(\int_0^\infty y^{\Phi(t)} e^{-y \eta(t; \theta^T)} F_Y(dy; \theta^Y) \right). \end{aligned} \quad (6.21)$$

But in many situations the sequence (T_n, X_n) cannot be fully observed up to fixed time t , e.g., if the system fails whenever the process $(Z(t))$ exceeds a certain value h resulting in censored data. Consider the $\{\mathcal{F}_t^\Phi\}$ -stopping time $\tilde{T} \wedge Z^h = \min\{\tilde{T}, Z^h\}$ where Z^h is given by

$$Z^h := \inf \left\{ t : Z(t) = \sum_{n=1}^\infty I(T_n \leq t) X_n \geq h \right\},$$

and \tilde{T} denotes an arbitrary $\{\mathcal{F}_t^\Phi\}$ -stopping time.

Our aim is to determine an estimator of θ based on $m \geq 1$ independent copies of the censored process Φ , where $T_{n,i}$, $X_{n,i}$, $\tilde{T}_i \wedge Z_i^h$, $\tilde{\lambda}_i$ as well as $\Phi_i(t)$ refer to the i th copy ($i = 1, \dots, m$). The $(P_\theta, \mathcal{F}_t^{\Phi_i})$ -intensity kernels of the censored processes are given by

$$I\left(t \leq \tilde{T}_i \wedge Z_i^h\right) \cdot \tilde{\lambda}_i(t, B; \theta), \quad B \in \mathcal{B}^+.$$

Anderson [3] (or direct calculations) show that we can replace in (6.21) t by $t \wedge \tilde{T}_i \wedge Z_i^h$. It is possible that the distributions of the independent stopping times \tilde{T}_i contain unknown parameters. In this case $L(t, \theta)$ is called a *partial likelihood function*. Using the notation

$$\Phi_i^c(t) = \Phi_i\left(t \wedge \tilde{T}_i \wedge Z_i^h\right) \quad \text{and} \quad \tilde{\lambda}_i(\cdot; \theta) = \tilde{\lambda}_i(\cdot, \mathbf{R}^+; \theta),$$

we get the maximum (partial) likelihood equations $\frac{\partial}{\partial \theta_r} \ln(L(t; \theta)) = 0$ with

$$\frac{\partial \ln L}{\partial \theta_j^Y} = \sum_{i=1}^m \frac{\partial}{\partial \theta_j^Y} \ln \left(\int_0^\infty y^{\Phi_i^c(t)} \exp(-y \eta(t \wedge \tilde{T}_i \wedge Z_i^h; \theta^T)) F_Y(dy; \theta^Y) \right), \quad (6.22)$$

$$\frac{\partial \ln L}{\partial \theta_k^T} = \sum_{i=1}^m \left\{ \sum_{n=1}^{\Phi_i^c(t)} \frac{\partial \ln(\xi(T_{n,i}; \theta^T))}{\partial \theta_k^T} - \frac{\partial \eta(t \wedge \tilde{T}_i \wedge Z_i^h; \theta^T)}{\partial \theta_k^T} \frac{\tilde{\lambda}_i(t_+ \wedge \tilde{T}_i \wedge Z_i^h; \theta)}{\xi(t \wedge \tilde{T}_i \wedge Z_i^h; \theta^T)} \right\}, \quad (6.23)$$

$$\frac{\partial \ln L}{\partial \theta_l^X} = \sum_{i=1}^m \sum_{n=1}^{\Phi_i^c(t)} \frac{\partial}{\partial \theta_l^X} \ln \left(G(T_{n,i}, \{X_{n,i}\}; \theta^X) \right) \quad (6.24)$$

provided the derivatives exist. In general these equations can be solved only by numerical methods. It can be shown that the estimator $\hat{\theta}^X$ does not depend on θ^T and θ^Y . Hence, the equation system (6.24) can be solved separately.

A simplification arises in the uncensored case where we have $t = t \wedge \tilde{T}_i \wedge Z_i^h$ for all $i = 1, \dots, m$. We use that the likelihood function is sufficient for θ . Exploiting the resulting sufficiency of the statistic

$$E_\theta \left[T^* \mid \sigma(\hat{\theta}) \right] \quad \text{with} \quad T^* = \frac{1}{m} \sum_{i=1}^m Y_i \cdot \eta(t; \theta^T),$$

we get

$$E_\theta \left[T^* \mid \sigma(\hat{\theta}) \right] = \frac{\eta(t; \hat{\theta}^T)}{\xi(t; \hat{\theta}^T)} \cdot \frac{1}{m} \sum_{i=1}^m \tilde{\lambda}_i(t_+; \hat{\theta}^Y, \hat{\theta}^T),$$

and

$$E_\theta \left[T^* \mid \sigma(\hat{\theta}) \right] = \frac{1}{m} \sum_{i=1}^m \Phi_i(t).$$

Hence, the following condition is satisfied $P_\theta - a.s.$:

$$\sum_{i=1}^m \frac{\tilde{\lambda}_i(t_+; \hat{\theta}^Y, \hat{\theta}^T)}{\xi(t; \hat{\theta}^T)} = \sum_{i=1}^m \frac{\int_0^\infty y^{\Phi_i(t)+1} e^{-y \eta(t; \hat{\theta}^T)} F_Y(dy; \hat{\theta}^Y)}{\int_0^\infty y^{\Phi_i(t)} e^{-y \eta(t; \hat{\theta}^T)} F_Y(dy; \hat{\theta}^Y)} = \sum_{i=1}^m \frac{\Phi_i(t)}{\eta(t; \hat{\theta}^T)}. \quad (6.25)$$

For more details we refer to Wendt [19]. An immediate consequence is the possibility to determine the estimator $\hat{\theta}^T$ of the system (6.23) without knowing the distribution of the random variable Y . Actually, if we take into consideration the relation (6.25) then the system of equation for $\hat{\theta}^T$ is given by

$$\sum_{i=1}^m \left\{ \sum_{n=1}^{\Phi_i(t)} \frac{\partial}{\partial \theta_k^T} \ln(\xi(T_{n,i}; \theta^T)) - \Phi_i(t) \cdot \frac{\partial}{\partial \theta_k^T} \ln(\eta(t; \theta^T)) \right\} \bigg|_{\theta^T = \hat{\theta}^T} = 0,$$

$k = 1, \dots, v$. Thus, for any distribution function F_Y the same estimator $\hat{\theta}^T$ is obtained as for the assumption of a nonhomogeneous Poisson process. Having determined $\hat{\theta}^T$, we can solve the system in (6.22) for θ^Y .

In the censored case, however, the ML-estimator $\hat{\theta}^T$ depends on the distribution function F_Y . Let us deal with an example.

Example 15. We consider the position-dependent marking $X_n = U_n \cdot e^{\delta T_n}$ from Example 5 where we make the assumptions:

- the i.i.d random variables U_n are exponentially distributed with the density $f_U(u) = I(u \geq 0) \cdot p e^{-p u}$,
- the $(P_\theta, \mathcal{F}_t)$ -intensity of the (uncensored) sequence (T_n) is given by $Y \cdot \xi(t; \theta^T)$ where the random variable Y is gamma distributed and the deterministic part is from Weibull type. That means, $\theta^Y = (c, b)$, $\theta^T = \alpha > -1$ with

$$f_Y(y) = I(y \geq 0) \frac{c^b}{\Gamma(b)} y^{b-1} e^{-c y} \quad \text{and} \quad \xi(t; \alpha) = (\alpha + 1) \cdot t^\alpha.$$

With respect to Remark 1 we have to replace in (6.24) the terms

$$\ln(G(T_{n,i}, \{X_{n,i}\}; \theta^X),$$

by

$$\ln(f_U(X_{n,i} \cdot e^{-\delta T_{n,i}})) - \delta T_{n,i}, \quad n = 1, \dots, \Phi_i^c(t), \quad i = 1, \dots, m.$$

After some calculations it can be shown that the ML-estimator $\hat{\theta}^X = (\hat{p}, \hat{\delta})$ is an unique solution of the system in (6.24). Here, \hat{p} is given as

$$\hat{p} = \frac{\sum_{j=1}^m \Phi_j^c(t)}{\sum_{i=1}^m \sum_{n=1}^{\Phi_i^c(t)} X_{n,i} \exp(-\hat{\delta} T_{n,i})},$$

where $\hat{\delta}$ is the solution of a non-linear equation.

From (6.22) and (6.23) we get the following equation system for the parameters α , c , and b :

$$\begin{aligned}
0 &= \sum_{i=1}^m \left\{ \frac{\Phi_i^c(t)}{\hat{\alpha} + 1} + \sum_{n=1}^{\Phi_i^c(t)} \ln(T_{n,i}) \right. \\
&\quad \left. - \left(t \wedge \tilde{T}_i \wedge Z_i^h \right)^{\hat{\alpha}+1} \ln \left(t \wedge \tilde{T}_i \wedge Z_i^h \right) \frac{\Phi_i^c(t) + \hat{b}}{\hat{c} + \left(t \wedge \tilde{T}_i \wedge Z_i^h \right)^{\hat{\alpha}+1}} \right\}, \\
0 &= \sum_{i=1}^m \left\{ \frac{\hat{b}}{\hat{c}} - \frac{\Phi_i^c(t) + \hat{b}}{\hat{c} + \left(t \wedge \tilde{T}_i \wedge Z_i^h \right)^{\hat{\alpha}+1}} \right\}, \\
0 &= \sum_{i=1}^m \left\{ \ln(\hat{c}) - \ln \left(\hat{c} + \left(t \wedge \tilde{T}_i \wedge Z_i^h \right)^{\hat{\alpha}+1} \right) + \sum_{n=1}^{\Phi_i^c(t)} \frac{1}{n - 1 + \hat{b}} \right\}.
\end{aligned}$$

In the uncensored case, however, we find after elementary calculations the estimator

$$\hat{\alpha} = \frac{\Phi_{\bullet}(t)}{\ln(t) \cdot \Phi_{\bullet}(t) - \sum_{i=1}^m \sum_{n=1}^{\Phi_i(t)} \ln(T_{n,i})} - 1 \quad \text{with} \quad \Phi_{\bullet}(t) = \sum_{i=1}^m \Phi_i(t),$$

which does not contain $\hat{\theta}^Y$. Let us again mention that the ML-estimator $\hat{\alpha}$ is not changed considering another distribution of Y .

Furthermore, we get in the uncensored case $\hat{c} = m \cdot \hat{b} \frac{t^{\hat{\alpha}+1}}{\Phi_{\bullet}(t)}$ where \hat{b} is the solution of

$$\sum_{i=1}^m \left\{ \sum_{n=1}^{\Phi_i(t)} \frac{1}{n - 1 + \hat{b}} + \ln \frac{m \hat{b}}{\Phi_{\bullet}(t) + m \hat{b}} \right\} = 0.$$

◁

Now, let us consider the interesting case where we want to estimate the parameters of the (nonobservable) frailty variable Y . From the intensity (6.20) we get that the essential part for estimating the parameter θ^Y is the last term

$$\lambda^* = \frac{\int_0^{\infty} y^{\Phi(t-)+1} e^{-y \eta(t; \theta^T)} F_Y(dy; \theta^Y)}{\int_0^{\infty} y^{\Phi(t-)} e^{-y \eta(t; \theta^T)} F_Y(dy; \theta^Y)},$$

which can be interpreted as the conditional expectation of Y given the history of observation. It is easy to see that this term depends only on θ^Y and θ^T .

Our aim is to determine an estimator of θ^Y based on $m \geq 1$ independent copies of the process Φ . Let $\Phi_i(t)$ be the observed number of shocks in the i th copy ($i = 1, \dots, m$). For the three special distributions of Y introduced in Section 6.3 we get the following essential parts of the process intensity and resulting maximum likelihood estimates.

Example 16. Let Y be uniformly distributed in $[a, b]$:

$$\lambda^* = \frac{\Phi(t-) + 1}{\eta(t; \theta^T)} \cdot \frac{\sum_{u=0}^{\Phi(t-)+1} \left(\frac{[b \eta(t; \theta^T)]^u}{u!} e^{-b \eta(t; \theta^T)} - \frac{[a \eta(t; \theta^T)]^u}{u!} e^{-a \eta(t; \theta^T)} \right)}{\sum_{u=0}^{\Phi(t-)} \left(\frac{[b \eta(t; \theta^T)]^u}{u!} e^{-b \eta(t; \theta^T)} - \frac{[a \eta(t; \theta^T)]^u}{u!} e^{-a \eta(t; \theta^T)} \right)}.$$

For this distribution we get two likelihood equations which are linear dependent and both lead to

$$\frac{1}{m} \sum_{i=1}^m \Phi_i(t) = \frac{\hat{a} + \hat{b}}{2} \cdot \eta(t; \hat{\theta}^T).$$

Consequently, it is not possible to estimate both parameters a and b . \triangleleft

Example 17. Let $(Y - y_0)$ be gamma distributed:

$$\lambda^* = \frac{\Phi(t-) + 1}{c + \eta(t; \theta^T)} \cdot \frac{\sum_{u=0}^{\Phi(t-)+1} \frac{\Gamma(\Phi(t-)+1-u+b)}{u! (\Phi(t-)+1-u)!} [y_0 (c + \eta(t; \theta^T))]^u}{\sum_{u=0}^{\Phi(t-)} \frac{\Gamma(\Phi(t-)-u+b)}{u! (\Phi(t-)-u)!} [y_0 (c + \eta(t; \theta^T))]^u}, \quad (0^0 := 1).$$

The likelihood equations can be found to be

$$0 = \sum_{i=1}^m \left\{ \frac{\hat{b}}{\hat{c}} - \frac{\Phi_i(t) + \hat{b}}{\hat{c} + \eta(t; \hat{\theta}^T)} + \hat{y}_0 \frac{\mathcal{U}_i(\Phi_i(t) - 1)}{\mathcal{U}_i(\Phi_i(t))} \right\}, \quad (6.26)$$

$$0 = \sum_{i=1}^m \left\{ -\eta(t; \hat{\theta}^T) + (\hat{c} + \eta(t; \hat{\theta}^T)) \frac{\mathcal{U}_i(\Phi_i(t) - 1)}{\mathcal{U}_i(\Phi_i(t))} \right\}, \quad (6.27)$$

$$0 = \sum_{i=1}^m \left\{ \ln(\hat{c}) - \ln(\hat{c} + \eta(t; \hat{\theta}^T)) + \frac{\mathcal{U}_i(\Phi_i(t) - 1)}{\mathcal{U}_i(\Phi_i(t))} (\hat{y}_0 [\hat{c} + \eta(t; \hat{\theta}^T)])^l \sum_{n=1}^{\Phi_i(t)-l} \frac{1}{n-1+\hat{b}} \right\}, \quad (6.28)$$

where

$$\mathcal{U}_i(n) = \sum_{l=0}^n \frac{\Gamma(n-l+\hat{b})}{\Gamma(n-l+1)l!} \cdot (\hat{y}_0 [\hat{c} + \eta(t; \hat{\theta}^T)])^l.$$

These equations must be solved numerically. For the special case of $b = 1$ (two parametric exponential distributions), (6.26) and (6.27) have to be solved with $\hat{b} = 1$ and $\mathcal{U}_i(n) = \sum_{l=0}^n \frac{(\hat{y}_0 [\hat{c} + \eta(t; \hat{\theta}^T)])^l}{l!}$. In the case of a two-parametric gamma distribution ($y_0 = 0$) we must consider (6.26) and (6.28), $\hat{y}_0 = 0$, and $\mathcal{U}_i(n) = \frac{\Gamma(n+\hat{b})}{n!}$. \triangleleft

Example 18. Let Y be inverse Gaussian distributed:

$$\lambda^* = \frac{\mu}{\sqrt{1 + 2\mu^2\eta(t; \theta^T)/\beta}} \cdot \frac{\sum_{u=0}^{\Phi(t-)} \frac{(\Phi(t-)+u)!}{(\Phi(t-)-u)! u!} \left[\frac{\mu}{2\beta \sqrt{1 + 2\mu^2\eta(t; \theta^T)/\beta}} \right]^u}{\sum_{u=0}^{\Phi(t-)-1} \frac{(\Phi(t-)-1+u)!}{(\Phi(t-)-1-u)! u!} \left[\frac{\mu}{2\beta \sqrt{1 + 2\mu^2\eta(t; \theta^T)/\beta}} \right]^u}.$$

The parameter μ can be found from

$$\hat{\mu} = \frac{1}{m \cdot \eta(t; \hat{\theta}^T)} \sum_{i=1}^m \Phi_i(t),$$

and β is the solution of

$$0 = \sum_{i=1}^m \left\{ \frac{1}{\hat{\mu}} \left(1 - \frac{1 + \eta(t; \hat{\theta}^T) \hat{\mu}^2 / \hat{\beta}}{\sqrt{1 + 2\eta(t; \hat{\theta}^T) \hat{\mu}^2 / \hat{\beta}}} \right) + \frac{\Phi_i(t) \cdot \eta(t; \hat{\theta}^T) \hat{\mu}^2 / \hat{\beta}^2}{1 + 2\eta(t; \hat{\theta}^T) \hat{\mu}^2 / \hat{\beta}} \right. \\ \left. - \mathcal{H}_i \cdot \frac{\hat{\mu} \cdot (1 + \eta(t; \hat{\theta}^T) \hat{\mu}^2 / \hat{\beta})}{2\hat{\beta}^2 (1 + 2\eta(t; \hat{\theta}^T) \hat{\mu}^2 / \hat{\beta})^{1.5}} \right\},$$

where

$$\mathcal{H}_i := \frac{\sum_{k=0}^{\Phi_i(t)-2} \frac{(\Phi_i(t) + k)!}{(\Phi_i(t) - 2 - k)! k!} \left(\frac{\hat{\mu}}{2\hat{\beta} \sqrt{1 + 2\eta(t; \hat{\theta}^T) \hat{\mu}^2 / \hat{\beta}}} \right)^k}{\sum_{k=0}^{\Phi_i(t)-1} \frac{(\Phi_i(t) - 1 + k)!}{(\Phi_i(t) - 1 - k)! k!} \left(\frac{\hat{\mu}}{2\hat{\beta} \sqrt{1 + 2\eta(t; \hat{\theta}^T) \hat{\mu}^2 / \hat{\beta}}} \right)^k}.$$

◁

6.5 The Large Sample Case

In this section we deal with the large sample properties of the maximum (partial) likelihood estimator $\hat{\theta} = (\hat{\theta}^Y, \hat{\theta}^T, \hat{\theta}^X)$ based on observations up to time $\tau < \infty$. It is found that the score vector process, i.e., the vector of the first derivatives of the log-likelihood process $(\ln L(t; \theta))$ with respect to $\theta \in R^p$ is a local square integrable $(P_\theta, \mathcal{F}_t^\Phi)$ -martingale. Under mild differentiation and continuity requirements on η , ξ , G , as well as on F_Y with respect to θ , Taylor series expansions on the score vector process are valid and further conditions ensure a proper behavior of the remaining terms in these Taylor expansions.

In particular, we shall assume that $\frac{1}{m} \sum_{i=1}^m \tilde{F}_i(s) \rightarrow \tilde{F}(s)$ where \tilde{F}_i denotes the distribution function of \tilde{T}_i and \tilde{F} is a distribution function for $m \rightarrow \infty$ and $s \in [0, \infty]$. Furthermore, we assume the existence of positive constants $k_j < 1$, $j = 1, 2$, such that the conditions $P_\theta(Z^h \leq \tau) \leq k_1$ and $\tilde{F}(\tau) \leq k_2$ are satisfied.

For additional information we have to distinguish between the regularity conditions for the distribution of marks and for the distribution of the process (T_n) . Here we want to give some notes for the sequence (T_n) . As for the other case we refer to Wendt [19].

An essential but hard to prove part of the regularity conditions is a condition on convergence in probability of the predictable covariation processes of the score vector process. Let $\tilde{\lambda}^m(t; \theta)$ denote the $(P_\theta, \mathcal{F}_t^\Phi)$ -stochastic intensity of the counting process $(\sum_{i=1}^m \Phi_i^c(t))$, i.e.,

$$\tilde{\lambda}^m(t; \theta) = \sum_{i=1}^m I\left(t \leq \tilde{T}_i \wedge Z_i^h\right) \cdot \tilde{\lambda}_i(t; \theta) \quad \text{with} \quad \tilde{\lambda}_i(t; \theta) = \xi(t; \theta^T) E_\theta[Y_i | \mathcal{F}_{t-}^{\Phi_i}].$$

Then we have to show that for $m \rightarrow \infty$

$$\frac{1}{m} \int_0^\tau \left(\frac{\partial}{\partial \theta_j} \ln \tilde{\lambda}^m(s, \theta) \right) \left(\frac{\partial}{\partial \theta_l} \ln \tilde{\lambda}^m(s, \theta) \right) \tilde{\lambda}^m(s, \theta) ds \Big|_{\theta=\theta_0} \xrightarrow{P_{\theta_0}} \sigma_{j,l}(\theta_0) ,$$

where $\sigma_{j,l}(\theta)$ are finite functions defined on a neighborhood Θ_0 of the true parameter θ_0 ($j, l = 1, \dots, u + v$). We make use of suitable criteria on the pointwise convergence in probability of a stochastic process to a deterministic function and then we apply *Helland's* condition to show the existence of the functions $\sigma_{j,l}(\theta)$ ([3], p. 85). But in general we are not able to determine these functions exactly. Additionally, we have to check the *Lindeberg*-type condition ensuring the jumps of normalized versions of the score vector process are asymptotically neglectable. Here we can again apply *Helland's* condition.

According to [3] (for (θ^Y, θ^T)) and Wendt [19] (for θ^X) it follows that under the regularity conditions with a probability tending to 1 likelihood equations have a consistent solution $\hat{\theta} = (\hat{\theta}^Y, \hat{\theta}^T, \hat{\theta}^X)$. Further, $\hat{\theta}$ is asymptotically multinormally distributed around the true value θ_0 . The asymptotic covariance matrix $[\sigma_{j,l}(\theta_0)]^{-1}$ may be estimated consistently by $\left[-\frac{1}{m} \frac{\partial^2 \ln L(\tau; \theta)}{\partial \theta_j \partial \theta_l} \Big|_{\theta=\hat{\theta}} \right]^{-1}$.

These propositions can be used to calculate confidence regions and acceptance areas for tests on the parameter vector θ applying the well-known Wald test statistic and the (partial) likelihood ratio test statistic, respectively.

Let us give some examples for such confidence regions. At first we consider the position-dependent marking of (T_n) from Example 15. We determine confidence estimations for the parameter $\theta^X = (p, \delta)$ based on $m = 25$ simulated realizations of $\Phi = ((T_n, X_n))$. The parameter α of the function $\eta(t) = t^{\alpha+1}$ is varying ($\alpha = -0.3$, $\alpha = 0$, and $\alpha = 0.3$). Hence, α has an influence on the sequence (T_n) and also on the confidence regions of θ^X .

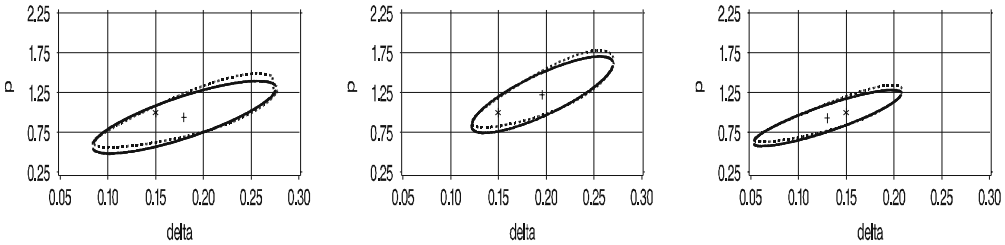


Figure 6.5. Confidence estimates for $\theta^X = (p, \delta)$ and different η

In Figure 6.5 the *solid line* represents the confidence region based on the Wald test statistic and the *broken line* those based on the likelihood ratio test statistic ($p_0 = 1$, $\delta_0 = 0.15$). Let us mention that $\frac{1}{m} \Phi_{\bullet}^c(10)$ was given by 2.1–7.1 (depending on α). Both statistics yield only little differences for the considered small sample size $m = 25$.

Next we want to give an example for θ^Y and θ^T . Here we assume an inverse Gaussian distribution for the random variable Y and a function η from logistic type, i.e.,

$$f_Y(y) = I(y \geq 0) \sqrt{\frac{\beta}{2\pi y^3}} \exp\left(-\frac{1}{2} \frac{\beta(y - \mu)^2}{\mu^2 y}\right),$$

and

$$\eta(t) = t [1 + \ln(1 + \alpha t^\gamma)].$$

In Figure 6.6 confidence estimates are plotted for μ , β and α , γ , respectively ($\mu_0 = 0.45$, $\beta_0 = 1$, $\alpha_0 = 0.5$, $\gamma_0 = 0.75$). Again, the solid line represents the confidence region based on the Wald test statistic. We have chosen the sample size $m = 100$ and $m = 250$.

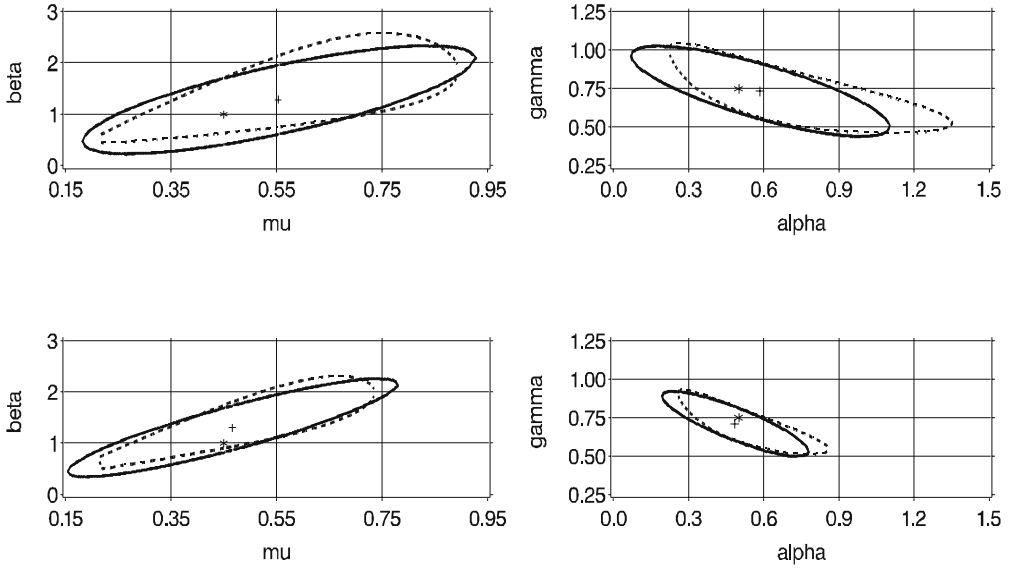


Figure 6.6. Confidence estimates for $\theta^Y = (\mu, \beta)$ and $\theta^T = (\alpha, \gamma)$

In the considered censored case we observed $\frac{1}{m} \Phi_{\bullet}^c(10) = 10.2$. Note that we assumed Y is not observable and the estimates are found from the observation of (T_n) . Furthermore, always two nuisance parameters have to be estimated.

Figure 6.6 shows both the confidence regions nearly coincide only in case of a large size m of realizations.

6.6 Moment Estimates

Let us consider again the case of observable counting process and unobservable frailty variable Y . Let $\mathbf{g}_k(W)$ and $\mathbf{z}_k(W)$ be the empirical ordinary and central moments, respectively, of a random variable W :

$$\mathfrak{g}_k(W) = \frac{1}{m} \sum_{i=1}^m W_i^k \quad \text{and} \quad \mathfrak{z}_k(W) = \frac{1}{m} \sum_{i=1}^m [W_i - \mathfrak{g}_1(W)]^k. \quad (6.29)$$

According to (6.2) we can express the k th ordinary and central moments of $\Phi(t)$ as linear combinations of moments of Y multiplied by powers of the deterministic function $\eta(t)$. As before in Section 6.3, let $S(k, u)$ denote the Stirling numbers of second kind. According to (6.10) we find for the first three moments

$$\begin{aligned} E_\theta[\Phi(t)] &= E_\theta[Y] \cdot \eta(t; \theta^T), \\ \mu_2^\theta(\Phi(t)) &= E_\theta[Y] \cdot \eta(t; \theta^T) + \mu_2^\theta(Y) \cdot \eta(t; \theta^T)^2, \\ \mu_3^\theta(\Phi(t)) &= E_\theta[Y] \cdot \eta(t; \theta^T) + 3\mu_2^\theta(Y) \cdot \eta(t; \theta^T)^2 + \mu_3^\theta(Y) \cdot \eta(t; \theta^T)^3, \end{aligned}$$

where $\mu_k(\cdot)$ denotes the k th central moment of a random variable. Let us further assume that the deterministic function $\eta(t)$ is known and that we are interested only in estimating the parameters of distribution of Y . The moments at the left-hand side are replaced by its empirical moments. Further, the moments of Y can be expressed in dependence of the moments of $\Phi(t)$ and the function η :

$$E_\theta[Y] = \eta(t; \theta^T)^{-1} \cdot E_\theta[\Phi(t)], \quad (6.30)$$

$$\mu_2^\theta(Y) = \eta(t; \theta^T)^{-2} \left\{ \mu_2^\theta(\Phi(t)) - E_\theta[\Phi(t)] \right\}, \quad (6.31)$$

$$\mu_3^\theta(Y) = \eta(t; \theta^T)^{-3} \left\{ \mu_3^\theta(\Phi(t)) - 3\mu_2^\theta(\Phi(t)) + 2E_\theta[\Phi(t)] \right\}. \quad (6.32)$$

Now it is possible to find moment estimates for all parameters of the distribution of Y . Let us consider again the three previous examples:

1. If Y is rectangular distributed in $[a, b]$ we get from (6.30) and (6.31) and taking into account $0 \leq a < b$

$$\hat{a} = \frac{\mathfrak{g}_1(\Phi(t))}{\eta(t; \hat{\theta}^T)} - \sqrt{3\mathcal{D}^2}, \quad \hat{b} = \frac{\mathfrak{g}_1(\Phi(t))}{\eta(t; \hat{\theta}^T)} + \sqrt{3\mathcal{D}^2},$$

with

$$\mathcal{D}^2 := \frac{\mathfrak{z}_2(\Phi(t)) - \mathfrak{g}_1(\Phi(t))}{\eta(t; \hat{\theta}^T)^2}.$$

In difference to the maximum likelihood method an unique admissible estimator exists if $\hat{a} \geq 0$ and $\mathcal{D}^2 > 0$. The assumption $\mathcal{D}^2 > 0$ is fulfilled for sufficient large values of m because \mathcal{D}^2 is a consistent estimate of the variance $\mu_2^\theta(Y)$ of Y .

2. If $Y - y_0$ is gamma distributed then it has the first three moments

$$E_\theta[Y] = \frac{b}{c} + y_0, \quad \mu_2^\theta(Y) = \frac{b}{c^2}, \quad \text{and} \quad \mu_3^\theta(Y) = 2 \frac{b}{c^3}.$$

From (6.30), (6.31), and (6.32) we get the unique moment estimators

$$\begin{aligned} \hat{c} &= 2 \frac{\mathfrak{z}_2(\Phi(t)) - \mathfrak{g}_1(\Phi(t))}{\mathcal{J}} \eta(t; \hat{\theta}^T), \\ \hat{y}_0 &= \frac{\mathfrak{g}_1(\Phi(t))}{\eta(t; \hat{\theta}^T)} - 2 \frac{[\mathfrak{z}_2(\Phi(t)) - \mathfrak{g}_1(\Phi(t))]^2}{\mathcal{J} \cdot \eta(t; \hat{\theta}^T)}, \\ \hat{b} &= 4 \frac{[\mathfrak{z}_2(\Phi(t)) - \mathfrak{g}_1(\Phi(t))]^3}{\mathcal{J}^2}, \end{aligned}$$

with

$$\mathcal{J} = \mathfrak{z}_3(\Phi(t)) - 3\mathfrak{z}_2(\Phi(t)) + 2\mathfrak{g}_1(\Phi(t)) .$$

3. For an inverse Gaussian distributed Y with $E_\theta[Y] = \mu$ and $\mu_2^\theta(Y) = \mu^3/b$ (6.30) and (6.31) give the unique estimators

$$\hat{\mu} = \frac{\mathfrak{g}_1(\Phi(t))}{\eta(t; \hat{\theta}^T)} \quad \text{and} \quad \hat{\beta} = \frac{\mathfrak{g}_1(\Phi(t))^3}{\eta(t; \hat{\theta}^T) \cdot [\mathfrak{z}_2(\Phi(t)) - \mathfrak{g}_1(\Phi(t))]} .$$

For this distribution we get the same $\hat{\mu}$ from the moment method as from the maximum likelihood method in Section 6.4.

The advantages of moment estimators in comparison to maximum likelihood estimators are its simple form and the fact that they can be found explicitly.

But it is well known that in general maximum likelihood estimates have better properties. Moment estimators are not asymptotically efficient in general, see [18]. In the next section we compare the two methods concerning to its bias and variance.

6.7 Comparison of Maximum Likelihood and Moment Estimates

Let $Y - y_0$ be gamma distributed and let η be Weibull, it is $\eta(t; \alpha) = t^{\alpha+1}$ with $\alpha > -1$. We have considered sample sizes of $m = 50$, $m = 100$, $m = 250$, and $m = 500$. For each realization a path of the degradation process was simulated with true parameter $\theta^Y = (c, y_0, b) = (2.4, .5, 1.2)$. The observation is assumed to continue up to time $t = 10$. We have considered three different values of the parameter α in the deterministic part $\eta(t; \alpha)$. For $\alpha = -0.3$ the derivative of $\eta(t; \alpha)$ is a decreasing function. The expected number of jumps up to time $t = 10$ is 5.01. If $\alpha = 0$ then we get a linear cumulative intensity or a constant hazard and the expected number of jumps up to time $t = 10$ is 10. Last, for $\alpha = 0.3$ the derivative of $\eta(t; \alpha)$ is increasing and the expected number of jumps up to time $t = 10$ is 19.95. Such a simulation was repeated 750 times and from these 750 parameter estimates the mean and the variance of the estimator were calculated. The results are shown in Table 6.1.

If the parameter α in the deterministic part is unknown, too, then it can be estimated by

$$\hat{\alpha} = \frac{\sum_{i=1}^m \Phi_i(t)}{\ln(t) \cdot \sum_{i=1}^m \Phi_i(t) - \sum_{i=1}^m \sum_{n=1}^{\Phi_i(t)} \ln(T_{n,i})} - 1 .$$

$\hat{\alpha}$ is a maximum likelihood estimator which does not contain other parameters. In Table 6.2 the results of the same simulation are shown with the difference that now α is unknown and has to be estimated.

From the simulation we get the following results:

1. Influence of α :

In both cases we can see that for $\alpha = 0.3$ the variances of the estimator \hat{y}_0 (in both cases, MLE and ME) are smaller than for $\alpha = 0$ or $\alpha = -0.3$. The variances of the moment estimators for \hat{b} and \hat{c} are also smaller for $\alpha = 0.3$ than for $\alpha = 0$

Table 6.1. Empirical moments of maximum likelihood (MLE) and moment (ME) estimators ($\theta_0^Y = (2.4, 0.5, 1.2)$)

		$m = 50$		$m = 100$		$m = 250$		$m = 500$	
		Mean variance		Mean variance		Mean variance		Mean variance	
$\alpha = -0.3$									
MLE	c	2.614	1.798	2.435	0.762	2.386	0.302	2.401	0.140
	y_0	0.599	0.018	0.561	0.016	0.520	0.008	0.508	0.004
	b	0.861	0.184	0.975	0.190	1.119	0.110	1.171	0.062
ME	c	2.910	3.902	2.824	2.208	2.826	1.475	2.758	0.893
	y_0	0.480	0.047	0.458	0.041	0.436	0.031	0.444	0.023
	b	1.680	2.552	1.708	2.026	1.754	1.507	1.650	1.052
$\alpha = 0.0$									
MLE	c	2.197	0.694	2.205	0.435	2.351	0.262	2.339	0.149
	y_0	0.574	0.013	0.560	0.010	0.532	0.005	0.536	0.003
	b	0.977	0.249	1.034	0.198	1.154	0.146	1.125	0.078
ME	c	3.189	2.533	2.846	1.325	2.771	0.876	2.613	0.427
	y_0	0.431	0.037	0.457	0.027	0.466	0.019	0.489	0.010
	b	1.982	2.030	1.716	1.257	1.621	0.855	1.431	0.375
$\alpha = 0.3$									
MLE	c	2.497	0.706	2.484	0.573	2.392	0.268	2.446	0.143
	y_0	0.521	0.010	0.519	0.007	0.523	0.004	0.497	0.002
	b	1.282	0.404	1.270	0.336	1.198	0.147	1.231	0.079
ME	c	3.361	2.066	3.133	1.399	2.758	0.672	2.602	0.347
	y_0	0.388	0.030	0.418	0.024	0.464	0.013	0.470	0.007
	b	2.276	2.110	2.004	1.457	1.596	0.625	1.417	0.279

Table 6.2. Empirical moments of maximum likelihood (MLE) and moment (ME) estimators ($\theta_0^Y = (2.4, 0.5, 1.2)$)

		$m = 50$		$m = 100$		$m = 250$		$m = 500$	
		Mean variance		Mean variance		Mean variance		Mean variance	
$\alpha = -0.3$									
MLE	c	2.187	3.217	2.000	0.913	2.125	0.481	2.197	0.407
	y_0	0.683	0.040	0.666	0.040	0.623	0.027	0.607	0.024
	b	0.818	0.157	0.891	0.141	1.026	0.084	1.072	0.054
	α	-0.328	0.011	-0.337	0.010	-0.337	0.009	-0.333	0.009
ME	c	2.801	3.942	2.729	3.067	2.641	1.622	2.607	1.144
	y_0	0.548	0.066	0.533	0.066	0.523	0.049	0.527	0.041
	b	1.621	2.179	1.739	2.357	1.669	1.410	1.572	0.916
$\alpha = 0.0$									
MLE	c	2.194	1.156	2.177	0.695	2.209	0.625	2.299	0.403
	y_0	0.594	0.029	0.594	0.028	0.602	0.028	0.592	0.019
	b	1.072	0.304	1.157	0.207	1.194	0.147	1.137	0.084
	α	-0.031	0.014	-0.040	0.014	-0.048	0.012	-0.036	0.010
ME	c	3.128	3.691	2.800	1.645	2.660	1.400	2.503	0.678
	y_0	0.438	0.051	0.471	0.041	0.516	0.038	0.532	0.025
	b	2.082	2.271	1.885	1.264	1.721	0.826	1.485	0.394
$\alpha = 0.3$									
MLE	c	2.571	0.924	2.652	1.001	2.394	0.495	2.363	0.286
	y_0	0.510	0.014	0.520	0.014	0.546	0.011	0.552	0.010
	b	1.424	0.516	1.455	0.378	1.269	0.148	1.160	0.059
	α	0.286	0.010	0.280	0.009	0.277	0.007	0.278	0.006
ME	c	3.324	2.478	3.234	1.748	2.687	0.900	2.494	0.497
	y_0	0.394	0.038	0.421	0.032	0.489	0.020	0.517	0.015
	b	2.297	1.926	2.171	1.535	1.639	0.703	1.409	0.250

or $\alpha = -0.3$, while α does not influence the variances of the maximum likelihood estimators.

2. The variances of the moment estimators are 2–4 times larger than the variances of the maximum likelihood estimators.
3. Both bias and variance are particularly visible smaller if the parameter θ^T is known (with the exception of the independent of η moment estimate \hat{b}). The ratio of the variances of the maximum likelihood estimators and moment estimators, however, is the same for known and for unknown θ^T .

The moment estimates are easy to calculate, where in many cases it is difficult to find maximum likelihood estimates. Moreover, there are problems in which the maximum likelihood estimate of the parameters does not exist.

On the other hand, the maximum likelihood estimators have a noticeable smaller variance as moment estimators. The ratio of the variances of the maximum likelihood estimators and moment estimators becomes smaller with increasing sample size.

6.8 Conclusion

In the chapter we considered the modeling of degradation by a relatively wide class of marked point processes. The degradation process (Z_t) was assumed to be generated by a position-dependent marking of a doubly stochastic Poisson process. Characteristics of this process were described.

For some parametric intensity kernels of the corresponding marked point process it was determined maximum likelihood estimations. Censored observations were taken into account. Moment estimations were found and compared with maximum likelihood estimations.

References

1. Albrecht, P. (1981). *Dynamische statistische Entscheidungsverfahren für Schadenzahlprozesse*, VVW Karlsruhe.
2. Aigner, M. (1993). *Diskrete Mathematik*, Vieweg, Germany.
3. Anderson, P., Borgan, Ø., Gill, R. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*, Springer, New York.
4. Aven, T. and Jensen, U. (1998). *Stochastic Models in Reliability*, Springer, New York.
5. Bremaud, P. (1981). *Point Processes and Queues*, Springer, New York.
6. Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of Royal Statistical Society B*, **17**, 129–164.
7. Cramer, H. (1969). On streams of random events. *Skand. AktuariTidskr. Suppl.*, **85**, 13–23.
8. Esary, J. D., Marshall, A. W. and Proshan, F. (1973). Shock models and wear processes. *The annals of probability*, **1**, 627–649.

9. Feng, W., Adachi, K. and Kowada, M. (1994). Optimal replacement under additive damage in a Poisson random environment, *Communications in statistics-stochastic models*, **10**, 679–700.
10. Grandell, J. (1991). *Aspects of Risk Theory*, Springer, New York.
11. Grandell, J. (1997). *Mixed Poisson Processes*, Chapman & Hall, London.
12. Kahle, W. and Wendt, H. (2006). Statistical Analysis of Some Parametric Degradation Models. In: Nikulin, M.S., Commenges, D., Huber, C. (eds), *Probability, Statistics and Modelling in Public Health*, Springer Science+Business Media, New York, pp. 266–279.
13. König, D. and Schmidt, V. (1992). *Zufällige Punktprozesse*, B.G. Teubner, Stuttgart.
14. Last, G. and Brandt, A. (1995). *Marked Point Processes on the Real Line*, Springer, New York.
15. Schröter, K. J. (1995). *Verfahren zur Approximation der Gesamtschadenverteilung*, VVW Karlsruhe.
16. Shaked, M. (1983). Wear and damage processes from shock models, In *Reliability theory and models*, Proc. Symp., Charlotte/N.C., 43–64.
17. Sobczyk, K. (1987). Stochastic models for fatigue damage of materials. *Advances in Applied Probability*, **19**, 652–673.
18. Voinov, V. G., Nikulin, M. S. (1996). *Unbiased Estimators and Their Applications: Volume 2: Multivariate Case*, Kluwer Academic Publishers, Hingham, MA.
19. Wendt, H. (1999). Parameterschätzungen für eine Klasse doppelt-stochastischer Poisson Prozesse bei unterschiedlichen Beobachtungsinformationen. *PhD-Thesis*.
20. Wendt, H. and Kahle, W. (2004). On a Cumulative Damage Process and Resulting First Passage Times. *Applied Stochastic Models in Business and Industry*, **20**, 17–26
21. Wendt, H. and Kahle, W. (2004) On Parameter Estimation for a Position-Dependent Marking of a Doubly Stochastic Poisson Process. In Nikulin et al. (eds), *Parametric and Semiparametric Models with Applications to Reliability, Survival Analysis, and Quality of Life*. Birkhauser Book Series, Statistics for Industry and Technology, 473–468.

Poisson Approximation of Processes with Locally Independent Increments and Semi-Markov Switching – Toward Application in Reliability

V.S. Koroliuk¹, N. Limnios², and I.V. Samoilenko¹

¹ Institute of Mathematics, Ukrainian National Academy of Science, Kiev, Ukraine

² Laboratoire de Mathématiques Appliquées, Université de Technologie de Compiègne, France

Abstract: In this chapter, the weak convergence of additive functionals of processes with locally independent increments and with semi-Markov switching in the scheme of Poisson approximation is investigated. Singular perturbation problem for the compensating operator of the extended Markov renewal process is used to prove the relative compactness. This approach can be used in applications and especially in shock and degradation in random environment arising in reliability.

Keywords and phrases: Poisson approximation, semimartingale, semi-Markov process, locally independent increments process, piecewise deterministic Markov process, weak convergence, singular perturbation, degradation, reliability

7.1 Introduction

Poisson approximation is still an active area of research in several theoretical and applied directions. Several recent works on this topic can be found in the literature: we can find the classical approach in [2–4] and the functional approach in [13, 14, 11, 17].

In particular in [13, 14] the stochastic additive functional

$$\xi(t) = \xi_0 + \int_0^t \eta(ds; x(s)), \quad t \geq 0, \quad (7.1)$$

of a jump Markov process with locally independent increments (PLII) [13] $\eta(t; \cdot), t \geq 0$ (also known as a piecewise deterministic Markov process – PDMP [7]), perturbed by the jump semi-Markov process $x(t), t \geq 0$, has been studied. The process (7.1) is studied in a (functional) Poisson approximation scheme, within an ad hoc timescaling as we can see below (7.2).

The process (7.1) includes as particular cases dynamical systems, additive functionals, compound Poisson process, and increment process in random environment [13]. The

latter one is important in reliability applications as, for example, in shock and degradation [1, 5, 8, 10, 12].

For example, the problem of shocks in a semi-Markov random environment is defined by the following stochastic process

$$\sum_{k=1}^{\nu(t)} \alpha_k(x_k), \quad t \geq 0,$$

where $\nu(t), t \geq 0$, is a counting process of shocks arriving on the system (here are the jump times of the semi-Markov process $x(t)$) and the positive random variables $\alpha_k(x)$ are the magnitudes of the shocks under the fixed environment condition $x_k = x$. The process $x_n, n \geq 0$, is the embedded Markov chain of the semi-Markov process $x(t), t \geq 0$ (see below).

In the Poisson approximation scheme, the jump values of the stochastic system are split into two parts: a small jump taking values with probabilities close to one and a big jump taken values with probabilities tending to zero together with the series parameter $\varepsilon \rightarrow 0$. So, in the Poisson approximation principle the probabilities (or intensities) of jumps are normalized by the series parameter $\varepsilon > 0$. Hence the time-scaled family of processes $\xi^\varepsilon(t), t \geq 0, \varepsilon > 0$ has to be considered.

However, the method used here to prove the weak convergence is quite different from the method proposed by other authors [9–22]: the main point is to prove convergence of predictable characteristics of semimartingales which are integral functionals of some switching semi-Markov processes.

We propose to study functionals of PLII [13] using a combination of two methods. The one based on semimartingales theory is combined with a solution of singular perturbation problem instead of ergodic theorem. So, the method includes two steps.

In the first step we prove the relative compactness of the semimartingales representation of the family $\xi^\varepsilon, \varepsilon > 0$, by proving the following two facts [9]:

$$\lim_{c \rightarrow \infty} \sup_{\varepsilon \leq \varepsilon_0} \mathbf{P}\{\sup_{t \leq T} |\xi^\varepsilon(t)| > c\} = 0,$$

known as the compact containment condition, and

$$\mathbf{E}|\xi^\varepsilon(t) - \xi^\varepsilon(s)|^2 \leq k|t - s|,$$

for some positive constant k .

In the second step we prove convergence of predictable characteristics of the semimartingales, which are integral functionals of the form:

$$\int_0^t a(\xi^\varepsilon(s), x^\varepsilon(s)) ds,$$

by using singular perturbation technique as presented in [13].

Finally, we apply Theorem IX.3.27 from [11] in order to prove the weak convergence of semimartingale.

The chapter is organized as follows. In Section 7.2 we present the timescaled additive functional (7.1), the PLII and the switching semi-Markov process. In the same section we present the main results of Poisson approximation. In Section 7.3 we present the proof of the theorem.

7.2 Main Results

Let us consider the space \mathbf{R}^d endowed with a norm $|\cdot|$ ($d \geq 1$) and (E, \mathcal{E}) , a *standard phase space* (i.e., E is a Polish space and \mathcal{E} its Borel σ -algebra). For a vector $v \in \mathbf{R}^d$ and a matrix $c \in \mathbf{R}^{d \times d}$, v^* and c^* denote their transpose, respectively. Let $C_3(\mathbf{R}^d)$ be a measure-determining class of real-valued bounded functions, such that $g(u)/|u|^2 \rightarrow 0$, as $|u| \rightarrow 0$ for $g \in C_3(\mathbf{R}^d)$ (see [11, 13]).

The additive functional $\xi^\varepsilon(t)$, $t \geq 0$, $\varepsilon > 0$, on \mathbf{R}^d in the series scheme with small series parameter $\varepsilon \rightarrow 0$ ($\varepsilon > 0$) is defined by the stochastic additive functional ([13], Section 3.3.1)

$$\xi^\varepsilon(t) = \xi_0^\varepsilon + \int_0^t \eta^\varepsilon(ds; x(s/\varepsilon)). \quad (7.2)$$

The family of the Markov jump processes with *locally independent increments* $\eta_\varepsilon(t; x)$, $t \geq 0$, $x \in E$, on \mathbf{R}^d is defined by the generators ([13], Section 3.3.1) (see also [14])

$$\mathbf{I}^\varepsilon(x)\varphi(u) = \int_{\mathbf{R}^d} [\varphi(u+v) - \varphi(u)] \Gamma^\varepsilon(u, dv; x), \quad x \in E, \quad (7.3)$$

or equivalently

$$\begin{aligned} \mathbf{I}^\varepsilon(x)\varphi(u) &= b_\varepsilon(u; x)\varphi'(u) + \frac{1}{2}c_\varepsilon(u; x)\varphi''(u) + \int_{\mathbf{R}^d} [\varphi(u+v) - \varphi(u) - v\varphi'(u) \\ &\quad - \frac{v^2}{2}\varphi''(u)] \Gamma^\varepsilon(u, dv; x), \end{aligned}$$

where $b_\varepsilon(u; x) = \int_{\mathbf{R}^d} v \Gamma^\varepsilon(u, dv; x)$, $c_\varepsilon(u; x) = \int_{\mathbf{R}^d} vv^* \Gamma^\varepsilon(u, dv; x)$, and $\Gamma^\varepsilon(u, dv; x)$ is the intensity kernel.

The switching semi-Markov process $x(t)$, $t \geq 0$ on the standard phase space (E, \mathcal{E}) is defined by the semi-Markov kernel

$$Q(x, B, t) = P(x, B)F_x(t), \quad x \in E, B \in \mathcal{E}, t \geq 0,$$

which defines the associated Markov renewal process $x_n, \tau_n, n \geq 0$:

$$Q(x, B, t) = P(x_{n+1} \in B, \theta_{n+1} \leq t | x_n = x) = P(x_{n+1} \in B | x_n = x)P(\theta_{n+1} \leq t | x_n = x).$$

Let the following conditions hold.

C1: The semi-Markov process $x(t)$, $t \geq 0$ is uniformly ergodic with the stationary distribution

$$\begin{aligned} \pi(dx)q(x) &= q\rho(dx), q(x) := 1/m(x), q := 1/m, \\ m(x) &:= \mathbf{E}\theta_x = \int_0^\infty \bar{F}_x(t)dt, m := \int_E \rho(dx)m(x), \\ \rho(B) &= \int_E \rho(dx)P(x, B), \rho(E) = 1. \end{aligned}$$

C2: *Poisson approximation.* The family of processes with locally independent increments $\eta^\varepsilon(t; x), t \geq 0, x \in E$ satisfies the Poisson approximation conditions ([13], Section 7.2.3).

PA1: Approximation of the mean values:

$$b_\varepsilon(u; x) = \int_{\mathbf{R}^d} v \Gamma^\varepsilon(u, dv; x) = \varepsilon[b(u; x) + \theta_b^\varepsilon(u; x)],$$

and

$$c_\varepsilon(u; x) = \int_{\mathbf{R}^d} vv^* \Gamma^\varepsilon(u, dv; x) = \varepsilon[c(u; x) + \theta_c^\varepsilon(u; x)].$$

PA2: Poisson approximation condition for intensity kernel

$$\Gamma_g^\varepsilon(u; x) = \int_{\mathbf{R}^d} g(v) \Gamma^\varepsilon(u, dv; x) = \varepsilon[\Gamma_g(u; x) + \theta_g^\varepsilon(u; x)]$$

for all $g \in C_3(\mathbf{R}^d)$, and the kernel $\Gamma_g(u; x)$ is bounded for all $g \in C_3(\mathbf{R}^d)$, that is,

$$|\Gamma_g(u; x)| \leq \Gamma_g \quad (\text{a constant depending on } g).$$

The above negligible terms $\theta_a^\varepsilon, \theta_b^\varepsilon, \theta_c^\varepsilon$ satisfy the condition

$$\sup_{x \in E} |\theta^\varepsilon(u; x)| \rightarrow 0, \quad \varepsilon \rightarrow 0.$$

In addition, the following conditions are used:

C3: *Uniform square integrability:*

$$\lim_{c \rightarrow \infty} \sup_{x \in E} \int_{|v| > c} vv^* \Gamma(u, dv; x) = 0,$$

where the kernel $\Gamma(u, dv; x)$ is defined on the class $C_3(\mathbf{R}^d)$ by the relation

$$\Gamma_g(u; x) = \int_{\mathbf{R}^d} g(v) \Gamma(u, dv; x), \quad g \in C_3(\mathbf{R}^d).$$

C4: Cramér's condition

$$\sup_{x \in E} \int_0^\infty e^{ht} F_x(dt) \leq H < +\infty, h > 0.$$

C5: *Linear growth:* there exists a positive constant L such that

$$|b(u; x)| \leq L(1 + |u|) \quad \text{and} \quad |c(u; x)| \leq L(1 + |u|^2),$$

and for any real-valued non-negative function $f(v), v \in \mathbf{R}^d$, such that $\int_{\mathbf{R}^d \setminus \{0\}} (1 + f(v)) |v|^2 dv < \infty$, we have

$$|A(u, v; x)| \leq Lf(v)(1 + |u|),$$

where $A(u, v; x)$ is the Radon–Nikodym derivative of $\Gamma(u, B; x)$ with respect to Lebesgue measure dv in \mathbf{R}^d , that is,

$$\Gamma(u, dv; x) = A(u, v; x) dv.$$

The main result of our work is the following.

Theorem 1. *Under conditions **C1** – **C5** the weak convergence*

$$\xi^\varepsilon(t) \Rightarrow \xi^0(t), \quad \varepsilon \rightarrow 0$$

takes place.

The limit process $\xi^0(t), t \geq 0$ is defined by the generator

$$\widehat{\Gamma}\varphi(u) = \widehat{b}(u)\varphi'(u) + \int_{\mathbf{R}^d} [\varphi(u+v) - \varphi(u) - v\varphi'(u)]\widehat{\Gamma}(u, dv), \quad (7.4)$$

where the average deterministic drift is defined by

$$\widehat{b}(u) = \int_E \pi(dx)b(u; x),$$

and the average intensity kernel is defined by

$$\widehat{\Gamma}(u, dv) = \int_E \pi(dx)\Gamma(u, dv; x).$$

Remark 1. The limit process $\xi^0(t), t \geq 0$ is a PLII (see, e.g., [13]) or a PDMP (see, e.g., [7]). The generator (7.4) can also be written as follows:

$$\widehat{\Gamma}\varphi(u) = \widehat{b}_0(u)\varphi'(u) + \int_{\mathbf{R}^d} [\varphi(u+v) - \varphi(u)]\widehat{\Gamma}(u, dv),$$

where $\widehat{b}_0(u) = \widehat{b}(u) - \int_{\mathbf{R}^d} v\widehat{\Gamma}(u, dv)$.

In what follows, we give a corollary of the above theorem where the limit process is a compound Poisson process.

Corollary 1. *In the particular case of the following Poisson approximation conditions:*

PA1': Here the mean value condition is

$$b_\varepsilon(u; x) = \int_{\mathbf{R}^d} v\Gamma^\varepsilon(dv; x) = \varepsilon[b(x) + \theta_b^\varepsilon(u; x)],$$

and

$$c_\varepsilon(u; x) = \int_{\mathbf{R}^d} vv^*\Gamma^\varepsilon(dv; x) = \varepsilon[c(x) + \theta_c^\varepsilon(u; x)].$$

PA2': For intensity kernel we have

$$\Gamma_g^\varepsilon(u; x) = \int_{\mathbf{R}^d} g(v)\Gamma^\varepsilon(u, dv; x) = \varepsilon[\Gamma_g(x) + \theta_g^\varepsilon(u; x)],$$

and the kernel $\Gamma_g(x)$ is bounded for all $g \in C_3(\mathbf{R}^d)$, that is,

$$|\Gamma_g(x)| \leq \Gamma_g \quad (\text{a constant depending on } g).$$

And the additional balance condition

PA3:

$$\int_{\mathbf{R}^d} v \Gamma(dv) = \int_E \pi(dx) b(x), \quad \Gamma(dv) = \int_{\mathbf{R}^d} \pi(dx) \Gamma(dv; x),$$

the limit process $\xi^0(t), t \geq 0$ is a compound Poisson process

$$\xi^0(t) = u + \sum_{k=1}^{\nu(t)} \alpha_k, \quad t \geq 0,$$

defined by the generator

$$\mathbf{G}\varphi(u) = \int [\varphi(u+v) - \varphi(u)] \Gamma(dv),$$

where

$$\Gamma(dv) = \int_E \pi(dx) \Gamma(dv; x), \quad \Gamma_g(x) = \int g(v) \Gamma(dv; x).$$

The sequence of random variables $\alpha_k, k = 1, 2, \dots$ is i.i.d. with common distribution function $\mathbf{P}(\alpha_k \in dv) = \Gamma(dv)/\Lambda, \quad \Lambda = \Gamma(\mathbf{R}^d)$. The time-homogeneous Poisson process $\nu(t), t \geq 0$ is defined by its intensity: $\Lambda > 0$.

7.3 Proof of Theorem 1

The proof of Theorem 1 is based on the semimartingale representation of the additive functional process (7.2). According to Theorems 6.27 and 7.16 [6] the predictable characteristics of the semimartingale (7.2) have the following representations:

- $B^\varepsilon(t) = \varepsilon^{-1} \int_0^t b_\varepsilon(\xi^\varepsilon(s); x_s^\varepsilon) ds = \int_0^t b(\xi^\varepsilon(s); x_s^\varepsilon) ds + \theta_b^\varepsilon,$
 - $C^\varepsilon(t) = \varepsilon^{-1} \int_0^t c_\varepsilon(\xi^\varepsilon(s); x_s^\varepsilon) ds = \int_0^t c(\xi^\varepsilon(s); x_s^\varepsilon) ds + \theta_c^\varepsilon,$
 - $\Gamma^\varepsilon(t) = \varepsilon^{-1} \int_0^t \int_{\mathbf{R}^d} g(v) \Gamma^\varepsilon(\xi^\varepsilon(s), dv; x_s^\varepsilon) ds = \int_0^t \int_{\mathbf{R}^d} g(v) \Gamma(\xi^\varepsilon(s), dv; x_s^\varepsilon) ds + \theta_g^\varepsilon,$
- where $x_t^\varepsilon := x(t/\varepsilon), t \geq 0$, and $\sup_{x \in E} |\theta^\varepsilon| \rightarrow 0, \quad \varepsilon \rightarrow 0$.

The jump martingale part of the semimartingale (7.2) is represented as follows:

$$\mu^\varepsilon(t) = \int_0^t \int_{\mathbf{R}^d} v [\mu^\varepsilon(ds, dv; x_s^\varepsilon) - \Gamma^\varepsilon(\xi^\varepsilon(s), dv; x_s^\varepsilon) ds].$$

Here $\mu^\varepsilon(ds, dv; x), x \in E$ is the family of counting measures with characteristics

$$\mathbf{E} \mu^\varepsilon(ds, dv; x) = \Gamma^\varepsilon(u, dv; x) ds.$$

We split the proof of Theorem 1 in to the following two steps.

STEP 1. In this step we establish the relative compactness of the family of processes $\xi^\varepsilon(t), t \geq 0, \varepsilon > 0$ by using the approach developed in [16]. Let us remind that the space of all probability measures defined on the standard space (E, \mathcal{E}) is also a Polish space; so the relative compactness and tightness are equivalent.

First, we need the following lemma.

Lemma 1. *Under assumption **C5** there exists a constant $k > 0$ independent of ε and dependent on T , such that*

$$\mathbf{E} \sup_{t \leq T} |\xi^\varepsilon(t)|^2 \leq k_T.$$

Corollary 2. *Under assumption **C5**, the following compact containment condition (CCC) holds*

$$\lim_{c \rightarrow \infty} \sup_{\varepsilon \leq \varepsilon_0} \mathbf{P}\{\sup_{t \leq T} |\xi^\varepsilon(t)| > c\} = 0.$$

PROOF. The proof of this corollary follows from Kolmogorov's inequality.

PROOF OF LEMMA 1. (following [16]). The semimartingale (7.2) has the following representation:

$$\xi^\varepsilon(t) = u + A_t^\varepsilon + M_t^\varepsilon, \quad (7.5)$$

where $u = \xi^\varepsilon(0)$; A_t^ε is the predictable drift (see [4]):

$$A_t^\varepsilon = \int_0^t b(\xi^\varepsilon(s), x_s^\varepsilon) ds + \int_0^t \int_{|v|>1} v \Gamma(\xi^\varepsilon(s), dv; x_s^\varepsilon) ds,$$

and M_t^ε is the locally square integrable martingale

$$M_t^\varepsilon = \int_0^t c(\xi^\varepsilon(s); x_s^\varepsilon) dw_s + \int_0^t \int_{|v| \leq 1} v [\mu^\varepsilon(ds, dv; x_s^\varepsilon) - \Gamma^\varepsilon(\xi^\varepsilon(s), dv; x_s^\varepsilon) ds],$$

where $w_t, t \geq 0$ is a standard Wiener process.

For a process $y(t), t \geq 0$, let us define the process $y_t^\dagger = \sup_{s \leq t} |y(s)|$, then from (7.5) we have

$$((\xi_t^\varepsilon)^\dagger)^2 \leq 3[u^2 + ((A_t^\varepsilon)^\dagger)^2 + ((M_t^\varepsilon)^\dagger)^2]. \quad (7.6)$$

Condition **C5** implies that

$$\begin{aligned} (A_t^\varepsilon)^\dagger &\leq L \int_0^t (1 + (\xi_s^\varepsilon)^\dagger) ds + \int_0^t \int_{|v|>1} |v| f(v) (1 + (\xi_s^\varepsilon)^\dagger) dv ds \\ &\leq L(1 + r_1) \int_0^t (1 + (\xi_s^\varepsilon)^\dagger) ds, \end{aligned} \quad (7.7)$$

where $r_1 = \int_{\mathbf{R}^d \setminus \{0\}} |v|^2 f(v) dv$.

Now, by Doob's inequality (see, e.g., [17, Theorem 1.9.2]),

$$\mathbf{E}((M_t^\varepsilon)^\dagger)^2 \leq 4 |\mathbf{E}\langle M^\varepsilon \rangle_t|,$$

and condition **C5** we obtain

$$\begin{aligned} |\langle M^\varepsilon \rangle_t| &= \left| \int_0^t c(\xi^\varepsilon(s); x_s^\varepsilon) c^*(\xi^\varepsilon(s); x_s^\varepsilon) ds + \int_0^t \int_{\mathbf{R}^d \setminus \{0\}} vv^* \Gamma^\varepsilon(\xi^\varepsilon(s), dv; x_s^\varepsilon) ds \right| \\ &\leq 2L(1 + r_1) \int_0^t [1 + ((\xi_s^\varepsilon)^\dagger)^2] ds. \end{aligned} \quad (7.8)$$

Inequalities (7.6), (7.7) and (7.8) and Cauchy–Bunyakovsky–Schwarz inequality $([\int_0^t \varphi(s)ds]^2 \leq t \int_0^t \varphi^2(s)ds)$ imply

$$\mathbf{E}((\xi_t^\varepsilon)^\dagger)^2 \leq k_1 + k_2 \int_0^t \mathbf{E}((\xi_s^\varepsilon)^\dagger)^2 ds,$$

where k_1 and k_2 are positive constants independent of ε .

By Gronwall inequality (see, e.g., [9, p. 498]), we obtain

$$\mathbf{E}((\xi_t^\varepsilon)^\dagger)^2 \leq k_1 \exp(k_2 t).$$

Hence the lemma is proved.

Remark 2. Another way to prove CCC is proposed in [13, Theorem 8.10] (see also [9, 22]). They use the function $\varphi(u) = \sqrt{1+u^2}$ and prove corollary for $\varphi(\xi_t^\varepsilon)$ by applying the martingale characterization of the Markov process.

This can be easily proved due to specific properties of $\varphi(u)$.

Lemma 2. *Under assumption **C5** there exists a constant $k > 0$ independent of ε such that*

$$\mathbf{E}|\xi^\varepsilon(t) - \xi^\varepsilon(s)|^2 \leq k|t - s|.$$

PROOF. In the same manner with (7.6), we may write

$$|\xi^\varepsilon(t) - \xi^\varepsilon(s)|^2 \leq 2|A_t^\varepsilon - A_s^\varepsilon|^2 + 2|M_t^\varepsilon - M_s^\varepsilon|^2.$$

By using Doob's inequality, we obtain

$$\mathbf{E}|\xi^\varepsilon(t) - \xi^\varepsilon(s)|^2 \leq 2\mathbf{E}\{|A_t^\varepsilon - A_s^\varepsilon|^2 + 8|\langle M^\varepsilon \rangle_t - \langle M^\varepsilon \rangle_s|\}.$$

Now (7.7), (7.8), and assumption **C5** imply

$$|A_t^\varepsilon - A_s^\varepsilon|^2 + 8|\langle M^\varepsilon \rangle_t - \langle M^\varepsilon \rangle_s| \leq k_3[1 + ((\xi_T^\varepsilon)^\dagger)^2]|t - s|,$$

where k_3 is a positive constant independent of ε .

From the last inequality and Lemma 1 the desired conclusion is obtained.

The conditions proved in Corollary 2 and Lemma 2 are necessary and sufficient for the compactness of the family of processes $\xi^\varepsilon(t), t \geq 0, \varepsilon > 0$.

STEP 2. The next step of proof concerns the convergence of the predictable characteristics. To do this, we apply the results of Sections 3.2–3.3 in [13] and the following theorem. $C_0^2(\mathbf{R}^d \times E)$ is the space of real-valued twice continuously differentiable functions on the first argument defined on $\mathbf{R}^d \times E$ and vanishing at infinity, and $C(\mathbf{R}^d \times E)$ is the space of real-valued continuous bounded functions defined on $\mathbf{R}^d \times E$.

Theorem 2. ([13, Theorem 6.3]) *Let the following conditions hold for a family of coupled Markov processes $\xi^\varepsilon(t), x^\varepsilon(t), t \geq 0, \varepsilon > 0$:*

CD1: There exists a family of test functions $\varphi^\varepsilon(u, x)$ in $C_0^2(\mathbf{R}^d \times E)$, such that

$$\lim_{\varepsilon \rightarrow 0} \varphi^\varepsilon(u, x) = \varphi(u),$$

uniformly on u, x .

CD2: The following convergence holds

$$\lim_{\varepsilon \rightarrow 0} \mathbf{L}^\varepsilon \varphi^\varepsilon(u, x) = \mathbf{L} \varphi(u)$$

uniformly on u, x . The family of functions $\mathbf{L}^\varepsilon \varphi^\varepsilon, \varepsilon > 0$ is uniformly bounded, and $\mathbf{L} \varphi(u)$ and $\mathbf{L}^\varepsilon \varphi^\varepsilon$ belong to $C(\mathbf{R}^d \times E)$.

CD3: The quadratic characteristics of the martingales that characterize a coupled Markov process $\xi^\varepsilon(t), x^\varepsilon(t), t \geq 0, \varepsilon > 0$ have the representation $\langle \mu^\varepsilon \rangle_t = \int_0^t \zeta^\varepsilon(s) ds$, where the random functions $\zeta^\varepsilon, \varepsilon > 0$, satisfy the condition

$$\sup_{0 \leq s \leq T} \mathbf{E} |\zeta^\varepsilon(s)| \leq c < +\infty.$$

CD4: The convergence of the initial values holds and

$$\sup_{\varepsilon > 0} \mathbf{E} |\zeta^\varepsilon(0)| \leq C < +\infty.$$

Then the weak convergence

$$\xi^\varepsilon(t) \Rightarrow \xi(t), \quad \varepsilon \rightarrow 0$$

takes place.

In the sequel the process $A^\varepsilon(t)$ will denote one of main parts of the above predictable characteristics $B^\varepsilon(t), C^\varepsilon(t), \Gamma^\varepsilon(t)$.

The extended Markov renewal process is considered as a four-component Markov chain

$$A_n^\varepsilon = A_0^\varepsilon(\tau_n^\varepsilon), \xi_n^\varepsilon, x_n^\varepsilon, \tau_n^\varepsilon, n \geq 0, \quad (7.9)$$

where $x_n^\varepsilon = x^\varepsilon(\tau_n^\varepsilon), x^\varepsilon(t) := x(t/\varepsilon), \xi_n^\varepsilon = \xi^\varepsilon(\tau_n^\varepsilon)$ and $\tau_{n+1}^\varepsilon = \tau_n^\varepsilon + \varepsilon \theta_n^\varepsilon, n \geq 0$, and

$$P(\theta_{n+1}^\varepsilon \leq t | x_n^\varepsilon = x) = F_x(t) = P(\theta_x \leq t).$$

Definition 1. [23] *The compensating operator \mathbf{L} of the Markov renewal process (7.9) is defined by the following relation:*

$$\mathbf{L}^\varepsilon \varphi(A_0^\varepsilon, \xi_0^\varepsilon, x_0, \tau_0) = q(x_0) \mathbf{E}[\varphi(A_1^\varepsilon, \xi_1^\varepsilon, x_1, \tau_1) - \varphi(A_0^\varepsilon, \xi_0^\varepsilon, x_0, \tau_0) | \mathcal{F}_0],$$

where

$$\mathcal{F}_t := \sigma(A^\varepsilon(s), \xi^\varepsilon(s), x^\varepsilon(s), \tau^\varepsilon(s); 0 \leq s \leq t).$$

Using Lemma 7.5 from [13] we obtain that the compensating operator of the extended Markov renewal process from Definition 1 can be defined by the relation (see also Section 2.8 in [13])

$$\begin{aligned} \mathbf{L}^\varepsilon \varphi(u, v; x, t) \\ = \left[\int_0^\infty F_x(ds) A_{\varepsilon s}(u; x) \Gamma_{\varepsilon s}(x) \int_E P(x, dy) \varphi(u, v; y, t + \varepsilon s) - \varphi(u, v; x, t) \right] / \varepsilon m(x). \end{aligned}$$

According to ([13], Lemma 7.9) the compensating operator of the extended Markov renewal process applied to functions $\varphi \in C^2(\mathbf{R}^d \times \mathbf{R}^d) \times \mathbf{B}(E)$ has the asymptotic representation

$$\begin{aligned} \mathbf{L}^\varepsilon \varphi(u, v; x) = \varepsilon^{-1} \mathbf{Q} \varphi(\cdot, \cdot; x) + \mathbf{A}(u, x) P \varphi(\cdot, v; \cdot) + \mathbf{\Gamma}^\varepsilon(x) P \varphi(u, \cdot; \cdot) \\ + \varepsilon \theta^\varepsilon(x) P \varphi(u, v; x), \end{aligned} \quad (7.10)$$

where P is a transition operator associated to $P(x, B)$, $\mathbf{\Gamma}^\varepsilon(x)$ is defined in (3),

$$\mathbf{Q} \varphi(x) = q(x) \int_E P(x, dy) [\varphi(\cdot; y) - \varphi(\cdot; x)],$$

and

$$\mathbf{A}(u; x) \varphi(v) = a(u; x) \varphi'(v).$$

In order to prove the convergence of predictable characteristics, it is sufficient to study the action of the generator \mathbf{L}^ε on test functions of two variables $\varphi(v, x)$.

Thus, it has the representation

$$\mathbf{L}^\varepsilon \varphi(v; x) = [\varepsilon^{-1} \mathbf{Q} + \mathbf{A}(u, x) P + \varepsilon \theta^\varepsilon(x) P] \varphi(v; x). \quad (7.11)$$

The solution of the singular perturbation problem at the test functions $\varphi^\varepsilon(v; x) = \varphi(v) + \varepsilon \varphi_1(v; x)$ in the form $\mathbf{L}^\varepsilon \varphi^\varepsilon = \widehat{\mathbf{L}} \varphi + \theta^\varepsilon \varphi$ can be found in the same manner with Proposition 5.1 in [13]. That is

$$\widehat{\mathbf{L}} = \widehat{\mathbf{A}}(u), \quad \widehat{\mathbf{A}}(u) := \Pi \mathbf{A}(u; x) \Pi, \quad (7.12)$$

or in other words $\widehat{\mathbf{A}}(u) = \int_E \pi(dx) \mathbf{A}(u, x)$. Thus $\widehat{\mathbf{A}}(u) \varphi(v) = \widehat{a}(u) \varphi'(v)$, where $\widehat{a}(u) := \int_E \pi(dx) a(u, x)$.

Now Theorem 2 can be applied.

We see from (7.10) and (7.12) that the solution of singular perturbation problem for $\mathbf{L}^\varepsilon \varphi^\varepsilon(u, v; x)$ satisfies the conditions **CD1**, **CD2**. Condition **CD3** of this theorem implies that the quadratic characteristics of the martingale corresponding to a coupled Markov process are relatively compact. The same result follows from the CCC (see Corollary 2 and Lemma 2) by [11]. Thus, the condition **CD3** follows from Corollary 2 and Lemma 2. As soon as $A^\varepsilon(0) \rightarrow A^0(0)$, $\xi^\varepsilon(0) \rightarrow \xi^0(0)$ we see that the condition **CD4** is also satisfied. Thus, all the conditions of Theorem 2 are satisfied, so the weak convergence $A^\varepsilon(t) \Rightarrow A^0(t)$ takes place.

The final step of the proof is achieved now by using Theorem IX.3.27 in [11]. Indeed all the conditions of this theorem are fulfilled.

As we mentioned, the square integrability condition 3.24 follows from CCC (see [11]). The strong majoration hypothesis is true with the majoration functions presented in the Condition **C5**. Condition **C5** implies the condition of big jumps for the last predictable measure of Theorem IX.3.27 in [11]. Conditions iv and v of Theorem IX.3.27 [11] are obviously fulfilled.

The weak convergence of predictable characteristics is proved by solving the singularly perturbation problem for the generator (7.11).

The last condition (3.29) of Theorem IX.3.27 is also fulfilled due to CCC proved in Corollary 2 and Lemma 2. Thus, the weak convergence is true.

Substituting the operators $\mathbf{B}(t)$, $\mathbf{C}(t)$, $\mathbf{\Gamma}(t)$ instead of $\mathbf{A}(t)$ into (7.11) we see that the limit Markov process is characterized by the following predictable characteristics:

$$B^0(t) = \int_0^t \widehat{b}(\xi^0(s))ds, \quad C^0(t) = \int_0^t \widehat{c}(\xi^0(s))ds, \quad \Gamma^0(t) = \int_0^t \widehat{\Gamma}_g(\xi^0(s))ds.$$

So, the limit Markov process $\xi^0(t)$ can be expressed by the generator (7.4). Theorem 1 is proved.

Acknowledgments. The authors thank University of Bielefeld and personally Prof. Yu.G.Kondratiev for hospitality and financial support by DFG project 436 UKR 113/80/04-07.

References

1. Abdel-Hameed M., Nakhi Y. (1991). Optimal replacement and maintenance of systems subject to semi-Markov damage, *Stochast. Process. Appl.*, 37, 141–160.
2. Barbour A.D., Holst L., Janson S. (1992). *Poisson Approximation*, Oxford University Press, Oxford.
3. Barbour A.D., Chen L.H.Y., (Eds.). (2005). *An introduction to Stein's method*. IMS Lecture Note Series Volume 4, World Scientific Press, Singapore.
4. Barbour A.D., Chen L.H.Y., (Eds.). (2005). *Stein's method and applications*. IMS Lecture Note Series Volume 5, World Scientific Press, Singapore.
5. Barlow R.E., Prochan F. (1975). *Statistical Theory of Reliability and Life Testing: Probability Models*, Holt, Rinehart and Winston, New York.
6. Çinlar E., Jacod J., Protter P. and Sharpe M.J. (1980). Semimartingale and Markov processes, *Z. Wahrschein. verw. Gebiete*, 54, 161–219.
7. Davis M.H.A. (1993). *Markov Models and Optimization*, Chapman & Hall, London.
8. Esary J.D., Marshal A.W., Proschan F. (1973). “Shock models and wear processes”, *Ann. Probab.*, 1, 627–649.
9. Ethier S.N., Kurtz T.G. (1986). *Markov Processes: Characterization and convergence*, J. Wiley, New York.
10. Gut A., Husler J. (2005). Realistic variation of shock models, *Statist. Probab. Lett.*, 74, 187–204.
11. Jacod J., Shiryaev A.N. (1987). *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin.

12. Kahle W., Wendt H. (2000). "Statistical analysis of damage processes", In *Recent Advances in Reliability Theory – Methodology, Practice and Inference*, N. Limnios, M.S. Nikulin (Eds.), Birkhauser, Boston, 199–212.
13. Koroliuk V.S., Limnios N. (2005). *Stochastic Systems in Merging Phase Space*, World Scientific Publishers, Singapore.
14. Korolyuk V.S., Limnios N. (2005). "Poisson approximation of increment processes with Markov switching", *Theor. Probab. Appl.*, 49(4), 629–644.
15. Kushner H.J. (1990). *Weak Convergence Methods and Singular Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston.
16. Liptser R.Sh. (1994). The Bogolubov averaging principle for semimartingales, *Proceedings of the Steklov Institute of Mathematics*, Moscow, No 4, 12 pages.
17. Liptser R.Sh., Shirayev A.N. (1989). *Theory of Martingales*, Kluwer Academic Publishers, Dordrecht, The Netherlands.
18. Stroock D.W., Varadhan S.R.S. (1979). *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin.
19. Skorokhod A.V. (1989). *Asymptotic Methods in the Theory of Stochastic Differential Equations*, AMS, vol. 78, Providence.
20. Skorokhod A.V., Hoppensteadt F.C., Salehi H. (2002). *Random Perturbation Method with Application in Science and Engineering*, Springer, Berlin.
21. Silvestrov D.S. (2004). *Limit Theorems for Randomly Stopped Stochastic Processes*. Series: Probability and its Applications, Springer, Berlin.
22. Sviridenko M.N. (1986). Martingale approach to limit theorems for semi-Markov processes, *Theor. Probab. Appl.*, 540–545.
23. Sviridenko M.N. (1998). Martingale characterization of limit distributions in the space of functions without discontinuities of second kind, *Math. Notes*, 43(5), 398–402.

On Some Shock Models of Degradation

Maxim Finkelstein¹ and Ji Hwan Cha²

¹ University of the Free State, Bloemfontein, Republic of South Africa and Max Planck Institute for Demographic Research, Rostock, Germany.

² Ewha Womans University, Seoul, Korea

Abstract: We discuss several shock models that describe univariate degradation. Poisson and renewal-type point processes of shocks are considered. We assume that degradation (damage) caused by each shock is accumulated by an overall degradation characteristic. The failure occurs when this accumulated degradation reaches the defined boundary. Asymptotic properties for the shot noise shock process and for the “imperfect repair-type” shock process are considered. The combined model when each shock results either in termination of the process or in the corresponding increment of degradation is discussed. A simple explicit formula for the probability of failure in this case is analyzed. Possible generalizations are suggested.

Keywords and phrases: Degradation, shocks, poisson process, shot noise process, random resource

8.1 Introduction

Stochastic degradation in engineering, ecological, and biological systems is naturally modeled by increasing (decreasing) stochastic processes. Denote the corresponding univariate stochastic process by $W_t, t \geq 0$. We will be interested in modeling stochastic degradation as such and in considering the first passage times, when this degradation reaches the predetermined level r . The latter interpretation can be useful for risk and safety assessment when r defines some critical safety region.

The simplest and the widely used approach in engineering practice for degradation modeling is the “path model”, whose stochastic properties are usually described either by the additive or by the multiplicative random variable:

$$W_t = d(t) + Z, \quad (8.1)$$

$$W_t = Zd(t), \quad (8.2)$$

where $d(t)$ is, for simplicity, a continuous, increasing function ($d(0) = 0, \lim_{t \rightarrow \infty} d(t) = \infty$) and Z is a nonnegative random variable with the Cdf $G(z)$. The failure occurs when

W_t reaches the predetermined threshold r . Denote by T_d the corresponding random variable with the Cdf $F(t)$. For instance, for the multiplicative model (8.2) it means that

$$\begin{aligned} F(t) &= \Pr(T_d \leq t) = \Pr(W_t \geq r) \\ &= \Pr\left(Z \geq \frac{r}{d(t)}\right) = 1 - G\left(\frac{r}{d(t)}\right), \end{aligned} \quad (8.3)$$

where $F(0)$ should be understood as the corresponding limit in the right-hand side of this equation.

Example 1. Inverse Weibull distribution. Let $d(t) = t$ and assume the Weibull law for $G(t)$, i.e., $G(z) = 1 - \exp\{-(\lambda z)^k\}$, $\lambda, k > 0$. Then, in accordance with (8.3),

$$F(t) = \exp\left\{-\left(\frac{\lambda r}{t}\right)^k\right\}.$$

Let the threshold R be now a random variable with the Cdf $F_0(r) = \Pr(R \leq r)$ and the wear process be described by the deterministic function $W(t)$ ($W(0) = 0, \lim_{t \rightarrow \infty} W(t) = \infty$). Then, similar to (8.3),

$$F(t) = \Pr(W(t) \geq R) = F_0(W(t)). \quad (8.4)$$

Therefore, Equation (8.4) can be interpreted as a general accelerated life model (ALM), whereas the function $W(t)$ defines the corresponding scale transformation. A generalization to W_t is as follows:

$$F(t) = E[F_0(W_t)], \quad (8.5)$$

where for convenience of notation the left-hand side is also denoted as $F(t)$.

Example 2. Exponential distributions. Let $F_0(r) = 1 - \exp\{-\lambda r\}$ and $W_t = Zd(t)$, where Z is also exponentially distributed with parameter μ . Direct integration in (8.5) gives

$$F(t) = 1 - \frac{\mu}{\mu + \lambda d(t)}.$$

As $d(0) = 0, \lim_{t \rightarrow \infty} d(t) = \infty$, this relationship defines a proper Cdf. We will come back to Equation (8.4) in Section 8.3.

A popular way of modeling monotone deterioration is via the corresponding process of shocks, which will be considered in the rest of this chapter. Shocks are usually defined as instantaneous external events resulting in an immediate failure or in a random amount of degradation. Therefore, a process of shocks can result in the accumulated degradation W_t . Numerous shock models were considered in the literature (see, e.g., [5] and references therein). Apart from degradation modeling, in various applications, shock models also present a convenient tool for analyzing certain nonparametric properties of distributions [1]. Explicit expressions for $E[\Pr(W_t \geq R)]$ can be usually obtained only for the simplified settings. In Section 8.3, we shall consider a specific but meaningful model that allows for an explicit relationship for this probability.

At many instances, it is convenient to model additive degradation via the corresponding compound process:

$$W_t = \sum_{i=1}^{N(t)} X_i, \quad (8.6)$$

where X_i , $i = 1, 2, \dots$, are nonnegative i.i.d. random variables and $N(t)$, $t \geq 0$, is an orderly point process independent of X_i , $i = 1, 2, \dots$ (we will also use notation N_t , $t \geq 0$ where appropriate). As it was mentioned, it is usually difficult to consider the corresponding first passage problems explicitly, whereas the central limit theorem obviously exists. Therefore, asymptotic (for $t \rightarrow \infty$) probability that the cumulative degradation (normalized) does not exceed a predetermined level usually can easily be obtained in this case.

In the next section, we will consider two asymptotic settings that present variations of the cumulative degradation model. In Section 8.4, the explicit exact solution for a meaningful specific problem will be discussed.

8.2 Asymptotic Properties for Two Special Shock Models

An important modification of the cumulative shock model (8.6) is given by the shot noise process [8]. In this model, the additive input of a shock of magnitude X_i in $[t, t+dt]$ in the cumulative degradation W_t is decreased in accordance with some decreasing (nonincreasing) response function $h(t)$. The corresponding cumulative degradation is defined as

$$W_t = \sum_{i=1}^{N(t)} X_i h(t - T_i), \quad (8.7)$$

where $T_1 < T_2 < T_3, \dots$ is the sequence of the corresponding arrival (waiting) times of i.i.d. shocks with magnitudes described by the generic random variable X . This setting has a lot of applications in electrical engineering, materials science, health sciences, and risk and safety analysis. For instance, cracks due to fatigue in some materials tend to close up after the material has borne a load which has caused the cracks to grow. Another example is the human heart muscle's tendency to heal after a heart attack [10].

Equivalently, definition (8.7) can be written as

$$\int_0^t X h(t - u) dN_u.$$

Assume that $E[X] < \infty$. As X_i , $i = 1, 2, \dots$, are independent of the point process N_t , $t \geq 0$,

$$E[W_t] = E[X] \int_0^t h(t - u) m(u) du, \quad (8.8)$$

where $m(u) = dE[N_u]/du$ is the rate of our point process of shocks. Specifically, for the homogeneous Poisson process, $m(u) = m$ and

$$E[W_t] = mE[X] \int_0^t h(u) du.$$

Therefore, distinct from Equation (8.6), the accumulated degradation in $[0, \infty)$, i.e., W_∞ , is finite when the corresponding integral of the response function is finite, which is usually the case in applications.

When the shock process is renewal with the renewal density function $m(u)$ and inter-arrival times with the mean $E[\tau]$, then applying the key renewal theorem [9] to Equation (8.8) results in the following asymptotic formula:

$$\lim_{t \rightarrow \infty} E[W_t] = \frac{E[X]}{E[\tau]} \int_0^\infty h(u) du.$$

It can be shown [7] that the central limit theorem holds, i.e., as $t \rightarrow \infty$:

$$\frac{W_t - E[W_t]}{(Var(W_t))^{1/2}} \longrightarrow N(0, 1), \quad (8.9)$$

where, e.g., for the Poisson process

$$Var(W_t) = mE[X^2] \int_0^t h^2(u) du.$$

Example 3. Exponential response function. Consider a specific case of the response function, i.e., $h(t) = \exp\{-\alpha t\}$, and the exponentially distributed (with parameter λ) increment X . In this case the problem can be solved explicitly. Straightforward calculations using the corresponding moment generating functions result [9] in the following limit distribution:

$$\lim_{t \rightarrow \infty} \Pr(W(t) \geq r) = \frac{\Gamma(m/\alpha, r\lambda)}{\Gamma(m/\lambda)}.$$

It is well known from the properties of the gamma distribution that as m/λ increases, it converges to the normal distribution; therefore, there is no contradiction between this result and asymptotic relationship (8.9).

The cumulative degradation in (8.6) obviously tends to ∞ (in expectation) as $t \rightarrow \infty$. On the other hand, the shot noise process model presents an option when expectation of the cumulative wear is finite as the number of shocks increases to ∞ . This occurs due to the decreasing response function $h(t)$. We will now consider another setting with the finite in the described sense cumulative wear.

Assume that shocks arrive in time in accordance with some orderly point process, but we are interested only in the amount of accumulated degradation after each shock and not in the corresponding arrival times. Let the magnitude of the first shock X_1 be described by the Cdf $F_1(w) = F_s(w)$. Assume that some repair (or anti-degradation) action decreases this degradation to qx_1 , where $0 < q < 1$ and x_1 is a realization of X_1 . Assume that the magnitude of the second shock is governed by the same generic Cdf $F_s(w)$, but the initial level of degradation is also taken into account in the following way: the degradation after the second shock is $qx_1 + x_2$, where x_2 is a realization of the random variable with the Cdf $F_2(w) = F_s(w|qx_1)$ and the analogue of the mean remaining lifetime in this case is

$$F_s(w|qx_1) = \frac{F_s(w + qx_1) - F_s(w)}{1 - F_s(qx_1)}.$$

The degradation after the third shock in this model is $q(qx_1 + x_2) + x_3$. Thus the sequence X_1, X_2, \dots is now dependent in the described way. This setting is the degradation-parameterized version of the Kijima Model 2 [6], which often describes imperfect repair in repairable systems. Our interest is in asymptotic properties of this shock model.

The interpreted in terms of our shock model result [4] states that the cumulative degradation after the i th shock, as $i \rightarrow \infty$, converges in distribution to a random variable with a finite mean. As in the case of the shot noise process, it means that given this distribution, the asymptotic probability that the cumulative degradation does not exceed a predetermined level can be easily obtained.

It should be noted that asymptotic models of this section are not of the “first passage type”. An explicit result for the first passage probability for the meaningful special case will be presented in the next section.

8.3 Terminating and “Accumulating” Shocks

In this section, for convenience, we will use a slightly different notation. Consider a degrading system in a baseline environment (without shocks) and denote its lifetime by R and the corresponding Cdf by $F(t)$. We interpret here R as some initial, random resource, which is “consumed” by a system in the process of its operation with rate 1 (the rate of degradation). Therefore, the age of our system in this case is equal to a calendar time t , and a failure occurs when this age reaches R .

Denote a system’s lifetime in the presence of shocks by $F_S(t)$. Assume that the i th shock, as in the Brown–Proschan model [2], causes immediate system’s failure with probability $p(t)$, but in contrast to this model, with probability $q(t) = 1 - p(t)$, it increases the degradation of a system on a random increment X_i , $i = 1, 2, \dots$. In accordance with this setting, the degradation of a system at time t is (compare with 8.6)

$$W_t = t + \sum_{i=0}^{N(t)} X_i,$$

where formally $X_0 = 0$ corresponds to the case when there are no shocks in $[0, t]$. The failure occurs when this random variable reaches the boundary R . Therefore,

$$\begin{aligned} \Pr(T_s > t | N(s), 0 \leq s \leq t, X_1, X_2, \dots, X_{N(t)}, R) \\ &= \prod_{i=0}^{N(t)} q(T_i) I(W_t \leq R) \\ &= \prod_{i=0}^{N(t)} q(T_i) I\left(\sum_{i=0}^{N(t)} X_i \leq R - t\right). \end{aligned}$$

This probability should be understood conditionally on the corresponding realizations of $N(t)$, X_i , and R . In order to proceed and finally obtain an explicit formula some additional assumptions should be made:

- $N(t)$, $t \geq 0$ is the nonhomogeneous Poisson process with rate $\nu(t)$.
- R is exponentially distributed with the corresponding failure rate λ .
- X_i , $i = 1, 2, \dots$, are i.i.d. random variables characterized by the moment generating function $M_X(t)$.
- $N(t)$, $t \geq 0$, X_i , $i = 1, 2, \dots$, and R are independent of each other.

Using these assumptions, R and X_i can be “integrated out” in the following way:

$$\begin{aligned}
 P(T_S > t | N(s), 0 \leq s \leq t, X_1, X_2, \dots, X_{N(t)}) \\
 &= \left(\prod_{i=1}^{N(t)} q(T_i) \right) \cdot \exp \left\{ - \int_0^{t + \sum_{i=1}^{N(t)} X_i} \lambda du \right\} \\
 &= \exp \left\{ -\lambda t - \lambda \sum_{i=1}^{N(t)} X_i + \sum_{i=1}^{N(t)} \ln q(T_i) \right\}. \tag{8.10}
 \end{aligned}$$

Thus we have

$$\begin{aligned}
 P(T_S > t | N(s), 0 \leq s \leq t) \\
 &= \exp\{-\lambda t\} \cdot \exp \left\{ \sum_{i=1}^{N(t)} \ln q(T_i) \right\} \cdot E \left[\exp \left\{ - \sum_{i=1}^{N(t)} \lambda X_i \right\} \right] \\
 &= \exp\{-\lambda t\} \cdot \exp \left\{ \sum_{i=1}^{N(t)} \ln q(T_i) \right\} \cdot (M_X(-\lambda))^{N(t)} \\
 &= \exp\{-\lambda t\} \cdot \exp \left\{ \sum_{i=1}^{N(t)} [\ln q(T_i) + \ln(M_X(-\lambda))] \right\}. \tag{8.11}
 \end{aligned}$$

“Integrating out the Poisson process” is more cumbersome. Omitting the technical details (which will be published elsewhere along with the other specific case of the fixed boundary), we arrive at the following relationship:

$$P(T_S > t) = \exp \left\{ -\lambda t - \int_0^t \nu(x) dx + M_X(-\lambda) \cdot \int_0^t q(x) \nu(x) dx \right\}. \tag{8.12}$$

Therefore, the failure rate function for our system is

$$\lambda_S(t) = \lambda + \left(1 - M_X(-\lambda) \cdot q(t) \right) \nu(t). \tag{8.13}$$

When X_i are distributed exponentially with mean μ , Equation (8.13) becomes

$$\lambda_S(t) = \lambda + \left(1 - \frac{q(t)}{\lambda\mu + 1} \right) \nu(t). \tag{8.14}$$

The qualitative analysis of these equations is really meaningful:

1. Equations (8.13) and (8.14) suggest that the failure rate $\lambda_S(t)$ can be interpreted as a failure rate of a series system with dependent (via R) components.
2. Consider the following special case. Let each shock in the equivalent to the Brown-Proschan model result either in the termination of the process with probability $p(t)$ or in the “no consequences event” with a complementary probability $q(t) = 1 - p(t)$. As $M_X(0) = 1$ for $\lambda = 0$, our failure rate $\lambda_S(t)$ reduces in this case to the corresponding failure rate in this simpler model, which is

$$\lambda_S(t) = p(t)\nu(t).$$

Note that the Brown-Proschan model was originally formulated [2] in terms of a repair model: with probability $p(t)$ the repair of a system with the failure rate $\nu(t)$ is perfect and with probability $1 - p(t)$ it is minimal. Then the time between two successive perfect repairs has the following Cdf:

$$1 - \exp\left\{-\int_0^t p(u)\nu(u)du\right\}.$$

3. When $\mu \rightarrow \infty$, from Equation (8.14), we get $\lambda_S(t) \rightarrow \lambda + \nu(t)$, which means that a failure occurs either in accordance with the baseline $F(t)$ or as a result of the first shock (competing risks). Note that, in accordance with the properties of Poisson processes, the rate $\nu(t)$ is equal to the failure rate which corresponds to the time to the first shock. Therefore, the two “components” of the described series system are asymptotically independent as $\mu \rightarrow \infty$.
4. When $\mu = 0$, which means that $X_i = 0, \forall i \geq 0$, Equation (8.14) turns to

$$\lambda_S(t) = \lambda + p(t)\nu(t).$$

Therefore, this specific case describes the series system with two independent components. The first component has the failure rate λ and the second component has the failure rate $p(t)\nu(t)$, which corresponds to the failure rate in the Brown-Proschan model.

5. It follows from (8.13) that when $p(t) = p$ and $\nu(t) = \nu$, the failure rate $\lambda_S(t)$ is also a constant which is a remarkable fact.
6. Let $q(t) = 1$ (no killing shocks) and $X_i, \forall i \geq 0$ be deterministic and equal to μ . Then $M_X(-\lambda) = \exp\{-\mu\lambda\}$ and Equation (8.13) turns to

$$\lambda_S(t) = \lambda + (1 - \exp\{-\mu\lambda\})\nu(t).$$

Note that this is an important special case of the “pure” cumulative degradation, which had not been considered before.

7. Equation (8.14) can be generalized to the case when X_i follow the gamma distribution with parameters α and μ , i.e.,

$$\lambda_S(t) = \lambda + \left(1 - \frac{q(t)}{(\lambda\mu + 1)^\alpha}\right)\nu(t).$$

8.4 Concluding Remarks

Shock models is a useful tool for describing degradation in technical systems caused by external factors. However, explicit formulas for the probability of failure due to accumulated degradation exist usually only in very simplified cases. In Section 8.3 we obtained the corresponding explicit formula for a meaningful special case. This special case can also be interpreted as a useful generalization of the Brown–Proschan model when instead of minimal repair, a repair action adding an increment of degradation is performed. An important simplification is due to the assumption that a random boundary is exponentially distributed. However, the method can be generalized to the case of the deterministic boundary.

Asymptotic approaches of Section 8.2 define asymptotic probabilities of exceeding the predetermined level for two important cases: the shot noise and the imperfect repair-type shock processes. It should be noted, however, that asymptotic models of this section are not of the “first passage type.”

References

1. Barlow, R. E. and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing*. Holt, Renerhart and Winston, New York.
2. Brown, M. and Proschan, F. (1983). Imperfect repair. *Journal of Applied Probability*, **20**, 851–862.
3. Finkelstein, M.S. (2008). Degradation and shock models. In: *Encyclopedia of Quantitative Risk Assessment and Analysis*, E. Melnick and B. Everitt (eds.). John Wiley & Sons, Chichester.
4. Finkelstein, M.S. (2007). On some ageing properties of general repair processes. *Journal of Applied Probability*, **44**, 506–513.
5. Gut, A. and Hustler, J. (2005). Realistic variation of shock models. *Statistics and Probability Letters*, **74**, 187–204.
6. Kijima, M. (1989). Some results for repairable systems with general repair. *Journal of Applied Probability*, **26**, 89–102.
7. Lund, R., McCormic, W. and Xiao, U. (2004). Limiting properties of Poisson shot noise processes. *Journal of Applied Probability*, **41**, 911–918.
8. Rice, J. (1977). On generalized shot noise. *Advances in Applied Probability*, **9**, 553–565.
9. Ross, S. (1996). *Stochastic Processes*. John Wiley & Sons, New York.
10. Singpurwalla, N. (1995). Survival in dynamic environment. *Statistical Science*, **10**, 86–103.

Degradation Models

The Wiener Process as a Degradation Model: Modeling and Parameter Estimation

Waltraud Kahle¹ and Axel Lehmann²

¹ Otto-von-Guericke-University, Magdeburg, Germany

² Leipzig University of Applied Sciences, Leipzig, Germany

Abstract: In this chapter we describe a simple degradation model based on the Wiener process. A failure occurs when the degradation reaches a given level for the first time. In this case, the time to failure is inverse Gaussian distributed. The parameters of the lifetime distribution can be estimated from observation of degradation only, from observation of failures, or from observation of both degradation increments and failure times. In the chapter, statistical methods for estimating the parameters of degradation processes for different data structures are developed and compared.

Keywords and phrases: Degradation process, failure time, likelihood function, maximum likelihood estimation, parameter estimation, process observation at discrete time points, Wiener process

9.1 Introduction

Investigating the reliability of products that are affected by degradation processes such as wear, corrosion, and crack growth, it is often necessary to observe the development of degradation processes which are characterized by a gradual drift of the mean value. The choice of the mathematical model for these degradation processes is based on the assumption of an additive accumulation of degradation with constant wear intensity. Regarding every degradation increment as an additive superposition of a large number of small effects, we can assume the degradation process to be normally distributed. Therefore, the degradation measure $Z(t)$ can be described by the following model [5]:

$$Z(t) = z_0 + \sigma W(t - t_0) + \mu \cdot (t - t_0), \quad t \geq t_0, \quad (9.1)$$

with z_0 – constant initial degradation ($z_0 \in \mathbb{R}$), t_0 – beginning of the degradation ($t_0 \in \mathbb{R}$), μ – drift parameter ($\mu \in \mathbb{R}$), σ – variance parameter ($\sigma > 0$), and $W(t)$ – standard Wiener process on $[0, \infty)$.

We suppose that a failure of a product will occur if the degradation process arrives at a certain boundary degradation level, which, in general, is unknown. For a given

boundary level h , the lifetime T_h of the product is then determined as the instant at which the degradation process $Z(t)$ exceeds the level h for the first time:

$$T_h = \inf\{t \geq t_0 : Z(t) \geq h\}. \quad (9.2)$$

It is well known that for $z_0 < h$ the lifetime T_h follows an inverse Gaussian distribution with the Lebesgue density function

$$f_{T_h}(t) = \frac{h - z_0}{\sqrt{2\pi\sigma^2(t - t_0)^3}} \exp\left(-\frac{(h - z_0 - \mu(t - t_0))^2}{2\sigma^2(t - t_0)}\right) I_{\{t > t_0\}}, \quad (9.3)$$

where $I_{\{\cdot\}}$ denotes an indicator variable, assuming the value 1 if the relation in brackets is true and the value 0 otherwise. Obviously, P^{T_h} is the Dirac measure ε_{t_0} concentrated at t_0 if $z_0 \geq h$.

The parameters of this distribution coincide with those of the degradation process. Note that the degradation level h can be an unknown parameter. To estimate the parameters of this model several observation schemes are possible.

1. The increments of the degradation process are observed.

In this case the likelihood function of the observation is a product of Gaussian densities. It is possible to estimate the parameters of the degradation process μ , σ^2 , x_0 , and t_0 , but it is impossible to estimate the degradation level h . Note that neither the same distance between observation points nor observation of all realizations at the same time points is required. Parameter estimations can be found if the observations are taken at individual time points in each realization.

2. The failure times are observed.

In this case we get the classical likelihood function from a sample of inverse Gaussian distributed random variables. From this likelihood function we can estimate the parameters μ , σ^2 , and t_0 . The degradation reserve $h - z_0$ must be given in advance.

3. The sample includes both process increments and failure times.

Assume that the sample consists of N failure times and $\sum_{i=1}^n m_i$ increments, being mutually independent. Then the likelihood function is the product of the likelihood functions mentioned above. In the case that failure times and process increments belong to the same realization of the degradation process, the likelihood function is more complicated. In this case the density of an observed process increment is a conditional density under the assumption that the process does not cross the degradation level in the current time interval. If the sample includes both process increments and failure times, then it is possible to estimate all parameters of the model.

These models and their parameter estimations are described in [5] and in [6]. A similar model and its application in medicine is described in [4]. Several generalizations of this simple model were given [7]. It is possible to include measurement errors [9] or to transform the timescale [10]. It is also possible to consider bivariate models [11].

In the present chapter we consider parameter estimation for different observation levels. In Section 9.2 only the degradation process is considered. We will find estimators for the parameters of a Wiener process with drift and with two parameters describing the starting time and the initial state of the process. Further, confidence regions and improved confidence regions are constructed. Especially the Bartlett correction to the

likelihood ratio statistic is calculated. In Section 9.3 we consider the classical sample of lifetimes. Finally, in Section 9.4 we consider the case that in each realization of the degradation process both process increments and a failure time are observable, and we estimate simultaneously the parameters μ, σ^2, t_0, z_0 , and h . By contrast with the previous models, here we observe for each sample path of the degradation process either a failure time or the process increments at fixed observation times under the condition that the process has not yet exceeded the degradation level h . Hence, these conditional process increments and the failure time T_h are dependent random variables. To compute the likelihood function, we have to find the conditional distribution of the process under the condition that the level h is not exceeded and the joint distribution of conditional process increments and the lifetime T_h .

9.2 Increments of the Degradation Process Are Observed

The observation of the underlying degradation process at fixed times t_j gives an alternative approach to parameter estimation which is particularly useful for products with high reliability since failure times have not been observed. If we ignore the boundary level h , the parameters μ, σ^2, t_0 , and z_0 may be estimated by the use of data of the degradation process before a failure occurs. For this case an example of a sample is given in Figure 9.1. The sample has the following form:

- n realizations of the degradation process $Z_i(t)$, $i = 1, \dots, n$, are observed.
- From each realization m_i observations $Z_i(t_{ij})$ at times t_{ij} , $j = 1, \dots, m_i$; $i = 1, \dots, n$, are given.

Let

$Y_{ij} = Z_i(t_{ij}) - Z_i(t_{ij-1})$ – increments of the process,

$s_{ij} = t_{ij} - t_{ij-1}$, $j = 2, \dots, m_i$ – distance between time points of observation,

$Z_{i1} = Z_i(t_{i1})$, $i = 1, \dots, n$,

$\bar{Z}_{.1} = \frac{1}{n} \sum_{i=1}^n Z_{i1}$ – average value of the first observed increments of the process,

$\bar{Z}_{.m} = \frac{1}{n} \sum_{i=1}^n Z_{im_i}$ – average value of the last observed process realizations,

$\bar{t}_{.1} = \frac{1}{n} \sum_{i=1}^n t_{i1}$ – average value of the first observation points,

$\bar{t}_{.m} = \frac{1}{n} \sum_{i=1}^n t_{im_i}$ – average value of the last observation points,

$\bar{m} = \frac{1}{n} \sum_{i=1}^n m_i$ – average value of the number of observations in one realization,

$\bar{A} = \frac{1}{n} \sum_{i=1}^n \frac{1}{t_{i1} - t_0}$, $\bar{B} = \frac{1}{n} \sum_{i=1}^n \frac{1}{(t_{i1} - t_0)^2}$.

Lower case letter notations like y_{ij} denote observed values of the corresponding random variables. In the case of the Wiener process the increments are Gaussian distributed:

$$\begin{aligned} Z_{i1} &\sim \mathcal{N}(z_0 + (t_{i1} - t_0)\mu, \sigma^2(t_{i1} - t_0)) \\ Y_{ij} &\sim \mathcal{N}(s_{ij}\mu, \sigma^2 s_{ij}) \end{aligned}$$

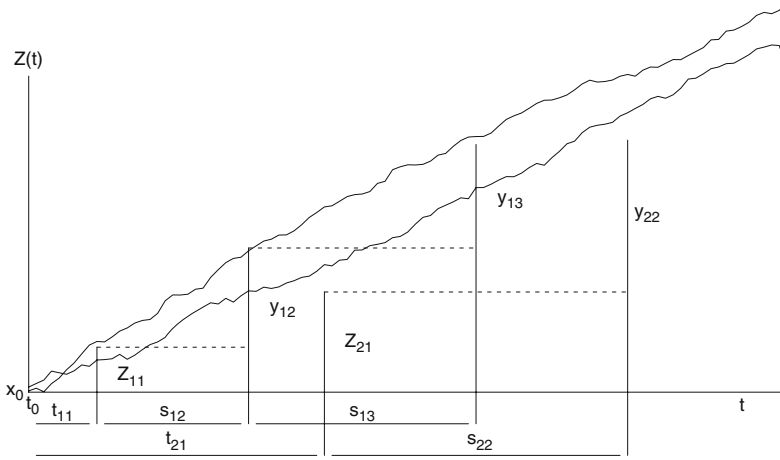


Figure 9.1. Sample for observations of process increments for the case $n = 2$, $m_1 = 3$, $m_2 = 2$

and the resulting likelihood function has the form

$$L = \prod_{i=1}^n \frac{1}{\sqrt{\sigma^2(t_{i1} - t_0)}} \varphi \left(\frac{z_{i1} - z_0 - \mu(t_{i1} - t_0)}{\sqrt{\sigma^2(t_{i1} - t_0)}} \right) \prod_{j=2}^{m_i} \frac{1}{\sqrt{\sigma^2 s_{ij}}} \varphi \left(\frac{(y_{ij} - \mu s_{ij})}{\sqrt{\sigma^2 s_{ij}}} \right), \quad (9.4)$$

where φ is the density function of the standard normal distribution. Here it has to be supposed that the first observation points in all realizations t_{i1} , $i = 1, \dots, n$, are greater than t_0 . In this case, which is formally obtained on setting $h = \infty$, we make neither use of information on failure times nor of the fact that the degradation process does not exceed the level h between two observation points. Note that the number of observations and the observation points may be different in each realization. The likelihood equations are

$$\begin{aligned} \hat{\mu} &= (\bar{z}_{.m} - \hat{z}_0) / (\bar{t}_{.m} - \hat{t}_0), \\ \hat{\sigma}^2 &= \frac{1}{mn} \left(\sum_{i=1}^n \left(\frac{(z_{i1} - \hat{z}_0)^2}{t_{i1} - \hat{t}_0} + \sum_{j=2}^{m_i} \frac{y_{ij}^2}{s_{ij}} \right) - \hat{\mu}n(\bar{z}_{.m} - \hat{z}_0) \right), \\ \hat{z}_0 &= \left(\frac{1}{n} \sum_{i=1}^n \frac{z_{i1}}{t_{i1} - \hat{t}_0} - \frac{\bar{z}_{.m}}{\bar{t}_{.m} - \hat{t}_0} \right) / \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{t_{i1} - \hat{t}_0} - \frac{1}{\bar{t}_{.m} - \hat{t}_0} \right), \\ \hat{\sigma}^2 \sum_{i=1}^n \frac{1}{t_{i1} - \hat{t}_0} &= \sum_{i=1}^n \frac{(z_{i1} - \hat{z}_0)^2}{(t_{i1} - \hat{t}_0)^2} - \hat{\mu}n. \end{aligned}$$

These equations must be solved numerically. We get an essential simplification if the first observation points are equal for all realizations of the degradation process: $t_{i1} = t_1$, $i = 1, \dots, n$. Then, an explicit solution of the likelihood equations can be found:

$$\begin{aligned}
\hat{\mu} &= (\bar{z}_{\cdot m} - \bar{z}_{\cdot 1}) / (\bar{t}_{\cdot m} - t_1) , \\
\hat{\sigma}^2 &= \frac{1}{\bar{m}(n-1)} \left(\sum_{i=1}^n \sum_{j=2}^{m_i} \frac{y_{ij}^2}{s_{ij}} - n \frac{(\bar{z}_{\cdot m} - \bar{z}_{\cdot 1})^2}{\bar{t}_{\cdot m} - t_1} \right) , \\
\hat{z}_0 &= (\bar{z}_{\cdot 1}(t_{\cdot m} - \hat{t}_0) - \bar{z}_{\cdot m}(t_1 - \hat{t}_0)) / (t_{\cdot m} - t_1) , \\
\hat{t}_0 &= t_1 - \frac{1}{\hat{\sigma}^2} \left(\sum_{i=1}^n z_{i1}^2 - \bar{z}_{\cdot 1}^2 \right) .
\end{aligned} \tag{9.5}$$

In the following, we will only consider the more simple case of equal first observation points. Then, the likelihood function has some nice properties. First, it can be proved [5] that the estimators (9.5) are asymptotically Gaussian distributed if

- either the number of observed processes n tends to ∞
- or we observe only one process, but \bar{m} and $(t_{\cdot m} - t_1)$ tend to ∞ .

The information matrix of the parameter $\theta = (\mu, \sigma^2, z_0, t_0)^T$ can be found to be

$$I = n \begin{bmatrix} \frac{t_{\cdot m} - t_0}{\sigma^2} & 0 & \frac{1}{\sigma^2} & -\frac{\mu}{\sigma^2} \\ 0 & \frac{\bar{m}}{2\sigma^4} & 0 & -\frac{1}{2\sigma^2} \bar{A} \\ \frac{1}{\sigma^2} & 0 & \frac{1}{\sigma^2} \bar{A} & -\frac{\mu}{\sigma^2} \bar{A} \\ -\frac{\mu}{\sigma^2} & -\frac{1}{2\sigma^2} \bar{A} & -\frac{\mu}{\sigma^2} \bar{A} & \frac{\mu}{\sigma^2} \bar{A} + \frac{\bar{B}}{2} \end{bmatrix} . \tag{9.6}$$

There are situations where only the parameters (μ, σ^2) are parameters of interest, while (z_0, t_0) which describe the begin of degradation are nuisance parameters. If the first observation points are equal for all realizations, then new nuisance parameters ν_1, ν_2 can be found, which are independent of the parameters of interest (μ, σ^2) . The method of reparameterization is used for conditional inference in [3]. The information matrix using new parameters $\nu_i = g_i(\theta)$, $(i = 1, \dots, 4)$ can be calculated by

$$I_g^{-1} = D_g I_\theta^{-1} D_g^T , \tag{9.7}$$

where $D_g = ((d_{ij}))_{i,j=1,\dots,4}$, $d_{ij} = \partial g_i / \partial \theta_j$. Our aim is now to find such new parameters that the estimators of these new parameters are asymptotically uncorrelated, that is, that the corresponding elements of the information matrix I_g tend to 0. For this purpose, a system of partial differential equations must be solved. In our case, it is possible to find such new parameters, which are $\nu_1, \nu_2, \mu, \sigma^2$ where

$$\begin{aligned}
\nu_1 &= z_0 + \mu(t_1 - t_0) , \\
\nu_2 &= \sigma^2(t_1 - t_0) .
\end{aligned}$$

Note that these new parameters describe the expectation and the variance of the degradation process at the first observation time t_1 . With the new parameters the likelihood function becomes

$$L = \left(\prod_{i=1}^n \frac{1}{\sqrt{\nu_2}} \varphi \left(\frac{z_{i1} - \nu_1}{\sqrt{\nu_2}} \right) \right) \left(\prod_{i=1}^n \prod_{j=2}^{m_i} \frac{1}{\sqrt{\sigma^2 s_{ij}}} \varphi \left(\frac{(y_{ij} - \mu s_{ij})}{\sqrt{\sigma^2 s_{ij}}} \right) \right) , \tag{9.8}$$

i.e., the likelihood function consists of two independent parts: one part includes the nuisance parameters ν_1, ν_2 and the other part includes the parameters of interest μ, σ^2 . Only the last part of the likelihood function must be taken into consideration.

The fact that the maximum likelihood estimator is asymptotically normally distributed can be used to construct simultaneous confidence regions for the parameters of the degradation process. An isoline of the k -dimensional Gaussian distribution, where k is the number of parameters, can be determined such that inside of this line there is $(1 - \alpha)$ of the probability mass of the distribution. The equation of this level line is given by the quadratic form [5]

$$(\hat{\theta} - \theta)^T I_{\theta}^{-1} (\hat{\theta} - \theta) < a^2, \quad (9.9)$$

with $\hat{\theta} = (\hat{\mu}, \hat{\sigma}^2, \hat{z}_0, \hat{t}_0)^T$, $\theta = (\mu, \sigma^2, z_0, t_0)^T$ where a^2 is the quantile of the χ^2 -distribution with $k = 4$ degrees of freedom and I_{θ} is the Fisher information matrix. Again, let us consider the two parameters of interest μ, σ^2 . In this case the confidence region is an isoline of the two-dimensional Gaussian distribution. For simplicity let again $t_{i1} = t_1$, $i = 1, \dots, n$. Then $\bar{A} = (t_1 - t_0)^{-1}$, $\bar{B} = (t_1 - t_0)^{-2}$, and the variances of the estimators are

$$\text{Var}(\hat{\mu}) = \frac{\sigma^2}{n(t_{.m} - t_1)}, \quad \text{Var}(\hat{\sigma}^2) = \frac{2\sigma^4}{n(\bar{m} - 1)}.$$

The joint confidence region has the form

$$\frac{(\hat{\mu} - \mu)^2}{\sigma^2} (t_{.m} - t_1) + \frac{(\hat{\sigma}^2 - \sigma^2)^2}{2\sigma^4} (\bar{m} - 1) \leq \frac{a^2}{n}, \quad (9.10)$$

with $a^2 = -2 \log \alpha$ - quantile of the χ^2 -distribution with 2 degrees of freedom.

It is well known that for many parameter estimators the rate of convergence to normality can be improved if a transformation of parameters can be found that the skewness of the estimation is minimal. Taking into consideration that $\hat{\mu}$ is already normally distributed, such a transformation must be found only for the parameter σ^2 . If $t_{i1} = t_1 = s_{ij} \ \forall i, j$, then $\hat{\sigma}^2$ is χ_f^2 -distributed with $f = 1$ degree of freedom. In this case the well-known Fisher transformation $\sigma_1^2 = (\sigma^2)^{1/3}$ improves the convergence to normality. If this transformation is used for the common case, the elements of the covariance matrix of $(\hat{\mu}, \hat{\sigma}_1^2)$ are given by

$$\text{Var}(\hat{\mu}) = \frac{(\sigma_1^{2/3})}{n(t_{.m} - t_1)}, \quad \text{Var}(\hat{\sigma}_1^2) = \frac{2}{9} \frac{\sigma_1^4}{n(\bar{m} - 1)}, \quad \text{Cov}(\hat{\mu}, \hat{\sigma}_1^2) = 0.$$

The improved joint confidence region is given by the inequality

$$\frac{(\hat{\mu} - \mu)^2}{\sigma_1^{2/3}} (t_{.m} - t_1) + \frac{9}{2} \frac{((\hat{\sigma}^2)^{1/3} - (\sigma^2)^{1/3})^2}{(\sigma^2)^{2/3}} (\bar{m} - 1) \leq \frac{a^2}{n}. \quad (9.11)$$

To improve the confidence regions for small samples the likelihood ratio can be used for the construction of joint confidence regions. It is known (see [1, 2]) that under some assumptions of regularity the log likelihood ratio

$$w = 2(\log L(\hat{\theta}) - \log L(\tilde{\theta})) \quad (9.12)$$

converges in distribution to a central χ^2 -distribution with k degrees of freedom, where k is the number of parameters of interest, $\hat{\theta}$ the maximum likelihood estimator of all parameters and $\tilde{\theta}$ the maximum likelihood estimator of the nuisance parameters, if the parameters of interest are fixed. The rate of convergence of the log likelihood ratio to the χ_k^2 -distribution is $O(n^{-1})$, where n is the sample size, and it may be improved to $O(n^{-3/2})$ by the Bartlett adjustment, if the fourth derivations of the log likelihood and there expectations exist. The log-likelihood ratio with Bartlett adjustment has the form

$$w' = w/(l + R/k),$$

where R is defined by

$$R = \sum (l_{rstu} - l_{rstuvw}), \quad (9.13)$$

with

$$\begin{aligned} l_{rstu} &= \lambda^{rs} \lambda^{tu} \left(\frac{1}{4} \lambda_{rstu} - \lambda_{rst}^{(u)} + \lambda_{rt}^{(su)} \right), \\ l_{rstuvw} &= \lambda^{rs} \lambda^{tu} \lambda^{vw} \left(\frac{1}{6} \lambda_{rtv} \lambda_{suw} + \frac{1}{4} \lambda_{rtu} \lambda_{svw} - \lambda_{rtv} \lambda_{sw}^{(u)} \right. \\ &\quad \left. - \lambda_{rtu} \lambda_{sw}^{(v)} + \lambda_{rt}^{(v)} \lambda_{sw}^{(u)} + \lambda_{rt}^{(u)} \lambda_{sw}^{(v)} \right), \\ \lambda_{rs} &= E \left(\frac{\partial^2 \log L}{\partial \theta_r \partial \theta_s} \right), \quad \lambda_{rst} = E \left(\frac{\partial^3 \log L}{\partial \theta_r \partial \theta_s \partial \theta_t} \right), \\ \lambda_{rstu} &= E \left(\frac{\partial^4 \log L}{\partial \theta_r \partial \theta_s \partial \theta_t \partial \theta_u} \right), \quad \lambda_{sw}^{(u)} = \frac{\partial \lambda_{sw}}{\partial \theta_u}, \text{ etc.} \end{aligned}$$

λ^{rs} is the (r, s) element of the inverse of the information matrix and the summation is taken over all integers

$$r, s, t, u, v, w = 1, \dots, k.$$

Cordeiro [2] derived correction factors for generalized linear models when the scale factor is assumed to be unknown. In the following, correction factors are derived for our degradation model with two parameters μ, σ^2 . The Bartlett factor in the form (9.13) includes 2^4 terms l_{rstu} with three terms each and 2^6 l_{rstuvw} with six terms each, that is, 432 terms have to be calculated. If there exists such a parameterization that $\lambda^{12} = 0$ (here the parameters are orthogonal), then there remain l_{1111} , l_{1122} , l_{2211} , l_{2222} , l_{111111} , l_{111122} , l_{112211} , l_{221111} , l_{112222} , l_{221122} , l_{222211} , l_{222222} and the Bartlett factor may be simplified to

$$\begin{aligned} R &= \frac{1}{\lambda_{11}^2} \left(\frac{1}{4} \lambda_{1111} - \lambda_{111}^{(1)} + \lambda_{11}^{(11)} \right) + \frac{1}{\lambda_{22}^2} \left(\frac{1}{4} \lambda_{2222} - \lambda_{222}^{(2)} + \lambda_{22}^{(22)} \right) \\ &\quad + \frac{1}{\lambda_{11} \lambda_{22}} \left(\frac{1}{2} \lambda_{1122} - \lambda_{112}^{(2)} - \lambda_{122}^{(1)} \right) \\ &\quad - \frac{1}{\lambda_{11}^3} \left(\frac{5}{12} \lambda_{1111}^2 - 2 \lambda_{111} \lambda_{11}^{(1)} + 2 (\lambda_{11}^{(1)})^2 \right) \\ &\quad - \frac{1}{\lambda_{22}^3} \left(\frac{5}{12} \lambda_{2222}^2 - 2 \lambda_{222} \lambda_{22}^{(2)} + 2 (\lambda_{22}^{(2)})^2 \right) \end{aligned}$$

$$\begin{aligned}
& -\frac{1}{\lambda_{11}^2 \lambda_{22}} \left(\frac{3}{4} \lambda_{112}^2 + \frac{1}{2} \lambda_{111} \lambda_{122} - \lambda_{112} \lambda_{11}^{(2)} - \lambda_{122} \lambda_{11}^{(1)} \right) \\
& -\frac{1}{\lambda_{11} \lambda_{22}^2} \left(\frac{3}{4} \lambda_{122}^2 + \frac{1}{2} \lambda_{222} \lambda_{112} - \lambda_{122} \lambda_{22}^{(2)} - \lambda_{112} \lambda_{22}^{(2)} \right).
\end{aligned} \tag{9.14}$$

In our model the maximum likelihood estimators of the parameters are not correlated and the Bartlett adjustment can be calculated by (9.14). It does not depend on the parameters and has the surprisingly simple form

$$R = 11 / (6n(\bar{m} - 1)). \tag{9.15}$$

In Section 9.5 confidence regions based on the asymptotic χ^2 -distribution of the likelihood ratio is given. The confidence region based on the log-likelihood ratio is the same as the improved confidence region based on the asymptotic normality. But with the likelihood ratio a general method for constructing confidence regions in the presence of nuisance parameters is found. By simulation studies it has been found that the confidence region (9.10), based on the asymptotic normality, can be used for $n(\bar{m} - 1) > 50$, the likelihood ratio (9.12) is usable for $n(\bar{m} - 1) > 25$ and the likelihood ratio with Bartlett adjustment is suitable for $n(\bar{m} - 1) > 18$.

9.3 Observation of Failure Times

When we know that a failure is the result of an underlying degradation process, the time to failure follows an inverse Gaussian distribution with the Lebesgue density function (9.3)

$$f_{T_h}(t) = \frac{h - z_0}{\sqrt{2\pi\sigma^2(t - t_0)^3}} \exp\left(-\frac{(h - z_0 - \mu(t - t_0))^2}{2\sigma^2(t - t_0)}\right) I_{\{t > t_0\}}.$$

Since the parameters of this time to failure correspond to the parameters of the underlying degradation process, it is possible to estimate the degradation parameters from a sample of failure times (τ_1, \dots, τ_n) .

In this case we get the classical likelihood function from a sample of inverse Gaussian distributed random variables:

$$L = \prod_{i=1}^n \frac{h - z_0}{\sqrt{2\pi\sigma^2(\tau_i - t_0)^3}} \exp\left(-\frac{(h - z_0 - \mu(\tau_i - t_0))^2}{2\sigma^2(\tau_i - t_0)}\right) I_{\{\tau_i > t_0\}}. \tag{9.16}$$

The estimators for the parameters μ, σ^2 can be calculated explicitly:

$$\hat{\mu} = \frac{h - z_0}{\frac{1}{n} \sum_{i=1}^n (\tau_i - t_0)}, \quad \hat{\sigma}^2 = (h - z_0)^2 \frac{1}{n} \sum_{i=1}^n \frac{1}{(\tau_i - t_0)} - \hat{\mu}(h - z_0).$$

Again, it can be shown that these estimators are asymptotically Gaussian distributed for $n \rightarrow \infty$ [5]. The problem of estimating t_0 is much more complicated, because it is a parameter which defines the support of the lifetime distribution.

Note, that here we can estimate the parameters μ , σ^2 , and t_0 only. The degradation reserve $h - z_0$ must be given in advance.

9.4 Observation of Both Degradation Increments and Failure Times

Let $Z(t) = z_0 + \sigma W(t - t_0) + \mu \cdot (t - t_0)$, $t \geq t_0$ be a degradation process on a probability space (Ω, \mathcal{F}, P) with values in $(\mathbb{R}, \mathcal{B})$ as defined in (9.1). Further, for $m \in \mathbb{N}$, let t_1, \dots, t_m be fixed observation points with $t_0 < t_1 < \dots < t_m < \infty$. To simplify the notation we consider only one realization of $Z(t)$ and drop the subscript i . We assume that a failure is observable at any time $t > t_0$ and that we stop observing the degradation process after a failure has occurred. Hence, in each time interval $(t_{j-1}, t_j]$, $j = 1, \dots, m$, we observe either a failure at time $\tau \in (t_{j-1}, t_j]$, i.e., a realization τ of the lifetime T_h in $(t_{j-1}, t_j]$, or we observe the degradation measure $z_j = Z(t_j)$ at t_j under the condition that the process has not yet exceeded the level h until the time t_j . The observable stopped process is given by

$$\tilde{Z}(t) = \begin{cases} Z(t) & , \quad t_0 \leq t \leq T_h \\ \infty & , \quad t > T_h \end{cases} \quad , \quad t \geq t_0.$$

Thus, $\tilde{Z}(t)$ is a functional of $\{Z(s); t_0 \leq s \leq t\}$,

$$\tilde{Z}(t) = F_t(T_h, Z(t)), \quad t \geq t_0, \quad (9.17)$$

where

$$F_t(\tau, z) = \begin{cases} z & , \quad \tau \geq t \\ \infty & , \quad \tau < t \end{cases} \quad , \quad t, \tau, z \in \mathbb{R},$$

and measurable with respect to the σ -algebra $\mathcal{F}_t = \sigma\{Z(s); t_0 \leq s \leq t\}$.

A sample of censored observations has the structure

$$X = F(T_h, Z(t_1), \dots, Z(t_m)) = (\min(T_h, t_m), \tilde{Z}(t_1), \dots, \tilde{Z}(t_m)) \quad (9.18)$$

with

$$F(\tau, z_1, \dots, z_m) = (F_0(\tau), F_{t_1}(\tau, z_1), \dots, F_{t_m}(\tau, z_m))$$

and

$$F_0(\tau) = \min(\tau, t_m),$$

and the statistical model is given by

$$(\mathbb{R} \times \overline{\mathbb{R}}^m, \mathcal{B} \otimes \overline{\mathcal{B}}^m, (P_\theta^F)_{\theta=(\mu, \sigma^2, z_0, t_0, h) \in \Theta \subset \mathbb{R}^5}),$$

where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ and $\overline{\mathcal{B}}$ is the σ -algebra of Borel sets in $\overline{\mathbb{R}}$.

To compute the likelihood function of X we have to find the Radon-Nikodym derivative of P_θ^F with respect to a dominating measure ν :

$$L(\tau, \mathbf{z}; \theta) = \frac{dP_\theta^F}{d\nu}(\tau, \mathbf{z}), \quad \tau \in \mathbb{R}, \quad \mathbf{z} = (z_1, \dots, z_m) \in \overline{\mathbb{R}}^m. \quad (9.19)$$

A crucial step in the derivation of $L(\tau, \mathbf{z}; \theta)$ is the calculation of the conditional density $f(s, x, t, z, h)$ of the process $Z(t)$ at time t starting at point $x = Z(s)$ under the condition that the level h is not exceeded in $[s, t]$. For that purpose we consider the Markov family $(Z(t), P_{s,x})$ with the transition function

$$P(s, x, t, B) = (2\pi\sigma^2(t-s))^{-1/2} \int_B \exp\left(-\frac{(y-x-\mu(t-s))^2}{2\sigma^2(t-s)}\right) dy$$

for $t > s \geq t_0$ and $B \in \mathcal{B}$, i.e., for all $s \geq t_0$ and $x \in \mathbb{R}$, $P_{s,x}$ is a probability measure on $\mathcal{F}_{\geq s} = \sigma\{Z(u); u \geq s\}$ with $P_{s,x}(Z(t) \in B) = P(s, x, t, B)$. Let T_h^s denote the first passage time of $Z(t)$ to the upper boundary h on $[s, \infty)$

$$T_h^s = \inf\{t \geq s : Z(t) \geq h\}.$$

In case $s = t_0$, we have $T_h^{t_0} = T_h$ according to definition (9.2). For every fixed $s \geq t_0$ and $x \in \mathbb{R}$, T_h^s is a random variable on the probability space $(\Omega, \mathcal{F}_{\geq s}, P_{s,x})$. Hence, for $x < h$ the induced probability measure $P_{s,x}^{T_h^s}$, i.e., the conditional distribution of T_h^s under the condition $\{Z(s) = x\}$, has the Lebesgue density

$$g(s, x, t, h) = \frac{h-x}{\sqrt{2\pi\sigma^2(t-s)^3}} \exp\left(-\frac{(h-x-\mu(t-s))^2}{2\sigma^2(t-s)}\right) I_{\{t>s\}}, \quad (9.20)$$

whereas for $x \geq h$, $P_{s,x}^{T_h^s}$ is the Dirac measure ε_s concentrated at s . Using total probability and the Markov property for Markov families (cf. Wentzell [8], p. 109) we get for $x < h$ and $s < t$

$$\begin{aligned} P_{s,x}(Z(t) \leq z, T_h^s \geq t) \\ &= P_{s,x}(Z(t) \leq z) - \int_s^t P_{s,x}(Z(t) \leq z | T_h^s = u) P_{s,x}^{T_h^s}(du) \\ &= P(s, x, t, (-\infty, z)) - \int_s^t P(u, h, t, (-\infty, z)) g(s, x, u, h) du \end{aligned}$$

and hence, applying Fubini's theorem and passing to densities

$$f(s, x, t, z, h) = p(s, x, t, z) - \int_s^t p(u, h, t, z) g(s, x, u, h) du, \quad (9.21)$$

where $p(s, x, t, z)$ is the transition density of $(Z(t), P_{s,x})$. Using (9.21), the following lemma provides an explicit representation for the density $f(s, x, t, z, h)$.

Lemma 1. *Let $t_0 \leq s < t$, $x < h$, and $z \in \mathbb{R}$. The density (9.21) is given by*

$$\begin{aligned} f(s, x, t, z, h) &= \frac{1}{\sigma\sqrt{t-s}} \phi\left(\frac{z-x-\mu(t-s)}{\sigma\sqrt{t-s}}\right) \\ &\quad \times \left[1 - \exp\left(-\frac{2(h-x)(h-z)}{\sigma^2(t-s)}\right)\right] I_{\{z \leq h\}}, \end{aligned} \quad (9.22)$$

where ϕ is the density of the standard normal distribution.

PROOF. Recalling (9.21) we have to compute the integral

$$\begin{aligned} I &= \int_s^t p(u, h, t, z) g(s, x, u, h) du \\ &= \int_s^t \frac{h-x}{2\pi\sigma^2\sqrt{(t-u)(u-s)^3}} \\ &\quad \times \exp\left[-\frac{1}{2\sigma^2}\left(\frac{(z-h-\mu(t-u))^2}{t-u} + \frac{(h-x-\mu(u-s))^2}{u-s}\right)\right] du \end{aligned}$$

for $t > s \geq t_0$ and $x < h$. Applying the substitution $y = ((t-u)/(u-s))^{1/2}$ for $s < u < t$, i.e., $u = (t+sy^2)/(1+y^2)$ with $+\infty > y > 0$, we obtain after some algebra

$$\begin{aligned} I &= \frac{h-x}{\pi\sigma^2(t-s)} \exp\left[-\frac{(z-h)^2+(h-x)^2-2\mu(t-s)(z-x)+\mu^2(t-s)^2}{2\sigma^2(t-s)}\right] \\ &\quad \times \int_0^\infty \exp\left[-\left(\frac{\beta^2}{y^2} + \alpha^2 y^2\right)\right] dy \end{aligned}$$

with $\beta^2 = (z-h)^2/(2\sigma^2(t-s))$ and $\alpha^2 = (h-x)^2/(2\sigma^2(t-s))$. Denoting the last integral with I_1 it can be easily checked that for $\alpha \neq 0$

$$\begin{aligned} I_1 &= \frac{\sqrt{\pi}}{4\alpha} \left[e^{2\alpha\beta} \left(\operatorname{erf}\left(\alpha z + \frac{\beta}{z}\right) - 1 \right) + e^{-2\alpha\beta} \left(\operatorname{erf}\left(\alpha z - \frac{\beta}{z}\right) + 1 \right) \right] \Bigg|_{z=0}^{z=\infty} \\ &= \frac{\sqrt{\pi}}{2|\alpha|} e^{-2|\alpha\beta|}, \end{aligned}$$

where $\operatorname{erf}(x) = 2\pi^{-1/2} \int_0^x e^{-t^2} dt$.

Hence, as $|\alpha| = |h-x|(2\sigma^2(t-s))^{-1/2}$ and $x < h$, we get

$$I = \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} \exp\left[-\frac{Q - 2\mu(t-s)(z-x) + \mu^2(t-s)^2}{2\sigma^2(t-s)}\right]$$

with

$$\begin{aligned} Q &= (z-h)^2 + 2|z-h|(h-x) + (h-x)^2 \\ &= \begin{cases} (z-2h+x)^2 = (z-x)^2 + 4(h-x)(h-z), & z \leq h \\ (z-x)^2, & z > h \end{cases} \end{aligned}$$

and finally

$$I = \frac{1}{\sigma\sqrt{t-s}} \phi\left(\frac{z-x-\mu(t-s)}{\sigma\sqrt{t-s}}\right) \times \begin{cases} \exp\left(-\frac{2(h-x)(h-z)}{\sigma^2(t-s)}\right), & z \leq h \\ 1, & z > h \end{cases},$$

where ϕ is the density of the standard normal distribution. This completes the proof. ■

The next lemma gives the density of the finite dimensional distributions of $Z(t)$ starting at $x = Z(s)$ under the condition that h is not exceeded in terms of f . Let λ^k be the k -dimensional Lebesgue measure on $\bar{\mathcal{B}}^k$ ($k \in \mathbb{N}$). We use the convention $\prod_{j=j_1}^{j_2} \alpha_j = 1$ if $j_2 < j_1$.

Lemma 2. Let $t_0 \leq s < t_1$, $\mathbf{t}_{1,k} = (t_1, \dots, t_k)$, $\mathbf{z}_{1,k} = (z_1, \dots, z_k) \in \mathbb{R}^k$, and $\mathbf{Z}_{\mathbf{t}_{1,k}} = (Z(t_1), \dots, Z(t_k))$ for $1 \leq k \leq m$. Then the probability measure $P_{s,x}(\mathbf{Z}_{\mathbf{t}_{1,k}} \in B, T_h^s \geq t_k)$, $B \in \mathcal{B}^k$, has the λ^k -density

$$f_k(s, x, \mathbf{t}_{1,k}, \mathbf{z}_{1,k}, h) = f(s, x, t_1, z_1, h) \prod_{j=2}^k f(t_{j-1}, z_{j-1}, t_j, z_j, h). \quad (9.23)$$

PROOF. For $k = 1$, clearly $f_1(s, x, \mathbf{t}_{1,1}, \mathbf{z}_{1,1}, h) = f(s, x, t_1, z_1, h)$ by the definition of f . For simplicity we will prove (9.23) only for $k = 2$. The general case can be shown analogously by induction. Let $B_1, B_2 \in \mathcal{B}$. Then, applying total probability and making use of the Markov property, we have

$$\begin{aligned} P_{s,x}(\mathbf{Z}_{\mathbf{t}_{1,2}} \in B_1 \times B_2, T_h^s \geq t_2) \\ &= P_{s,x}(Z(t_2) \in B_2, T_h^{t_1} \geq t_2, Z(t_1) \in B_1, T_h^s \geq t_1) \\ &= \int_{B_1} P_{s,x}(Z(t_2) \in B_2, T_h^{t_1} \geq t_2 \mid Z(t_1) = z_1) f(s, x, t_1, z_1, h) dz_1 \\ &= \int_{B_1} P_{t_1, z_1}(Z(t_2) \in B_2, T_h^{t_1} \geq t_2) f(s, x, t_1, z_1, h) dz_1 \\ &= \int_{B_2 \times B_1} f(t_1, z_1, t_2, z_2, h) f(s, x, t_1, z_1, h) dz_2 dz_1. \end{aligned}$$

■

Based on these preliminary lemmas we can derive the likelihood function $L(\tau, \mathbf{z}; \cdot)$. We shall show that P_θ^F is for all $\theta \in \Theta$ absolutely continuous with respect to the measure

$$\nu = (\lambda + \varepsilon_{t_m}) \otimes (\lambda + \varepsilon_\infty)^m, \quad (9.24)$$

where λ is the Lebesgue measure on $\overline{\mathcal{B}}$ and, for every $x \in \overline{\mathbb{R}}$, ε_x is the Dirac measure on $\overline{\mathcal{B}}$ concentrated at x .

Theorem 1. Let ν be defined as in (9.24) on $\mathcal{B} \otimes \overline{\mathcal{B}}^m$. Then, for all $\theta \in \Theta$, P_θ^F is absolutely continuous with respect to ν and has the Radon–Nikodym derivative

$$\begin{aligned} L(\tau, \mathbf{z}; \theta) &= \frac{dP_\theta^F}{d\nu}(\tau, \mathbf{z}) \\ &= \prod_{k=1}^m \left(g(t_{k-1}, z_{k-1}, \tau, h) \prod_{j=1}^{k-1} f(t_{j-1}, z_{j-1}, t_j, z_j, h) \prod_{j=k}^m I_{\{z_j = \infty\}} \right)^{I_{\{t_{k-1} \leq \tau < t_k\}}} \\ &\quad \times \left(\prod_{j=1}^m f(t_{j-1}, z_{j-1}, t_j, z_j, h) \right)^{I_{\{\tau = t_m\}}} I_{\{t_0 \leq \tau \leq t_m\}}, \end{aligned} \quad (9.25)$$

where $\tau \in \mathbb{R}$ and $\mathbf{z} = (z_1, \dots, z_m) \in \overline{\mathbb{R}}^m$.

PROOF. Let $1 \leq i \leq j \leq n$, $(y_i, \dots, y_j) \in \overline{\mathbb{R}}^{j-i+1}$, and $t \in \mathbb{R}$. For the sake of brevity we define $\mathbf{t}_{i,j} = (t_i, \dots, t_j)$, $\mathbf{y}_{i,j} = (y_i, \dots, y_j)$, $(-\infty, \mathbf{y}_{i,j}] = (-\infty, y_i] \times \dots \times (-\infty, y_j]$,

and $E = (-\infty, t] \times (-\infty, \mathbf{y}_{1,m}]$. Further, we use the abbreviation $\{\mathbf{Z}_{\mathbf{t}_{i,j}} \leq \mathbf{y}_{i,j}\} = \{Z(t_i) \leq y_i, \dots, Z(t_j) \leq y_j\}$ and an analogous one for $\tilde{\mathbf{Z}}(t)$. Recalling (9.18) we have for all $\theta \in \Theta$

$$\begin{aligned} P_\theta^F(E) &= P_{t_0, z_0}(\min(T_h, t_m) \leq t, \tilde{\mathbf{Z}}_{\mathbf{t}_{1,m}} \leq \mathbf{y}_{1,m}) \\ &= \sum_{k=1}^m I_{\{t \geq t_{k-1}\}} P_{t_0, z_0}(t_{k-1} \leq T_h \leq \min(t, t_k), \tilde{\mathbf{Z}}_{\mathbf{t}_{1,m}} \leq \mathbf{y}_{1,m}) \\ &\quad + I_{\{t \geq t_m\}} P_{t_0, z_0}(T_h \geq t_m, \tilde{\mathbf{Z}}_{\mathbf{t}_{1,m}} \leq \mathbf{y}_{1,m}). \end{aligned} \quad (9.26)$$

To calculate the first m terms of (9.26) take $t_{k-1} \leq t < t_k$ with $1 \leq k \leq m$ and define $\bigcap_{k=k_1}^{k_2} A_k = \Omega$ if $k_2 < k_1$. Then, for $s \geq t_{k-1}$ by the definition of $\tilde{\mathbf{Z}}(s)$

$$\begin{aligned} &P_{t_0, z_0}(t_{k-1} \leq T_h \leq s, \tilde{\mathbf{Z}}_{\mathbf{t}_{1,m}} \leq \mathbf{y}_{1,m}) \\ &= P_{t_0, z_0}(\{t_{k-1} \leq T_h \leq s, \mathbf{Z}_{\mathbf{t}_{1,k-1}} \leq \mathbf{y}_{1,k-1}\} \cap \{\tilde{\mathbf{Z}}_{\mathbf{t}_{k,m}} \leq \mathbf{y}_{k,m}\}) \\ &= P_{t_0, z_0}(T_h^{t_{k-1}} \leq s, T_h^{t_0} \geq t_{k-1}, \mathbf{Z}_{\mathbf{t}_{1,k-1}} \leq \mathbf{y}_{1,k-1}) \prod_{j=k}^m I_{\{y_j = \infty\}}. \end{aligned} \quad (9.27)$$

Applying total probability, the Markov property, and Lemma 2 yields

$$\begin{aligned} &P_{t_0, z_0}(T_h^{t_{k-1}} \leq s, T_h^{t_0} \geq t_{k-1}, \mathbf{Z}_{\mathbf{t}_{1,k-1}} \leq \mathbf{y}_{1,k-1}) \\ &= \int_{(-\infty, \mathbf{y}_{1,k-1}]} P_{t_0, z_0}(T_h^{t_{k-1}} \leq s \mid Z(t_{k-1}) = z_{k-1}) \\ &\quad \times f_{k-1}(t_0, z_0, \mathbf{t}_{1,k-1}, \mathbf{z}_{1,k-1}, h) \lambda^{k-1}(d\mathbf{z}_{1,k-1}) \end{aligned}$$

and, moreover, recalling the distribution of $P_{s,x}^{T_h^s}$

$$\begin{aligned} P_{t_0, z_0}(T_h^{t_{k-1}} \leq s \mid Z(s_{k-1}) = z_{k-1}) &= P_{t_{k-1}, z_{k-1}}(T_h^{t_{k-1}} \leq s) \\ &= \int_{t_{k-1}}^s g(t_{k-1}, z_{k-1}, \tau, h) \lambda(d\tau), \end{aligned}$$

if $z_{k-1} < h$. Finally, the product of indicator functions in (9.27) can be written as

$$\prod_{j=k}^m I_{\{y_j = \infty\}} = \int_{(-\infty, \mathbf{y}_{k,m}]} \prod_{j=k}^m I_{\{z_j = \infty\}} \varepsilon_\infty^{m-k+1}(d\mathbf{z}_{k,m}).$$

Combining the above equations we obtain

$$\begin{aligned} &P_{t_0, z_0}(t_{k-1} \leq T_h \leq s, \tilde{\mathbf{Z}}_{\mathbf{t}_{1,m}} \leq \mathbf{y}_{1,m}) \\ &= \int_{(-\infty, \mathbf{y}_{1,m}]} \int_{t_{k-1}}^s g(t_{k-1}, z_{k-1}, \tau, h) \lambda(d\tau) f_{k-1}(t_0, z_0, \mathbf{t}_{1,k-1}, \mathbf{z}_{1,k-1}, h) \\ &\quad \times \prod_{j=k}^m I_{\{z_j = \infty\}} (\lambda^{k-1} \otimes \varepsilon_\infty^{m-k+1})(d\mathbf{z}_{1,m}). \end{aligned} \quad (9.28)$$

Now, from (9.22) and (9.23) it is obvious that the measure $\lambda^{k-1} \otimes \varepsilon_\infty^{m-k+1}$ in (9.28) can be replaced by

$$(\lambda + \varepsilon_\infty)^m = \sum_{(\mu_1, \dots, \mu_m) \in \{\lambda, \varepsilon_\infty\}^m} \bigotimes_{i=1}^m \mu_i$$

without altering the value of the integral, since integrating with respect to the summands of $(\lambda + \varepsilon_\infty)^m$ gives zero except for the case $\mu_i = \lambda^{k-1} \otimes \varepsilon_\infty^{m-k+1}$. Hence, going back to (9.26) and using (9.27) and Lemma 2 we obtain

$$\begin{aligned} P_\theta^F(E) &= \int_{(-\infty, \mathbf{y}_{1,m}]} \left(\int_{t_0}^t \sum_{k=1}^m \left[I_{\{t_{k-1} \leq \tau < t_k\}} g(t_{k-1}, z_{k-1}, \tau, h) \lambda(d\tau) \right. \right. \\ &\quad \times f_{k-1}(t_0, z_0, \mathbf{t}_{1,k-1}, \mathbf{z}_{1,k-1}, h) \prod_{j=k}^m I_{\{z_j = \infty\}} \left. \right] \\ &\quad + \int_{t_0}^t I_{\{\tau = t_m\}} \varepsilon_{t_m}(d\tau) f_m(t_0, z_0, \mathbf{t}_{1,m}, \mathbf{z}_{1,m}, h) \left. \right) (\lambda + \varepsilon_\infty)^m(d\mathbf{z}_{1,m}) \\ &= \int_{(-\infty, t] \times (-\infty, \mathbf{y}_{1,m}]} \left(\sum_{k=1}^m \left[I_{\{t_{k-1} \leq \tau < t_k\}} g(t_{k-1}, z_{k-1}, \tau, h) \right. \right. \\ &\quad \times \prod_{j=1}^{k-1} f(t_{j-1}, z_{j-1}, t_j, z_j, h) \prod_{j=k}^m I_{\{z_j = \infty\}} \left. \right] \\ &\quad + I_{\{\tau = t_m\}} \prod_{j=1}^m f(t_{j-1}, z_{j-1}, t_j, z_j, h) \left. \right) \nu(d(\tau, \mathbf{z}_{1,m})), \end{aligned}$$

where $\nu = (\lambda + \varepsilon_{t_m}) \otimes (\lambda + \varepsilon_\infty)^m$.

Thus, the integrand of the last equation is a version of the Radon–Nikodym derivative of P_θ^F with respect to ν , since $\mathcal{B} \otimes \overline{\mathcal{B}}^m$ is generated by the sets E . To complete the proof note that this integrand and the right-hand side of (9.25) are equivalent. ■

In general we observe not only one realization of the degradation process but the realizations of n independent degradation processes $Z_i(t)$, $1, \dots, n$, corresponding to n independent items. Let t_{i1}, \dots, t_{im_i} with $t_0 < t_{i1} < \dots < t_{im_i} < \infty$ be the m_i fixed observation points of the realization of $Z_i(t)$, $M = \sum_{i=1}^n m_i$, and $T_h^{(i)}$ the lifetime of $Z_i(t)$. The censored observations in this realization have the form

$$\begin{aligned} &(\tau_i, z_{i1}, \dots, z_{im_i}) \\ &= (\min(T_h^{(i)}, t_{im_i}), F_{t_{i1}}(T_h^{(i)}, Z_i(t_{i1})), \dots, F_{t_{im_i}}(T_h^{(i)}, Z_i(t_{im_i}))). \end{aligned}$$

Then the statistical model is given by

$$(\mathbb{R}^n \times \overline{\mathbb{R}}^M, \mathcal{B}^n \otimes \overline{\mathcal{B}}^M, (\otimes_{i=1}^n P_\theta^F)_{\theta=(\mu, \sigma^2, z_0, t_0, h) \in \Theta \subset \mathbb{R}^5})$$

with the likelihood function

$$L(\tau, \mathbf{z}^{(1)}, \dots, \mathbf{z}^{(n)}; \theta) = \prod_{i=1}^n L(\tau_i, \mathbf{z}^{(i)}; \theta),$$

where $\tau = (\tau_1, \dots, \tau_n) \in \mathbb{R}^n$ and $\mathbf{z}^{(i)} = (z_{i1}, \dots, z_{im_i}) \in \overline{\mathbb{R}}^{m_i}$, $i = 1, \dots, n$.

9.5 An Example

Substituting formulae (9.20) and (9.22) for the functions g and f in (9.25) and setting $t_{i0} = t_0$ and $z_{i0} = z_0$ we get the likelihood function of the i th realization ($i = 1, \dots, n$)

$$L(\tau_i, \mathbf{z}^{(i)}; \theta) = \prod_{j=1}^{k-1} \left(\frac{1}{\sqrt{\sigma^2(t_{ij} - t_{ij-1})}} \phi \left(\frac{(z_{ij} - z_{ij-1}) - \mu(t_{ij} - t_{ij-1})}{\sqrt{\sigma^2(t_{ij} - t_{ij-1})}} \right) \right. \\ \cdot \left[1 - \exp \left(-\frac{2(h - z_{ij-1})(h - z_{ij})}{\sigma^2(t_{ij} - t_{ij-1})} \right) \right] \prod_{j=k}^{m_i} I_{\{z_j = \infty\}} \\ \times \frac{h - z_{ik-1}}{\sqrt{\sigma^2(\tau_i - t_{ik-1})}^3} \phi \left(\frac{h - z_{ik-1} - \mu(\tau_i - t_{ik-1})}{\sqrt{\sigma^2(\tau_i - t_{ik-1})}} \right), \quad (9.29)$$

if $t_{k-1} \leq \tau_i < t_k$ and $1 \leq k \leq m_i$ and

$$L(\tau_i, \mathbf{z}^{(i)}; \theta) = \prod_{j=1}^{m_i} \left(\frac{1}{\sqrt{\sigma^2(t_{ij} - t_{ij-1})}} \phi \left(\frac{(z_{ij} - z_{ij-1}) - \mu(t_{ij} - t_{ij-1})}{\sqrt{\sigma^2(t_{ij} - t_{ij-1})}} \right) \right. \\ \cdot \left[1 - \exp \left(-\frac{2(h - z_{ij-1})(h - z_{ij})}{\sigma^2(t_{ij} - t_{ij-1})} \right) \right] \Bigg), \quad (9.30)$$

if $\tau_i = t_{m_i}$. To compare our estimates based on the likelihood function in (9.29) and (9.30) with previous results, we consider the case that only degradation process increments are observed. As shown in Section 9.2, in this case the likelihood function of the i th realization is given by

$$L_D(\mathbf{z}^{(i)}; \theta) = \prod_{j=1}^{m_i} \frac{1}{\sqrt{\sigma^2(t_{ij} - t_{ij-1})}} \phi \left(\frac{(z_{ij} - z_{ij-1}) - \mu(t_{ij} - t_{ij-1})}{\sqrt{\sigma^2(t_{ij} - t_{ij-1})}} \right). \quad (9.31)$$

We notice that the likelihood function (9.31) can be obtained from (9.30) if h tends to infinity; this means that if a boundary level h does not exist then, consequently, exceeding the boundary is impossible.

As an example, we have simulated a sample of $n = 5$ process realizations with $m_i = 10$ observation points t_{ij} in each realization, which have equal distances $t_{ij} - t_{ij-1} = 1$ and with true parameter values $\mu = 1.0$, $\sigma^2 = 0.25$, $t_0 = 0$, $z_0 = 0$, and $h = 5.0$. The simulated data are shown in the following table:

i	1	2	3	4	5
$z_{ij} - z_{ij-1}$	1.290	0.397	0.744	0.984	1.223
	1.096	1.619	1.125	0.904	0.769
	0.715	0.427	1.047	1.384	1.775
	0.339	1.574	0.530	1.000	0.567
	0.028		0.123		0.426
	1.231		1.346		
τ_i	6.559	4.967	6.018	4.461	5.138

Suppose that the parameter values $t_0 = 0$ and $z_0 = 0$ are known, we first estimate the parameters μ and σ^2 maximizing the likelihood function (9.31), this means, we ignore that the degradation process does not exceed the boundary level between observation points. Using maximum likelihood estimators described in Section 9.2 we get

$$\hat{\mu}_D = 0.9198, \quad \hat{\sigma}_D^2 = 0.1983.$$

In Figure 9.2, 0.95- and 0.99-confidence regions for μ and σ^2 based on the asymptotic χ_2^2 -distribution of the likelihood ratio are shown for this model.

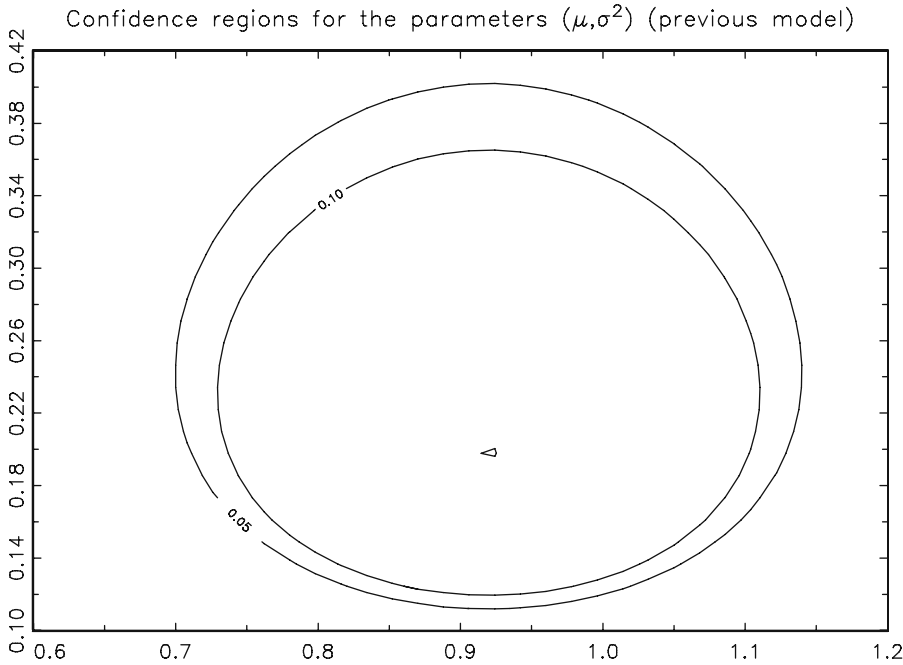


Figure 9.2. Confidence estimates for (μ, σ^2)

Now, again supposing that $t_0 = z_0 = 0$ are known, we compute maximum likelihood estimates for the parameters μ , σ^2 , and h based on the likelihood function (9.29) of the exact model. The maximization of the likelihood function and also the calculation of confidence regions were done by the computer program GAUSS and yielded the point estimates

$$\hat{\mu} = 0.9222, \quad \hat{\sigma}^2 = 0.1915, \quad \hat{h} = 5.0047.$$

The estimates turned out to be difficult to find so that a good initial value of the parameters has to be known. Such initial values may be the estimates of μ and σ^2 from the likelihood function (9.31) and an initial value of h can be found from a likelihood function based on the inverse Gaussian density (9.3). Figures 9.3, 9.4, and 9.5 show 0.95- and 0.99-confidence regions based on the asymptotic χ_2^2 -distribution of the likelihood ratio for the parameters (μ, σ^2) , (μ, h) , and (σ^2, h) , respectively. In all examples considered both the estimates of the parameter σ^2 and the confidence regions for the exact model are smaller than the estimates based on (9.31). Note that in this case it is also possible to estimate the parameter h .

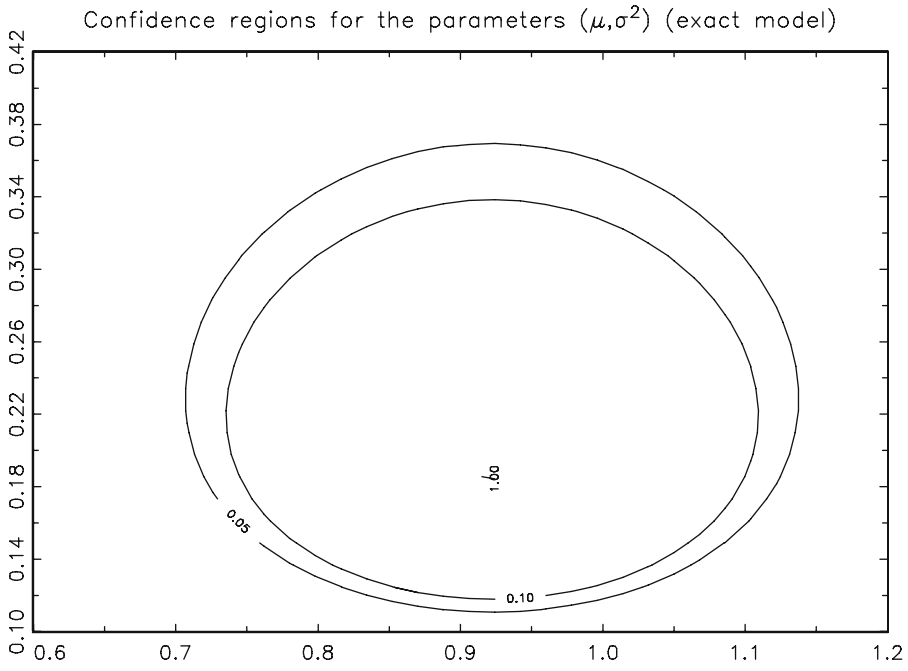


Figure 9.3. Confidence estimates for (μ, σ^2)

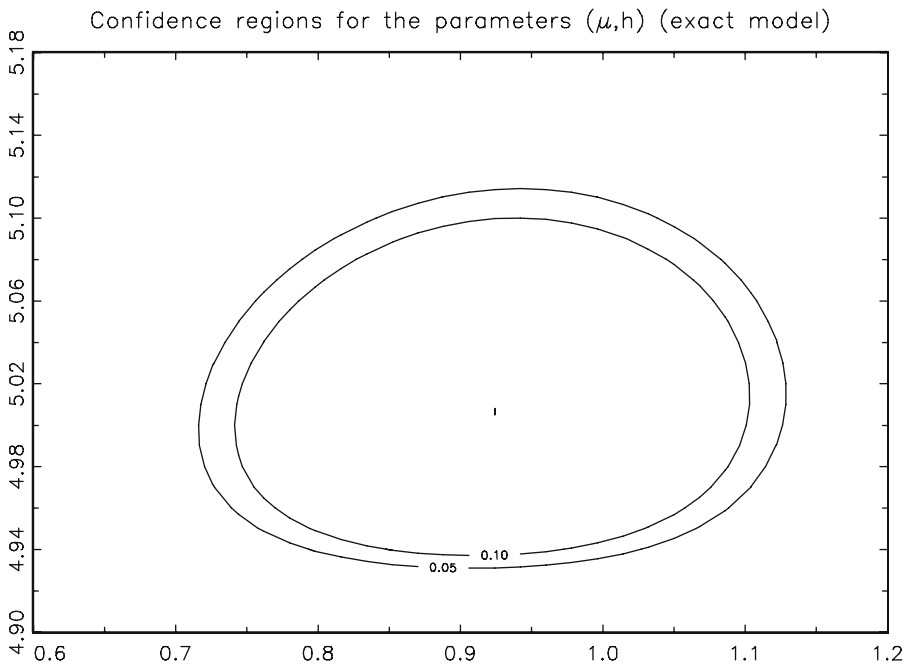


Figure 9.4. Confidence estimates for (μ, h)

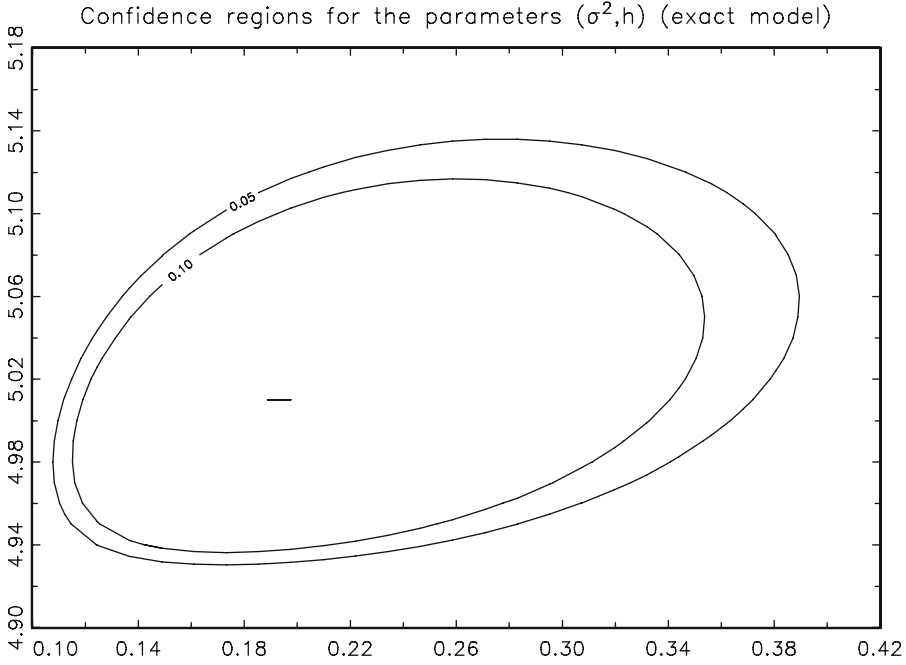


Figure 9.5. Confidence estimates for (σ^2, h)

9.6 Simulation Study

To find a maximum likelihood estimator $\hat{\theta}$ of θ , the log-likelihood equations $\partial \log L(X; \theta) / \partial \theta_i = 0$ have to be solved numerically. In a simulation study we have estimated the partial parameter vector $\theta^* = (\mu, \sigma^2, h)$, whereas t_0 and z_0 were assumed to be known. Given an observation vector $X = (X_1, \dots, X_n)$ with $X_i = (\min(\tau_i, t_{im_i}), z_{i1}, \dots, z_{iM_i})$, only the maximum likelihood estimator for the drift parameter μ from the likelihood functions (9.29) and (9.30) can be explicitly obtained by

$$\hat{\mu} = \frac{\sum_{i=1}^n \left(z_{iM_i} - z_{i0} + I_{\{\tau_i < t_{im_i}\}} (\hat{h} - z_{iM_i}) \right)}{\sum_{i=1}^n \left(t_{iM_i} - t_{i0} + I_{\{\tau_i < t_{im_i}\}} (\tau_i - t_{iM_i}) \right)},$$

where $M_i = \max\{k : t_{ik} < T_h^{(i)}\}$. The estimators $\hat{\sigma}^2$ and \hat{h} were determined numerically using the optimization module of the software package GAUSS.

The following table summarizes the estimated mean-squared errors of the maximum likelihood estimators $\hat{\mu}$, $\hat{\sigma}^2$, and \hat{h} from 1000 simulation runs and compares them with those of the estimators $\hat{\mu}_D$ and $\hat{\sigma}_D^2$ from the likelihood function (9.31) which served as initial values for the optimization procedure. In each simulation run, $n = 50$ process realizations were simulated with $m_i = 10$ observation points t_{ij} and equal distances $t_{ij} - t_{ij-1} = 1$. The true parameter values were $\mu = 5$; $\sigma^2 = 1, 2, 5$; $t_0 = 0$; $z_0 = 0$; and

$h = 20, 40, 50$. Additionally, the table gives the average number \bar{d} of observation points t_{ij} per realization before a failure.

Table 9.1 shows that for all parameter choices in the simulation study the mean-squared errors of the estimators $\hat{\mu}$, $\hat{\sigma}^2$, and \hat{h} based on the likelihood functions (9.29) and (9.30) are smaller than those of the estimators $\hat{\mu}_D$ and $\hat{\sigma}^2_D$ based on the likelihood function (9.31). In particular, the general model based on the observation of degradation process increments and failure times provides better estimates of the process parameters μ and σ^2 if in nearly all realizations failures have occurred and the number of observation points t_{ij} , where actually degradation levels could be observed, is small in comparison with the number of planned observation points.

Table 9.1. Mean-squared errors of maximum likelihood estimators in 1000 simulation runs

μ	σ^2	h	\bar{d}	Mean-squared errors				
				$\hat{\mu}$	$\hat{\mu}_D$	$\hat{\sigma}^2$	$\hat{\sigma}^2_D$	\hat{h}
5	2	20	3.50	0.0107	0.0198	0.0347	0.0440	0.0028
		40	7.55	0.0046	0.0062	0.0191	0.0196	0.0027
		50	9.34	0.0041	0.0042	0.0161	0.0172	0.0103
5	1	20	3.51	0.0046	0.0085	0.0093	0.0114	0.0015
	2		3.50	0.0107	0.0198	0.0347	0.0440	0.0028
	5		3.45	0.0358	0.0454	0.2992	0.5238	0.0043

References

1. Barndorff-Nielsen, O. E., Blaesild, P. (1986). A note on the calculation of Bartlett adjustments, *Journal of Royal Statistical Society Series B* **48**, 353–358.
2. Cordeiro, G. M. (1987). On the corrections to the likelihood ratio statistics, *Biometrika* **74**, 265–274.
3. Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference, *Journal of Royal Statistical Society Series B* **49**, 1–39.
4. Doksum, K. A. and Normand, S. T. (1996) Models for degradation processes and event times based on gaussian processes. In: Jewell, N. P. et al. (eds), *Lifetime Data: Models in Reliability and Survival Analysis*. Netherlands: Kluwer academic publishers, 85–91.
5. Kahle, W. (1994). Simultaneous confidence regions for the parameters of damage processes. *Statistical Papers* **35**, 27–41.
6. Kahle, W. Lehmann, A. (1998) Parameter estimation in damage processes: Dependent observations of damage increments and first passage time, In: Kahle, W. et al. (eds), *Advances in Stochastic Models for Reliability, Quality and Safety*. Boston: Birkhauser, pp 139–152.
7. Lu, C. J. and Meeker, W. Q. (1993) Using degradation measures to estimate a time-to-failure distribution, *Technometrics*, **35**, 161–174.
8. Wentzell, A. D. (1979). *Theorie zufälliger Prozesse*. Berlin: Akademie-Verlag.

9. Whitmore, G. A. (1995) Estimating degradation by a Wiener diffusion process subject to measurement error, *Lifetime Data Analysis*, **1**, 307–319.
10. Whitmore, G. A., and Schenkelberg, F. (1997) Modelling accelerated degradation data using Wiener diffusion with a time scale transformation, *Lifetime Data Analysis*, **3**, 27–45.
11. Whitmore, G. A., Crowder, M. J., and Lawless, J. F. (1998) Failure inference from a marker process based on a bivariate Wiener model, *Lifetime Data Analysis*, **4**, 229–251.

On the General Degradation Path Model: Review and Simulation

Firoozeh Haghighi, Nazanin Noorae, and Narges Nazeri Rad

University of Tehran, Department of Mathematics, Statistics, and Computer Sciences,
Tehran, Iran

Abstract: In this chapter we review the estimation of survival function based on the general path model, with noise that is studied by Bagdonavičius and Nikulin (Lifetime Data Analysis 10: 65–81, 2004). We also discuss the nonparametric, semi-parametric, and parametric estimation of survival function in degradation model with multiple failure modes originated by the ignorance of measurement error described in Bagdonavičius and Nikulin (Lifetime Data Analysis 10: 65–81, 2004), Bagdonavičius et al. (Communication in Statistics-Theory and Methods 34: 1771–1793, 2005), and Haghighi et al. (Proceedings of Second International Conference on Accelerated Life Testing in Reliability and Quality Control, 2008), respectively. Ultimately, different simulation studies are conducted and the comparison of proposed estimation methods are illustrated.

Keywords and phrases: Degradation model, survival function, intensity, parametric, semi-parametric and non-parametric estimation methods, natural and traumatic failures

10.1 Introduction

Many materials degrade over time before they break down or fail. The degradation may involve, for instance, physical changes such as wearing, fracturing, fatigue of components, or chemical changes brought about by corrosion. Our knowledge about degradation is obtained by statistical modeling of degradation process. The model can be used to predict the lifetime of an item or the extent of degradation of an item at a specified future time. To make inference about the failure of such materials, we consider two models in which degradation is measured with or without errors. Furthermore, level of degradation often can be observed during the operation of unit, and in addition to observed failure time, the degradation value provides valuable information that should not be ignored. As Park and Padgett [17] mention, early work on the degradation models is referenced by Nelson [16], while more recent results are mentioned by Bagdonavičius and Nikulin [6–8] and Bagdonavičius et al. [2–5]. Also, the degradation models applied to specific problems in engineering have been presented by Meeker and Escobar

[15]. In particular, Whitmore [19] described the degradation process as a Wiener diffusion process and the extended version of his work was presented in Whitmore and Schenkelgerg [21], whereas the results of degradation process with a bivariate Wiener process could be referred to Whitmore et al. [20]. Self and Pawitan [18] and Wolfsohn and Tsiatis [22] have provided a methodology in medical application. Using regression-type method, general degradation path models studied by several authors, including Lu and Meeker [13] and Meeker et al. [15]. Bagdonavičius et al. [2] introduced the nonparametric estimations of reliability characteristics in a linear degradation model with multiple failure modes. Further studies on the semi-parametric and parametric analyses of this model have been considered by Bagdonavičius et al. [5] and Haghighi et al. [10], respectively. Some simulation studies based on the results of Bagdonavičius and Nikulin's works are considered by Couallier [9] and Haghighi (in press). Statistical analysis of degradation failure time data with renewal is given in Lehmann [12], Kahle [11], and Bagdonavičius et al. [3]. The chapter is organized as follows. The general framework of degradation model including degradation models with noise and negligible measurement errors is presented in Section 10.2. A review on various estimation procedures of survival function which consist of the nonparametric and semi-parametric methods is also considered. Section 10.3 is devoted to simulation study in order to investigate the performance of the proposed approaches.

10.2 Degradation Model

Meeker and Escobar [14] introduced the real degradation process by

$$Z_r(t) = g(t, A),$$

where $A = (A_1, \dots, A_n)$ is a random vector with positive components and the distribution function F_A and g is a specified increasing and continuously differentiable function of real time t . We denote by h the inverse function of g with respect to the first argument and by h' its partial derivative.

For more product there is a gradual loss of performance. An item is regarded as failed when degradation level reaches a critical threshold level z_0 . We call this a natural failure and the corresponding time to this failure is denoted by T^0 .

In addition to the natural failure, there are other failures whose origins are related to aging that are called traumatic failures. Several traumatic failure types are possible and we denote by T^k ($k = 1, \dots, s$) the failure time corresponding to the k th traumatic failure mode. Hence, the moment of the failure is

$$\tau = \min(T^0, T^1, \dots, T^s).$$

We suppose that the random variables T^1, T^2, \dots, T^s are conditionally independent (given $A = a$). The conditional survival function of T^k is

$$S^k(t|a) = \mathbf{P}\{T^k > t|A = a\} = e^{-\int_0^t \lambda^k(s/a)ds} = e^{-A^k(t/a)}, \quad (10.1)$$

where λ^k being the intensity function ($k = 1, 2, \dots, s$) and $A^k(z) = \int_0^z \lambda^k(y)dy$.

10.2.1 Degradation model with noise

Practically, the observed degradation $Z(t)$ deviates from the real degradation $Z_r(t)$. Thus a complete statistical model must take the source of this variation into account. Here, we describe one such model. The degradation model with noise $U(t)$ is given by

$$Z(t) = Z_r(t)U(t), \quad t \geq 0. \quad (10.2)$$

The noise in the model is taken to be

$$U(t) = \exp\{\sigma W(c(t))\}, \quad (10.3)$$

where W is the standard Wiener process that is independent of the degradation process and c is a non-negative, specified continuous, and increasing time function with $c(0) = 0$. Clearly, if $\sigma = 0$ then we have a particular case of the well-known *General Path Model* developed by Meeker and Escobar [14]. In the degradation model with noise, the component $Z_r(t)$ explains the interior degradation process which is different for each unit and the noise explains the complementary influence on the obtained disability score.

We consider the model (10.2) with only one traumatic failure mode ($k = 1$). Thus the moment of the failure is $\tau = \min(T^0, T)$. Bagdonavičius and Nikulin assumed that the conditional hazard rate is given by a proportional hazard model

$$\lambda_T(t|A) = \lambda_0(s, \alpha)\lambda(g(s, A)), \quad (10.4)$$

where λ_0 and λ are the baseline hazard function and the intensity function, respectively. So the conditional hazard rate depends on the degradation through the intensity function. If we consider the degradation process as an internal covariate this model can be compared with the well-known Cox model. In the Cox model the baseline hazard is taken to be unknown, but in (10.4) the base line hazard belongs to a parametric family. However, in (10.4) no assumption is made about the form of λ , that is, the advantage of (10.4), because we do not usually know how the degradation influence on the lifetime of an unit.

The conditional survival function of T can be written as

$$\begin{aligned} S_T(t|A) &= P\{T > t | g(s, A), 0 \leq s \leq t\} \\ &= \exp \left\{ - \int_0^t \lambda_0(s, \alpha) \lambda(g(s, A)) \, ds \right\} \\ &= \exp \left\{ - \int_{g(0, A)}^{g(t, A)} \lambda_0(h(z, A), \alpha) h'(z, A) d\Lambda(z) \right\}. \end{aligned} \quad (10.5)$$

Then the survival function of the time to failure can result as

$$\begin{aligned} S_\tau(t) &= P(\tau > t) \\ &= \int_{g(t, a) \leq z_0} S_T(t|A = a) dF_A(a). \end{aligned} \quad (10.6)$$

We suppose that $\lambda_T(t|A) = \lambda(g(s, A))$ and F_A , λ are completely unknown. It means that the conditional hazard rate depends only on the degradation values through the

intensity while no assumption is made about the form of the intensity. In this case, a nonparametric estimation of (10.6) is obtained as follows:

$$\hat{S}_\tau(t) = \int_{g(t,a) \leq z_0} \hat{S}_T(t|a) d\hat{F}_A(a), \quad (10.7)$$

where $\hat{F}_A(a)$ is the empirical distribution function, $\hat{S}_T(t|a)$ is obtained by substituting the Nelson–Aalen-type estimator $\hat{A}(z) = \sum_{Z_i \leq z, \delta_i=1} \left(\sum_{Z_j \geq z_i} h'(z, \hat{A}_i) \right)^{-1}$ in (10.5), and \hat{A}_i is the predictor of A_i .

10.2.2 Degradation model without measurement error

The measurement errors created by imperfect instruments, procedures, and environments cause a deviation from real values. Here, we assume that the measurement errors are negligible, i.e.,

$$Z(t) = g(t, A). \quad (10.8)$$

We consider that several traumatic failure modes can occur. Using the conditional survival function (10.1) the survival function of the random variable τ results as

$$\begin{aligned} S_\tau(t) &= P(\tau \geq t) = P\left\{ \min(T^0, T^1, \dots, T^s) \geq t \right\} \\ &= \int_{g(t,a) < z_0} \exp\left\{ - \int_0^z h'(y, a) d\Lambda^{(\cdot)}(y) \right\} dF_A(a), \end{aligned} \quad (10.9)$$

where $\Lambda^{(\cdot)}(z) = \sum_{k=1}^s \Lambda^{(k)}(z)$.

Various assumptions about $F_A(a)$ and $\lambda(z)$ may lead to different methods for estimating the survival function. Some of these probable assumptions are as follows.

1. **F_A and λ are completely unknown (Bagdonavičius et al. [2])**

In this case, the intensities and $F_A(a)$ are estimated by a nonparametric method. Optimal estimators of the cumulative intensities are of the Nelson–Aalen type, for more details see [1], and a nonparametric estimator of $F_A(a)$ is the empirical distribution function. From (10.9) the nonparametric estimation of survival function is obtained by

$$\hat{S}(t) = \frac{1}{n} \sum_{\frac{t}{\hat{A}_i} < z_0} \exp\left\{ -A_i \sum_{j: V_j \neq 0, Z_j \leq \frac{t}{\hat{A}_i}} \frac{1}{Y(Z_j)} \right\}, \quad (10.10)$$

where $Y(z) = \sum_{Z_i \geq z} A_i$.

2. **F_A is unknown and λ is taken from a specified parametric family (Bagdonavičius et al. [5])**

Here, we suppose that the intensities $\lambda^{(k)}(z)$ belong to a parametric class $\lambda^{(k)}(z) = \lambda^{(k)}(z, \gamma_k)$, $0 \leq z \leq z_0$, $k = 1, \dots, s$, where γ_k is a multi-dimensional parameter.

A semi-parametric estimation of survival function can result from substituting the maximum likelihood estimator of $\Lambda^{(k)}(z, \gamma_k)$ and the empirical distribution function $\hat{F}_A(a)$ in (10.9).

3. Both F_A and λ are taken from specified parametric families (Nikulin and Haghighi [in press] and Haghighi et al. [10])

The intensities $\lambda^{(k)}(z)$, $k = 1, \dots, s$ are chosen from a parametric class similar to one described in previous case and $F_A(a)$ is taken from a particular family, i.e., $F_A(a, \eta)$, where η is a multidimensional parameter. The parametric estimator of survival function is obtained by substituting the maximum likelihood estimators of $\Lambda^{(k)}(z, \gamma_k)$, $k = 1, \dots, s$ and $F_A(a, \eta)$ in (10.9). Hence,

$$\hat{S}_\tau(t) = \int_{g(t, a) \leq z_0} \exp \left\{ - \int_0^z h'(y, a) d\hat{\Lambda}^{(\cdot)}(y) \right\} dF(a, \hat{\eta}), \quad (10.11)$$

where $\hat{\Lambda}^{(\cdot)}(y) = \sum_{k=1}^s \Lambda^{(k)}(y, \hat{\gamma}_k)$ and $F_A(a, \hat{\eta})$ are the maximum likelihood estimators.

10.3 Simulation

To illustrate the statistical methods discussed in previous sections, we conduct some simulation experiments. First, we consider the degradation model with noise (10.2). Our aim is to estimate the survival function from simulated data using the method discussed in (1.2.1). The real degradation is taken to be $Z_r(t) = e^{A_1} (1+t)^{A_2}$. We have simulated $A = (A_1, A_2)$ based on a bivariate normal distribution for $n = 20, 50, 200$ items with the parameter set $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) = (1.8, 2.3, 0.2, 0.2, 0)$.

We suppose that the conditional hazard rate depends only on the degradation values through the intensity and the intensity is taken to be $\lambda(z, \theta, \nu) = (\frac{Z_r(t)}{\theta})^\nu$. Using conditional survival function (10.5), times to failure are obtained as follows:

$$T = \left[1 - \left(\frac{\theta}{e^{A_1}} \right)^\nu (\alpha + A_2 \nu + 1) U \right]^{\frac{1}{\alpha + A_2 \nu + 1} - 1},$$

where $U \sim (0, 1)$. We generate times to failure setting $\nu = 3.6$ and $\theta = 140$. The Wiener process (10.3) is generated by $W_i = W_{i-1} + \xi_i$, where ξ_i 's are taken to be independent and identically distributed (i.i.d) as $N(0, c_i - c_{i-1})$. In addition, $c_i = \ln(1 + t_i)$ and $c_0 = \ln(1 + t_0)$.

Figure 10.1 shows simulated observed paths for different values of σ based on the model (10.2).

We obtained the nonparametric estimated survival function based on the model (10.2) for simulated data using (10.7). The results are compared with the true survival function in Figure 10.2 for different sample sizes.

Here, we consider the degradation model without noise and measurement errors (10.8). We are interested in comparing the parametric and nonparametric estimated

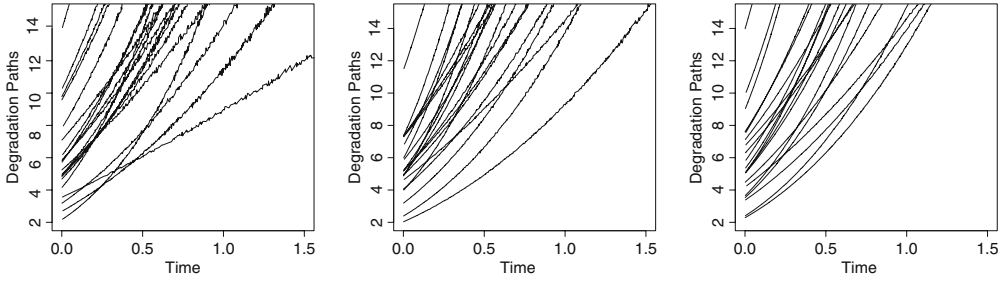


Figure 10.1. Simulated observed paths for 20 units. Above *left* ($\sigma = 0.01$), above *right* ($\sigma = 0.0055$), and *Bottom* ($\sigma = 0.0025$) [model (10.2)]

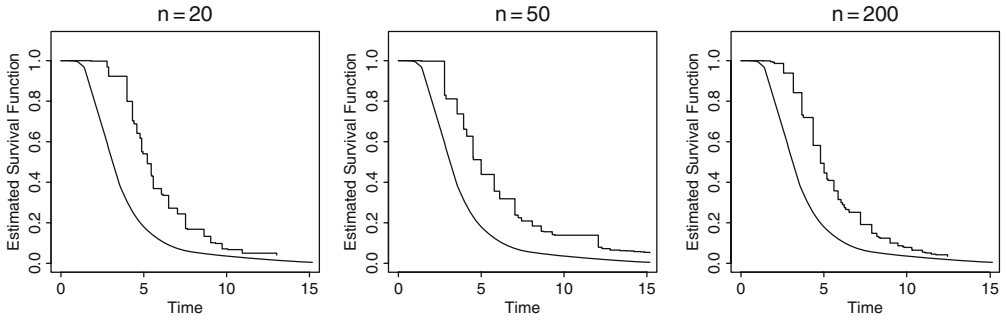


Figure 10.2. True survival function (*solid line*) versus the nonparametric estimator of survival function for $n = 20, 50, 100$ [model (10.2)]

survival function in a linear degradation model. We suppose that A to be a univariate random variable from the Weibull distribution with the parameter set $(5, 2)$. Setting $z_0 = 10$, $\lambda^{(k)}(z) = (\theta_k z)^{\nu_k}$, $k = 1, 2$, $(\theta_1, \theta_2, \nu_1, \nu_2) = (0.028, 0.075, 2, 5)$ from conditional survival function (10.1), we obtain times to failure for each mode as follows:

$$T_i^k = [-(\nu_k + 1)(\log \omega) \left(\frac{A_i}{\theta_k} \right)^{\nu_k}]^{\frac{1}{\nu_k + 1}}, \quad k = 1, 2, \quad (10.12)$$

where $\omega \sim U(0, 1)$.

Simulated observed paths are shown in Figure 10.3. From (10.11), it is observed that the parametric estimation of survival function cannot be expressed in explicit form. In order to prevail this problem, we proposed the Monte Carlo simulation method to approximate the integral in (10.11). The bias-corrected and accelerated bootstrap procedure is used for constructing pointwise confidence interval. Each Bootstrap estimation was computed using 4,000 samples and each Monte Carlo simulation involved 2000 replications. The results are shown in Figure 10.4.

A comparison between parametric and nonparametric estimations of survival function for $(n = 50, 100, 500, 1000)$ is shown in Figure 10.5. It has been illustrated that for the sample size $n = 50$, the nonparametric estimations fluctuate around the

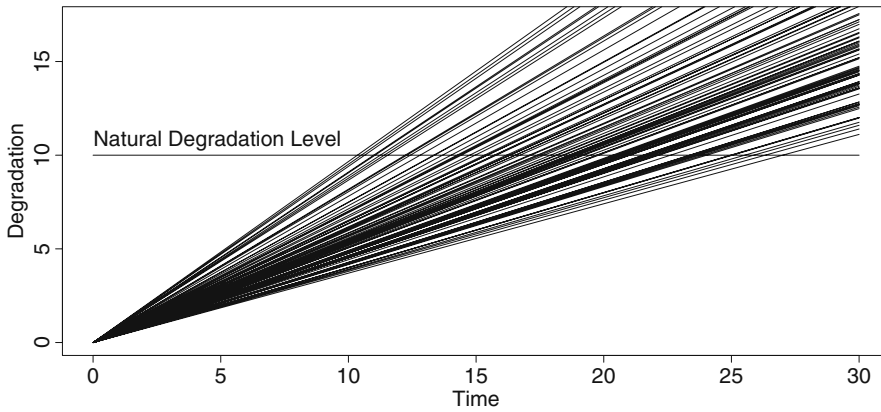


Figure 10.3. Simulated sample paths without measurement error [model (10.8)]

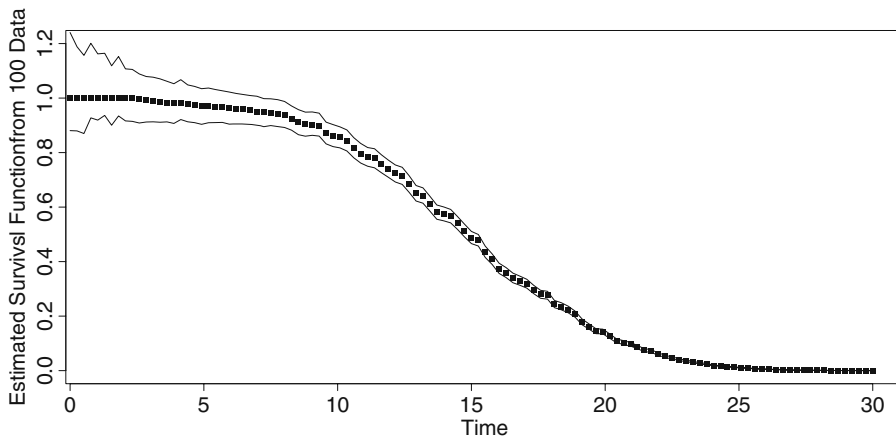


Figure 10.4. Parametric estimators of $S(t)$ and its 90% confidence intervals [model (10.8)]

curve of the parametric estimations. As the sample size increases, the fluctuation reduces, and there is a close agreement between the two methods in large sample sizes. In other words, if the sample size is small and the underlying distribution assumptions are known, the parametric approach must legitimately be used and this would not be a situation requiring the nonparametric methods. On the other hand, in the large samples there would not be any noticeable differences between the two approaches.

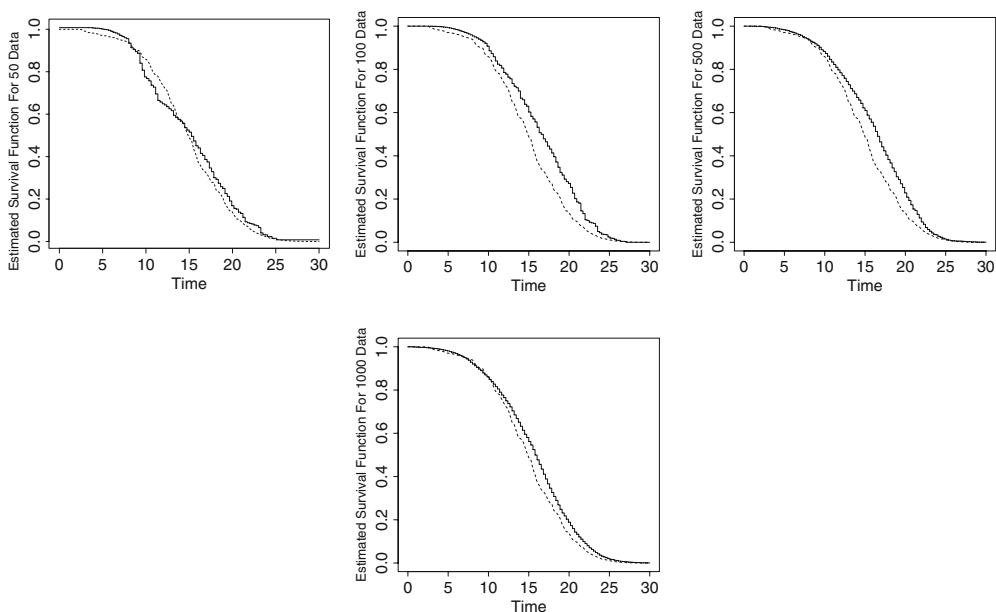


Figure 10.5. The comparison between parametric (*dotted line*) and nonparametric (*solid line*) estimators of $S(t)$ for different sample sizes [model (10.8)]

References

1. Anderson, P.K., Borgan, O., Gill, R.D., and Keiding, N.: Statistical models based on the counting processes. Springer-Verlag, New York (1993)
2. Bagdonavičius, V., Bikelis, A., and Kazakevičius, V.: Statistical and failure time data with multiple failure modes. *Lifetime Data Analysis*, **10**, 65–81 (2004)
3. Bagdonavičius, V., Bikelis, A., Kazakevičius, V., and Nikulin, M.S.: Analysis of joint multiple failure mode and linear degradation data with renewals. *Journal of Statistical Planning and Inference*, **137**, 2191–2207 (2007)
4. Bagdonavičius, V., Bikelis, A., Kazakevičius, V., and Nikulin, M.S.: Estimation from simultaneous degradation and failure time data. In: Lindqvist, B., Doksum, K.A. (ed) *Mathematical and Statistical Methods in Reliability*. World Scientific Publishing, Series on Quality, Reliability and engineering Statistics (2003)
5. Bagdonavičius, V., Haghighi, F., and Nikulin, M.S.: Statistical analysis of degradation and failure time data using the general path model. *Communication in Statistics—Theory and Methods*, **34**, 1771–1793 (2005)
6. Bagdonavičius, V. and Nikulin, M.S.: Semi-parametric analysis of degradation and failure time data with covariates. In: Nikulin, M.S., Balakrishnan, N., Mesbah, M., Limnios, N. (ed) *Parametric and Semi-parametric Models with Applications to Reliability, Survival Analysis, and Quality of Life*. Birkhäuser, Boston (2004)
7. Bagdonavičius, V. and Nikulin, M.S.: *Accelerated Life Models: Modeling and Statistical Analysis*. Chapman and Hall/CRC, Boca Raton (2002)
8. Bagdonavičius, V. and Nikulin, M.S.: Estimation in degradation models with explanatory variables. *Lifetime Data Analysis*, **7**, 83–103 (2001)

9. Couallier, V.: A competing risk model for degradation and traumatic failure time. In: Vonta, F., Nikulin, M.S., Limnios, N., Huber-Carol, C. (ed) *Statistical Models and Methods for Biomedical and Technical System*. Birkhäuser, Boston (2008)
10. Haghighi, F., Nazeri rad, N., and Noorae. N.: On the parametric degradation model with multiple failure modes. In: *Proceedings of Second International Conference on Accelerated Life Testing in Reliability and Quality Control*. IMB, University Victor Segalen Bordeaux 2, Bordeaux, France (2008)
11. Kahle, W.: Statistical models for the degree of repair in incomplete repair models. In: *Proceedings of the International Symposium on Stochastic Model in Reliability, Safety, Security and Logistics*. Sami Shamoon College of Engineering, Beer Sheva (2005)
12. Lehmann, A.: On a degradation-failure models for repairable items. In: Nikulin, M.S., Balakrishnan, N., Mesbah, M., Limnios, N. (ed) *Parametric and Semi-parametric Models with Applications to Reliability, Survival Analysis, and Quality of Life*. Birkhäuser, Boston (2004)
13. Lu, C.J. and Meeker, W.Q.: Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, **35**, 161–174 (1993)
14. Meeker, W. and Escobar, L.: *Statistical Models for Reliability Data*. Chapman and Hall/CRC, Boca Raton (1998)
15. Meeker, W.Q., Escobar, L.A., and Lu, C.J.: Accelerated degradation tests: Modeling and analysis. *Technometrics*, **40**, 89–99 (1998)
16. Nelson, W.: *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. John Wiley & Sons, New York (1990)
17. Park, C. and Padgett, W.J.: Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. *Lifetime Data Analysis*, **11**, 511–527 (2005)
18. Self, S. and Pawitan, Y.: Modeling a marker of disease progression and onset of disease. In: Jewell, N., Dietz, K., Farewell, V. (ed) *AIDS Epidemiology: Methodological Issues*. Birkhäuser: Boston (1992)
19. Whitmore, G.A.: Estimating degradation by a Wiener diffusion process subject to measurement error. *Lifetime Data Analysis*, **1**, 307–319 (1995)
20. Whitmore, G.A., Crowder, M.I., and Lawless, J.F.: Failure inference from a marker process based on a bivariate model. *Lifetime Data Analysis*, **4**, 229–251 (1998)
21. Whitmore, G.A. and Schenkelburg, F.: Modeling accelerated degradation data using Wiener diffusion with a scale transformation. *Lifetime Data Analysis*, **3**, 27–45 (1997)
22. Wolfshon, M.S. and Tisatis, A.A.: A joint model for survival and longitudinal data measured with error. *Biometrics*, **53**, 330–339 (1997)

A Closer Look at Degradation Models: Classical and Bayesian Approaches

Marta A. Freitas¹, Thiago R. dos Santos², Magda C. Pires², and Enrico A. Colosimo²

¹ Department of Industrial Engineering, Federal University of Minas Gerais, Brazil

² Department of Statistics, Federal University of Minas Gerais, Brazil

Abstract: Traditionally, reliability assessment of devices has been based on (accelerated) life tests. However, for highly reliable products, little information about reliability is provided by life tests in which few or no failures are typically observed. Since most failures arise from a degradation mechanism at work for which there are characteristics that degrade over time, one alternative is to monitor the device for a period of time and assess its reliability from the changes in performance (degradation) observed during that period. The goal of this chapter is to illustrate how degradation data can be modeled and analyzed by using “classical” and Bayesian approaches. Four methods of data analysis based on classical inference are presented. Next we show how Bayesian methods can also be used to provide a natural approach to analyzing degradation data. The approaches are applied to a real data set regarding train wheels degradation.

Keywords and phrases: Bayesian approach, degradation data analysis, reliability

11.1 Introduction

11.1.1 Background and literature

Much of the literature focuses on the use of lifetime to make reliability assessments. For products that are highly reliable, assessing reliability with lifetime data is problematic, however. For a practical test duration (or a fixed observation period of the performance on the field), few or perhaps no failures may occur so that most of the observations are censored. Such data provide little information about the proportion of products, surviving a warranty period that is orders of magnitude longer than the test duration.

Recently, degradation data have shown to be a superior alternative to lifetime data in such situations because they are more informative [2, 3, 14, 15, 33]. Most failures arise from a degradation mechanism at work, such as the progression of a chemical reaction, for which there are characteristics that degrade (or grow) over time. We consider the situation in which failure is defined in terms of an observable characteristic. For example, a crack grows over time, and failure is defined to occur when the crack

reaches a specified length. Another example is the brightness of fluorescent lights that decreases over time. Its failure is defined to occur when the lights luminosity degrades to 60% or less of its luminosity at 100 hours of use. Such failures are referred to as “soft” failures because the units are still working, but their performance has become unacceptable.

To conduct a degradation test, one has to prespecify a threshold level of degradation, obtain measurements of degradation at different fixed times, and define that failure occurs when the amount of degradation of an experimental unit exceeds that level. Thus, these degradation measurements may provide some useful information to assess reliability even when failures do not occur during the test period.

There are important references that have used degradation data to assess reliability. Nelson [22] discussed a special situation in which the degradation measurement is destructive (only one measurement could be made on each item). Nelson [23, Chapter 11] reviewed the degradation literature, surveyed applications, described basic ideas and using a specific example showed how to analyze a type of degradation data. In the literature, there are two major aspects of modeling for degradation data. One approach is to assume that the degradation is a random process in time. Doksum [6] used a Wiener process model to analyze degradation data. Tang and Chang [32] modeled nondestructive accelerated degradation data from power supply units as a collection of stochastic processes. Whitmore and Shenkelberg [34] considered that the degradation process in the model is taken to be a Wiener diffusion process with a timescale transformation. Their model and inference methods were illustrated with a case application involving self-regulating heating cables.

An alternative approach is to consider more general statistical models. Degradation in these models is modeled by a function of time and some possibly multidimensional random variables. These models are called general degradation path models. Lu and Meeker [14] developed statistical methods using degradation measures to estimate a time-to-failure distribution for a broad class of degradation models. They considered a nonlinear mixed-effects model and used a two-stage method to obtain point estimates and confidence intervals of percentiles of the failure time distribution. Tseng et al. [33] presented a case study which used degradation data and a fractional factorial design to improve the reliability of fluorescent lamps. Yacout et al. [38] used degradation data to estimate the time-to-failure distribution of metallic Integral Fast Reactor fuel pins irradiated in Experimental Breeder Reactor II. The time-to-failure distribution for the fuel pins was estimated based on a fixed threshold failure model and the two-stage estimation approach proposed by [14]. Lu et al. [16] proposed a model with random regression coefficients and standard deviation function for analyzing linear degradation data from semiconductors. Su et al. [30] considered a random coefficient degradation model with random sample size and used maximum likelihood for parameter estimation. A data set from a semiconductor application was used to illustrate their methods. Wu and Shao [36] established the asymptotic properties of the (weighted) least square estimators under the nonlinear mixed-effect model. They used these properties to obtain point estimates and approximate confidence intervals for percentiles of the failure time distribution. They applied the proposed methods to metal film resistor and metal fatigue crack length data sets.

A good reference on degradation path models is [18, Chapter 13]. Wu and Tsai [37] presented a fuzzy-weighted estimation method to modify the two-stage procedure proposed by [14]. The proposed method and the two-stage one were both studied on

the example of the metal film resistor of [36]. The former seemed to reduce the affection of different patterns of degradation paths and improve the estimation results of time-to-failure distribution providing much tighter confidence intervals. Crk [5] proposed a methodology that encompasses many of the known and published ones and went a step further by considering a component or a system performance degradation function whose parameters may be random, correlated, and stress dependent (in the case of accelerated degradation tests). Jiang and Zhang [12] presented a dynamic model of degradation data. Random fatigue crack growth was illustrated in detail as an example of degradation data problem. Important issues related to the design of degradation tests can be found, for example, in [39, 35, 40, 17, 41].

11.1.2 The problem

In a degradation test, measurements of performance are obtained as it degrades over time for a random sample of test units. Thus, the general approach is to model the degradations of the individual units using the same functional form and differences between individual units using random effects. The model is

$$y_{ij} = D_{ij} = D(t_{ij}; \alpha; \beta_i) + \varepsilon_{ij}, \quad (11.1)$$

$i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m_i$, where $D(t_{ij}; \alpha; \beta_i)$ is the actual degradation path of unit i at a prespecified time t_{ij} ; $\alpha = (\alpha_1; \alpha_2; \dots; \alpha_p)^t$ is a vector of fixed effects that describes population characteristics (they remain constant for all units); $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})^t$ is a vector of the i th unit random effects that representing the individual unit's characteristics (variations in the properties of the raw material, in the production process, in the component dimensions, etc.), and ε_{ij} is the associated random error of the i th unit at time t_{ij} .

The deterministic form of $D(t_{ij}; \alpha; \beta_i)$ might be based on empirical analysis of the degradation process under study, but whenever possible it should be based on the physical-chemical phenomenon associated with it. The ε_{ij} are assumed to be independently and identically distributed (*i.i.d*) with mean zero and unknown variance σ_ε^2 .

The random vectors $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})^t$ ($i = 1, \dots, n$) are independently distributed as $A_\beta(\theta)$, where $A_\beta(\theta)$ is a multivariate distribution function, which may depend on an unknown parameter vector $\theta = (\theta_1, \dots, \theta_q)^t$ that must be estimated from the data, and $\{\varepsilon_{ij}\}$ and $\{\beta_i\}$ are assumed independent. It is also assumed that y and t are in appropriately transformed scales, if needed.

The proportion of failures at time t is equivalent to the proportion of degradation paths that exceed the critical level D_f at that time t . Thus, it is possible to define the distribution of the time-to-failure T for model (11.1) as follows:

$$F_T(t) = F_T(t; \alpha; A_\beta(\cdot); D_f; D) = P(T \leq t) = P[D(t; \alpha; \beta) \geq D_f]$$

when the degradation measurements are increasing with time or

$$F_T(t) = F_T(t; \alpha; A_\beta(\cdot); D_f; D) = P(T \leq t) = P[D(t; \alpha; \beta) \leq D_f]$$

when the degradation measurements are decreasing with time.

Under this degradation model, one has to get the estimates of α (the vector of fixed effects) and θ , the parameter vector of the random effects distribution Λ_β , in order to estimate the percentiles of failure time distribution.

For simple path models, once $\Lambda_\beta(\theta)$ is known, the distribution function $F_T(t)$ can be expressed in a closed form. For many path models, however, this is not possible. When the functional form of $D(t; \alpha; \beta)$ is nonlinear and the model has more than one random parameter (in other words, the parameter vector β has dimension >1), the specification of $F_T(t)$ becomes quite complicated. Usually, one will have to evaluate the resulting forms numerically. More generally, one can obtain numerically the distribution of T for any specified α , $\Lambda_\beta(\theta)$, D_f , and D (i.e., the model parameters, the critical degradation level, and the degradation path model) by using Monte Carlo simulation. However, this can only be done if the fixed parameters α and the parameters θ of the random effects distribution $\Lambda_\beta(\theta)$ could be somehow estimated. So, even for a given $\Lambda_\beta(\theta)$ the problem remains on the parameter estimation.

The authors in [14] worked on the problem of parameter estimation and proposed a two-stage method for the case where the vector of random effects β or some appropriate reparametrization follows a multivariate normal distribution (MVN) with mean μ_β and variance-covariance matrix Σ_β . In other words, in this case, $\Lambda_\beta(\theta) = \Lambda_\beta(\mu_\beta; \Sigma_\beta) = MVN(\mu_\beta; \Sigma_\beta)$. Since full maximum likelihood estimation of random effect parameters μ_β and Σ_β is, in general, algebraically intractable and computationally intensive when they appear nonlinearly in the path model, the authors proposed this two-stage method as an alternative to the computationally intensive ones. Simulation studies showed that the method compared well with the more computationally intensive methods.

Pinheiro and Bates [24] used the method of [13] to obtain an approximated maximum likelihood estimate of the parameters μ_β , Σ_β , and σ_ε^2 . The LME (linear mixed-effects models) and NLME (nonlinear mixed-effects models) functions, written in the S-PLUS language, were developed to attain this goal [25]. In other words, these functions were developed for the specific case where the random effects follow a multivariate normal distribution.

Meeker and Escobar [18] used the numerical method with the NLME function developed by [24, 25] in a number of examples. In all of them, the failure time distribution was estimated numerically using Monte Carlo simulation. In addition, the authors presented two other methods of degradation data analysis, namely the approximate and the analytical method. Both of them are difficult to apply when the degradation path model is nonlinear and has more than one random parameter.

The methods described so far rely on maximum likelihood or least squares estimation of the model parameters (the so-called classical inference procedures) and Monte Carlo simulation. An alternative approach to degradation data analysis is to use Bayesian methods. In particular, because reliability is a function of the parameters of the degradation model, the posterior distribution for reliability at a specified time is straightforward to obtain from the posterior distribution of the model parameters. Hamada [11] used a Bayesian approach for analyzing a laser degradation data but the author did not compare the results with the non-Bayesian approaches available.

The goal of this chapter is to illustrate how degradation data can be modeled and analyzed by using “classical” and Bayesian approaches. We use the general degradation path model to model degradation data and the mixed-effect model proposed by

[14]. Four methods of data analysis are implemented: the approximate, the analytical, the numerical (as presented by [18]), and the two-stage method [14]. Next we show how Bayesian methods can also be used to provide a natural approach to analyzing degradation data. The approaches are applied to a real data set regarding train wheels degradation.

The outline of the chapter is as follows: Section 11.2 presents the real motivating situation. Three methods based on “classical” inference as well as the Bayesian approach are briefly presented in Section 11.3. The “Train Wheel degradation data” is analyzed in Section 11.4. Conclusions and final comments end the chapter in Section 11.5.

11.2 Train Wheel Degradation Data

Wheel failures, which account for half of the train derailments, cost billions of dollars to the global rail industry. Wheel failures also accelerate rail deterioration. To minimize rail breaks and help avoid catastrophic events such as derailments, railways are now closely monitoring the performance of wheels and trying to remove them before they start badly affecting the rails.

Most railways keep in a database detailed descriptions of all maintenance actions performed on their trains. The data used in this chapter is just a small subset of such database. It refers to a larger study being conducted by a Brazilian railway company. The complete database includes, among other information, the diameter measurements of the wheels, taken at 13 equally spaced inspection times: $t_0 = 0$ km; $t_1 = 50,000$ km; $t_2 = 100,000$ km; \dots ; $t_{13} = 600,000$ km. These measurements were recorded for 14 trains, each one composed of 4 cars (CA1, CA2, CA3, CA4). A wheel’s location in a particular car within a given train is specified by an axle number (1, 2, 3, 4 – number of axles on the car) and the side of the wheel on the axle (right or left).

In this preliminary study, special attention was given to the CA1 cars because these are the ones responsible for pushing the other three cars in a given train. It is known that the operating mode of such cars accelerates the degradation process of its wheels. Therefore, the data used in this chapter refers to the diameter measurements of the wheels located on the left side of axle number 1 of each one of the CA1 cars.

The diameter of a new wheel is 966 mm. When the diameter reaches 889 mm the wheel is replaced by a new one. Figure 11.1 presents the degradation profiles of the 14 wheels under study. Instead of plotting the diameters itself, the curves were constructed using the degradation observed at each evaluation time t (i.e., $966 - [\text{observed diameter measured at time } t]$). “Failure” of the wheel is then defined to occur when the degradation reaches the threshold level $D_f = 77$ mm. Note that 3 out of 14 units studied achieved the threshold level during the observation time.

The main purpose here is to use the degradation measurements to estimate the lifetime distribution of those train wheels. Once this distribution is obtained, one can get estimates of other important characteristics such as the MTTF (mean-time-to-failure or, specifically, mean covered distance), quantiles of the lifetime distribution. The profiles are shown in Figure 11.1.

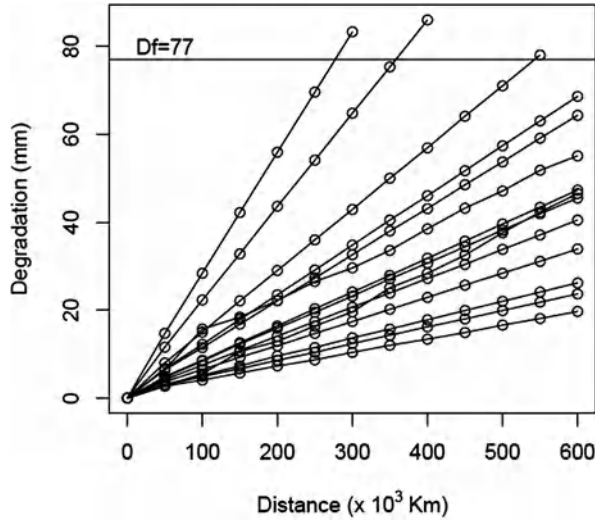


Figure 11.1. Plot of the wheel degradation data

11.3 Statistical Methods for Degradation Data Analysis

In this section, “classical” and Bayesian methods are presented. First, the four methods based on “classical” inference are briefly presented (Section 11.3.1). Next, in Section 11.3.2 the Bayesian approach is described.

11.3.1 Methods based on “classical” inference

The main purpose of a statistical analysis of degradation data is to get an estimate of the failure time distribution $F_T(t)$. Therefore, for a given degradation path model, two main steps are involved in such analysis: (1) the estimation of model parameters and (2) the evaluation of $F_T(t)$.

For some particularly simple path models, $F_T(t)$ can be expressed as a simple function, and simple methods, such as the approximate and the analytical, can be used to estimate $F_T(t)$. The two-stage and the numerical methods are more complete and make the estimation of $F_T(t)$ possible in any situation. These methods are described below.

The approximate method

Consider the general degradation model (11.1) given in Section 11.1.2. The approximate method comprises two steps. The first one consists of a separate analysis for each unit to predict the time at which the unit will reach the critical degradation level (D_f) corresponding to failure. These times are called “pseudo” failure times. In the second step, the “pseudo” failure times are analyzed as a complete sample of failure times to estimate. Formally, the method is as follows:

- For the unit i , use the path model $y_{ij} = D_{ij} + \varepsilon_{ij}$ and the sample path data $(t_{i1}, y_{i1}), \dots, (t_{im_i}, y_{im_i})$ to find the (conditional) maximum likelihood estimate of $\alpha_i = (\alpha_{i1}; \alpha_{i2}; \dots; \alpha_{ip})^t$ and $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})^t$, say $\hat{\alpha}_i$ and $\hat{\beta}_i$. This can be done using least squares (linear or nonlinear depending on the functional form of the degradation path).
- Solve the equation $D(t, \hat{\alpha}_i, \hat{\beta}_i) = D_f$ for t and call the solution \hat{t}_i .
- Repeat the procedure for each sample path to obtain the pseudo failure times $\hat{t}_1, \dots, \hat{t}_n$.
- Do a single distribution time-to-failure analysis [23] of the data $\hat{t}_1, \dots, \hat{t}_n$ to estimate $F_T(t)$.

The approximate method is simple and intuitively appealing. However, it is only adequate in cases where the degradation path $D(t)$ is relatively simple; the degradation model considered is sufficiently appropriate; there is enough degradation data to accurately estimate α_i and β_i ; the magnitude of the errors is small; and finally, the magnitude of the extrapolation needed to predict the failure times is small [18]. Note that this method considers the model parameters as fixed. Moreover, Meeker and Escobar [18] state that the method presents the following problems: (1) it ignores the uncertainty in the prediction of the “pseudo” failure times \hat{t}_i and does not consider the errors involved in the observed degradation values; (2) the distribution of the “pseudo” failure times does not generally correspond to the one that would be indicated by the degradation model; and (3) in some cases, the volume of degradation data collected can be insufficient for estimating all the model parameters. The authors also emphasize that in these scenarios, it might be necessary to fit different models for different units to predict the “pseudo” failure times.

The analytical method

For some simple path models, $F_T(t)$ can be expressed in a closed form. The following example provides an illustration of such a case.

Suppose that the actual degradation path of a particular unit is given by

$$D(t) = \alpha + \beta t,$$

where α is fixed and represents the common initial amount of degradation of all the test units at the beginning of the test ($D(0) = \alpha$), and therefore corresponds to a fixed effect; β is the degradation rate that varies from unit to unit and corresponds to a random effect.

Assuming that β varies from unit to unit according to a lognormal distribution with parameters μ and σ ($\beta \sim \text{lognormal}(\mu_\beta, \sigma_\beta)$), it is possible to define the distribution function of T by

$$\begin{aligned} F_T(t) &= P\left(\frac{D_f - \alpha}{\beta} \leq t\right) = P\left(\beta \geq \frac{D_f - \alpha}{t}\right) = 1 - \Phi_{\text{nor}}\left(\frac{\log(D_f - \alpha) - \log(t) - \mu}{\sigma}\right) \\ &= \Phi_{\text{nor}}\left(\frac{\log(t) - [\log(D_f - \alpha) - \mu]}{\sigma}\right), \quad t > 0, \end{aligned}$$

where Φ_{nor} is the cumulative distribution function of a standard normal distribution. In this case, T is also lognormal with location and scale parameters given by $\mu_T =$

$[\log(D_f - \alpha) - \mu]$ and $\sigma_T = \sigma_\beta$. Other distributions can be used and the same procedures for obtaining the distribution for T can also be applied. Results with other distributions like the Weibull, normal (Gaussian), and the Bernstein distribution can be found in [14].

The two-stage method

To carry out the two-stage method of parameter estimation, the following steps should be implemented:

Stage 1.

1. In the first stage, for each sampled unit, fit the degradation model (11.1) (using least squares) and obtain the stage 1 estimates of the model parameters $(\alpha; \beta_i)$. In other words, for each unit i ($i = 1, 2, \dots, n$) $(\hat{\alpha}_i; \hat{\beta}_i)$ are the least squares estimators of the fixed and random model parameters. In addition, an estimator of the error variance σ_ε^2 obtained from the i th unit is the mean square error $\hat{\sigma}_{\varepsilon_i}^2 = [\sum_{j=1}^{m_i} \{y_{ij} - D(t_{ij}; \hat{\alpha}_i; \hat{\beta}_i)\}^2 / (m_i - q)]$, where $q = p + k$.
2. Assume that by some appropriate reparameterization (e.g., using a Box-Cox transformation) $\hat{\varphi}_i = H(\hat{\beta}_i)$ is approximately multivariate normally distributed with the asymptotic mean μ_φ and variance-covariance matrix Σ_φ .

Stage 2.

In the second stage, the unconditional estimators, from the preceding discussion, $(\hat{\alpha}_i, \hat{\varphi}_i)$ ($i = 1, 2, \dots, n$) are combined to construct the two-stage estimators of the path-model parameters. The two-stage estimators of the path-model parameters α, μ_φ , and Σ_φ are, respectively,

$$\hat{\alpha} = \frac{\sum_{i=1}^n \hat{\alpha}_i}{n}$$

and

$$\hat{\Sigma}_\varphi = (\sum_{i=1}^n (\hat{\varphi}_i - \hat{\mu}_i)(\hat{\varphi}_i - \hat{\mu}_i)^t / (n - 1)) - (\sum_{i=1}^n \hat{v} \hat{a} r_\varepsilon(\hat{\varphi}_i) / n).$$

The point estimate $\hat{F}_T(t)$ can be evaluated to any desired degree of precision by using Monte Carlo simulation. This is done by generating a sufficiently large number of random sample paths from the assumed path model with the estimated parameters and using the proportion failing as a function of time as an estimate of $F_T(t)$. The basic steps are the following:

1. Estimate the path-model parameters α, μ_φ , and Σ_φ from the n sample paths by using the two-stage method, giving $\hat{\alpha}$, $\hat{\mu}_\varphi$, and $\hat{\Sigma}_\varphi$.
2. Generate N simulated realizations $\tilde{\varphi}$ of φ from $N(\hat{\mu}_\varphi, \hat{\Sigma}_\varphi)$ and obtain the corresponding N simulated realizations β^* of β from $H^{-1}(\tilde{\varphi})$, where N is a large number (e.g., $N = 100,000$) and H^{-1} is the inverse transformation of H . Note that in the cases where the distribution of $F_\beta(\cdot)$ of β is known, N simulated realizations β^* of β can be generated directly from this distribution. These values can then be used in steps 3 and 4 described below.
3. Compute the N simulated failure times t^* by substituting β^* into $D_f = D(\beta; \hat{\alpha}; t)$.

4. Estimate $F_T(t)$ from the simulated empirical distribution

$$\hat{F}_T(t) = [\text{Number of simulated first crossing times } t^* \leq t] / N$$

for any desired values of t .

The Monte Carlo approximation error is easy to evaluate by using the binomial distribution. This error can be made arbitrarily small by choosing the Monte Carlo sample size N to be large enough. Pointwise confidence intervals can be constructed by the bootstrap procedures [7].

The numerical method

This method has been presented by [18]. It is useful specially in practical situations described by nonlinear models which include more than one random effect. In such cases, estimation of the model parameters needs to be done by maximization of the likelihood function numerically, since it is very difficult to get a closed-form expression for $F_T(t)$. A brief presentation of the method follows next. For more details and theoretical aspects, the reader should refer to [18].

Suppose that in the general degradation path model (1), the parameter vector $\Theta^t = (\alpha; \beta) = (\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_k)$ follows a multivariate normal distribution (MVN), with mean vector μ_Θ and variance-covariance matrix Σ_Θ . In addition, suppose that the random errors ε_{ij} follow a normal distribution with mean zero and constant variance σ_ε^2 . The assumption of MVN distribution for Θ allows the information of the unit path $D(t)$ to be concentrated only on the parameters μ_Θ and Σ_Θ without loss of information. For the fixed effects components of Θ , the values are set equal to the proper effects and the respective variance and covariance terms involving the fixed effects are set equal to zero.

The estimation of μ_Θ , Σ_Θ , and σ_ε^2 is carried out from the following likelihood function:

$$L(\mu_\Theta, \Sigma_\Theta, \sigma_\varepsilon | Data) = \prod_{i=1}^n \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left[\prod_{j=1}^{m_i} \frac{1}{\sigma_\varepsilon} \Phi(\zeta_{ij}) \right] f_\Theta(\Theta; \mu_\Theta; \Sigma_\Theta) d\Theta, \quad (11.2)$$

where $\zeta_{ij} = [y_{ij} - D(t_{ij}, \Theta_i)] / \sigma_\varepsilon = [y_{ij} - D(t_{ij}, \alpha_{i1}, \dots, \alpha_{ip}, \beta_{i1}, \dots, \beta_{ik})] / \sigma_\varepsilon$ and $f_\Theta(\Theta, \mu_\Theta, \Sigma_\Theta) = f_\Theta(\alpha_{i1}, \dots, \alpha_{ip}, \beta_{i1}, \dots, \beta_{ik}; \mu_\Theta, \Sigma_\Theta)$ is the multivariate normal density function.

Pinheiro and Bates [24] used the results developed by [13] to obtain the approximation maximum likelihood estimate of the parameters μ_Θ , Σ_Θ , and σ_ε^2 . The LME (linear mixed-effects models) and NLME (nonlinear mixed-effects models) functions, written in the S-PLUS language, were developed to attain this goal [25].

After the estimation of μ_Θ , Σ_Θ , and σ_ε^2 , $F_T(t)$ can be obtained numerically by direct integration. The amount of computational time needed to evaluate the multi-dimensional integral will, however, increase exponentially with the dimension of the integral. An alternative procedure is to evaluate $F_T(t)$ numerically using Monte Carlo simulation. This simulation is carried out using the estimates of the parameters μ_Θ , Σ_Θ , and σ_ε^2 that are supplied by the LME or NLME functions. N possible degradation paths $D(t)$ are generated and for each one of them the “failure time” (crossing time

or, equivalently, the time when the degradation path first crosses the line $y = D_f$ is obtained to calculate the values of $\hat{F}_T(t)$ using the expression

$$\hat{F}_T(t) = \frac{\text{Number of simulated first crossing times} \leq t}{N}, \quad (11.3)$$

where t is a fixed instant of time and N must be a large number (usually $N \geq 10^5$).

To simulate the N paths of $D(t)$, it is necessary to generate N possible realizations of the vector $\Theta^t = (\alpha; \beta) = (\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_k)$ from a MVN distribution with mean $\hat{\mu}_\Theta$ and variance-covariance matrix $\hat{\Sigma}_\Theta$. The last step consists of applying (3). An algorithm showing the whole sequence of estimation steps for the numerical method was presented by [38]. Confidence intervals can be obtained using a resample method, as the bootstrap [7].

11.3.2 Bayesian inference

Consider the general degradation path model given by expression (11.1). For that model the β_i 's ($i = 1, 2, \dots, n$) are assumed to be independently distributed as $A_\beta(\theta)$, where $A_\beta(\theta)$ is a multivariate distribution function, which may depend on an unknown parameter vector $\theta = (\theta_1, \dots, \theta_q)^t$ that must be estimated from the data. In addition, ε_{ij} and β_i are assumed independent and the random errors ε_{ij} 's ($i = 1, \dots, n; j = 1, \dots, m_i$) are assumed to be independently and identically distributed (iid) with mean zero and unknown variance. Under this degradation model, one has to get the estimates of the unknown model parameters $\eta = (\alpha; \theta; \sigma_\varepsilon^2)$ in order to estimate the failure time distribution.

Bayesian inference provides a way to estimate the unknown model parameters and to assess their uncertainty through the resulting parameter posterior distribution. It does so by combining prior information about $\eta = (\alpha; \theta; \sigma_\varepsilon^2)$ with the information about $\eta = (\alpha; \theta; \sigma_\varepsilon^2)$ contained in the data. The prior information is described by a probability density function $\pi(\eta)$ known as the prior, and the information provided by the data is captured by the data sampling model $l(Data|\eta) = l(y|\eta) = l(y_{11}, \dots, y_{1m_1}, \dots, y_{n1}, \dots, y_{nm_n}|\eta)$ known as the likelihood. The combined information is then described by another probability density function $\pi(\eta|y)$ called the posterior. Bayes theorem provides the way to calculate the posterior, namely

$$\pi(\eta|y) = l(y|\eta)\pi(\eta) / \int l(y|\omega)\pi(\omega)d\omega, \quad (11.4)$$

where $\int l(y|\omega)\pi(\omega)d\omega$ is the marginal density of y .

The problem here is usually to calculate the integral in (11.4) as well as those necessary to get marginal posteriors from $\pi(\eta|y)$. However, recent advances in Bayesian computing make it easy to sample from the posterior of the model parameters [10, 9, 1, 4]. The sampling is accomplished through Markov chain Monte Carlo (MCMC) simulation [8]. It also turns out that it is more convenient to work with samples to provide inference for the reliability function. For the wheel degradation data, the Bayesian software WinBugs [29] was used to carry out Bayesian inference. WinBugs is freely available from the Web at <http://www.mrc-bsu.cam.ac.uk/bugs/> and can easily implement MCMC.

Although the posterior distribution $\pi(\eta|y)$ summarizes all the information about η once the data y is observed, in some cases it is convenient to summarize this information

in a single quantity. In a Bayesian framework it is necessary to first specify the losses consequent on making a decision d when various values of the parameter η pertain.

For a **real-valued** parameter η and a loss function $L(\eta, d)$, the Bayes estimator is the value d which minimizes the posterior expected loss. In other words,

$$\hat{\eta}_B = \min_d E(L(\eta, d)|y) = \min_d \int L(\eta, d)\pi(\eta|y)d\eta. \quad (11.5)$$

Different loss functions lead to different Bayes estimators. If the quadratic loss function $L(\eta, d) = (\eta - d)^2$ is used, then $d = \hat{\eta}_B = E(\eta|y)$ (the posterior mean). It can be shown that the choices $L(\eta, d) = |\eta - d|$ and the “0–1” loss generate, respectively, the posterior median and the posterior mode as Bayesian estimators [19].

In addition to point summaries, it is always important to report posterior uncertainty. The usual approach is to present quantiles of the posterior distribution of the quantities of interest. A slightly different method of summarizing posterior uncertainty is to compute a region of highest posterior density (HPD): the region of values that contains $100(1 - \alpha)\%$ of the posterior probability and also has the characteristic that the density within the region is never lower than that outside.

High posterior density regions (HPD) were calculated using the package Coda [26] implemented in the software R [27].

11.4 The Wheel Degradation Data Revisited

In this section, we return to the situation described in Section 11.2 using the degradation data presented in Figure 11.1. Recall that each one of the wheels has the same initial diameter (degradation = 0) at the beginning of the observation period. We are interested in the distribution of the time it takes to achieve a degradation threshold level of $D_f = 77$ mm. Only 3 out of 14 units studied achieved the threshold level during the observation time. Therefore, in a traditional failure time analysis, 11 out of 14 observations (78%) would be considered as censored. First, the results of the degradation data analysis using the approximate and the numerical method are presented. Then, these results are compared to the traditional failure time analysis (parametric models are fitted to the censored and uncensored time-to-failure data). The approximate degradation analysis and the failure time data analysis were done with Minitab [20] (v.15). The numerical analysis was done with R (v. 2.4, 2006) [27].

The statistical model for the data displayed in Figure 11.1 can be succinctly stated as

$$Y_{ij} = \frac{1}{\beta_i} t_{ij} + \varepsilon_{ij}, i = 1, \dots, 14 \text{ (wheels)}; j = 1, 2, \dots, 12 \text{ (measurement times)}, \quad (11.6)$$

where the reciprocal slope β_i is the i th unit random effect and represents an individual unit's characteristics (variations in the properties of the raw material, in the production process, in the component dimensions, etc.) and ε_{ij} is the associated random error of the i th unit at time t_{ij} . It is assumed that (1) the β_i 's and the ε_{ij} 's are independent and (2) the ε_{ij} 's are independently and identically distributed $N(0, \sigma_\varepsilon^2)$.

In this section, the wheel degradation data is analyzed. Sections 11.4.1–11.4.4 describe the data analysis based on each one of the “classical” methods. All the results are summarized in Table 11.2. Figures 11.4–11.6 summarize them graphically, including the confidence intervals that have been constructed in each case. Comments regarding the comparison of these results are left to Section 11.4.5. The Bayesian approach to this practical situation is described in Section 11.4.6.

11.4.1 Estimation of $F_T(t)$ using the approximate method

A separate degradation model given by the expression (11.6) was fitted to each sample unit i and least squares estimators $\hat{\beta}_i$ of β_i ($i = 1, 2, \dots, 14$) were calculated. Note that, by doing this, the model parameter β_i is assumed to be fixed. The calculation of the “pseudo” failure distance for each wheel unit was carried out from the values of $\hat{\beta}_i$, using the expression $\hat{t}_i = D_f/\hat{\beta}_i = 77/\hat{\beta}_i$. The results of this step are displayed in Table 11.1.

Table 11.1. Pseudo failure distances

Wheel Pseudo failure distance ($\times 10^3\text{km}$)	
1	2296.14
2	1912.68
3	1345.96
4	1004.08
5	669.13
6	968.66
7	1046.43
8	1732.50
9	713.79
10	1132.67
11	356.51
12	275.66
13	800.08
14	546.70

Probability plots and residual analysis were used to investigate the adequacy of several distributions to the data displayed in Table 11.1 (pseudo failure distances). Figure 11.2 shows probability plots for the Weibull and lognormal distributions. Figure 11.3 compares the Kaplan–Meier nonparametric point estimates ($\hat{R}_{KM}(t)$) of the reliability function $R_T(t)$ [18] and the maximum likelihood (ML) estimates $\hat{R}_W(t)$ and $\hat{R}_{LN}(t)$ obtained from the Weibull and lognormal models, respectively, all evaluated at the pseudo failure times (pseudo failure distances) shown in Table 11.1.

Some observations from Figures 11.2 and 11.3 are the following:

- Figure 11.2 shows that either the lognormal or Weibull distribution can be used to fit the data displayed in Table 11.1.
- Comparing Figures 11.3(a,b) shows that the points on Figure 11.3(b) (lognormal) lie closer to the line “ $y = x$ ” than the points on the Weibull plot. This pattern

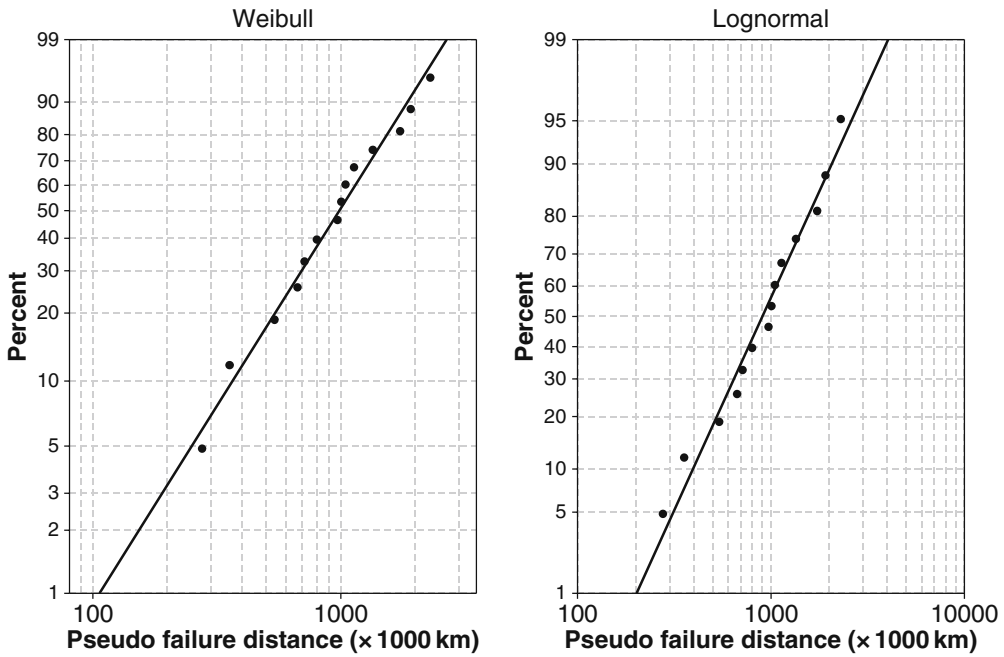


Figure 11.2. Probability plots for the pseudo failure distances

indicates that the parametric point estimates provided by the lognormal model are closer to the empirical estimates than the ones provided by the Weibull model,

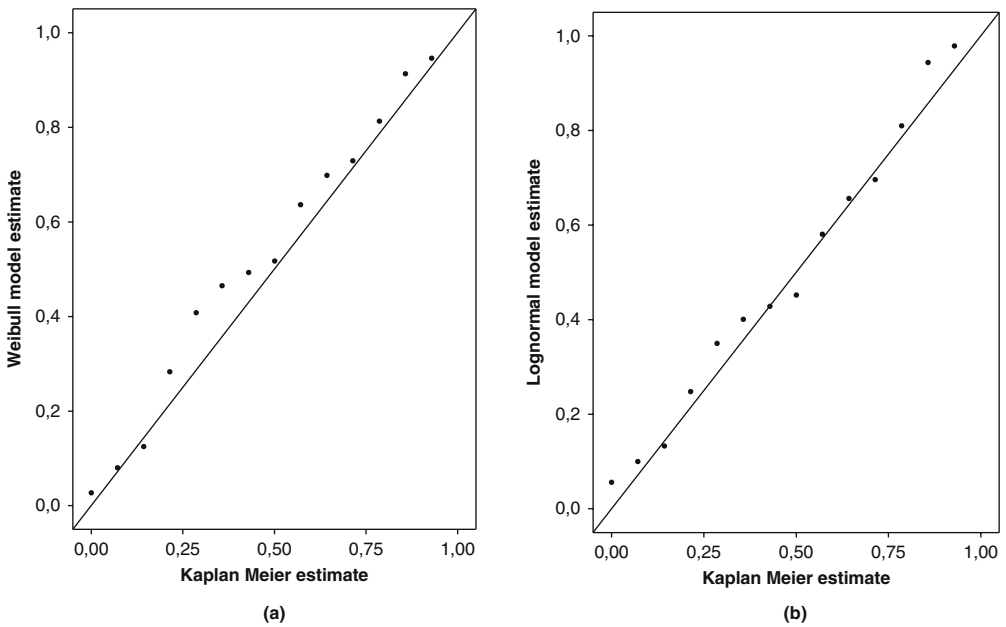


Figure 11.3. Comparison of the parametric and nonparametric (Kaplan–Meier) estimates of $R_T(t)$ at each evaluation point

indicating a slightly better performance of the former. However, due to sample size restrictions (only 14 wheels), it was decided to go on the analysis considering both distributions.

Table 11.2 shows the results obtained from the Weibull and lognormal models. The estimated values of the average distance covered by the wheel (MTTF) are 1,060,880 Km (Weibull model) and 1,071,870 km (lognormal model). Other quantities of interest are the median distance ($t_{0.50}$) and the reliability at 300,000 km. Table 11.2 displays these estimates as well as the 95% asymptotic confidence intervals. We note that the widths of the confidence intervals provided by the lognormal model are smaller than the respective ones calculated by the Weibull model, indicating a higher precision of the estimated values provided by the former.

Table 11.2. Interval and point estimates obtained by each method

Method	Distribution	Estimates			
		$MTTF^{(1)}$	$t_{0.10}^{(1)}$	$t_{0.50}^{(1)}$	$R(300,000)$
Approximate ⁽²⁾	Weibull	1060.88	383.35	994.25	0.937
		[803.86; 1400.08]	[207.92; 706.82]	[727.36; 1359.07]	[0.771; 0.984]
	Lognormal	1071.87	426.28	902.64	0.970
		[769.02; 1494.00]	[281.77; 644.91]	[664.40; 1227.14]	[0.844; 0.997]
Analytical ⁽³⁾	Weibull	1060.20	383.7	994.91	0.937
		[784.3; 1370.4]	[240.7; 684.41]	[730.0; 1329.41]	[0.860; 0.989]
	Lognormal	1071.91	426.32	903.02	0.970
		[773.22; 1401.62]	[287.5; 713.9]	[657.8; 1228.1]	[0.889; 1.00]
Two stages ⁽³⁾	Weibull	1061.80	365.8	983.9	0.930
		[785.5; 1374.4]	[223.3; 627.53]	[732.5; 1322.91]	[0.858; 0.983]
	Lognormal	1109.21	395.33	899.94	0.956
		[729.43; 1422.81]	[264.82; 687.9]	[648.81; 1217.23]	[0.876; 0.985]

(1) values should be multiplied by 10^3 km (2) point estimates and asymptotic 95% confidence intervals (3) point estimates and (nonparametric) bootstrap 95% confidence intervals.

11.4.2 Estimation of $F_T(t)$ using the analytical method

1. For each sampled unit, the degradation model (11.6) was fitted to the sample paths and the estimates of the model parameters were obtained using the least squares estimation method. These estimated values ($\hat{\beta}_i$) are exactly the ones that have already been calculated in order to get the pseudo covered distances displayed in Table 11.1.
2. The degradation model (11.1) postulated for the wheel profiles is very simple since it is a straight line with one random parameter only. In addition, the analysis of probability plots constructed to the $\hat{\beta}_i$ (they are not shown here) indicated that either the lognormal or Weibull distribution could be used to fit those values. Therefore, the failure time distribution $F_T(t)$ can be obtained directly, using the following relationships:

- Weibull case

$$\beta \sim Weibull(\delta_\beta, \lambda_\beta) \Leftrightarrow T \sim Weibull(\delta_T, \lambda_T),$$

where

$$\lambda_T = D_f \lambda_\beta; \delta_T = \delta_\beta$$

and

$$f_T(x) = \frac{\delta_T}{\lambda_T^{\delta_T}} x^{\delta_T-1} \exp \left\{ - \left(\frac{x}{\lambda_T} \right)^{\delta_T} \right\}.$$

- Lognormal case

$$\beta \sim lognormal(\mu_\beta, \sigma_\beta) \Leftrightarrow T \sim lognormal(\log D_f + \mu_\beta, \sigma_\beta).$$

Consequently, the data analysis was carried on according to the following steps:

1. Lognormal and Weibull models were fitted to the $\hat{\beta}_i$ values and the maximum likelihood estimates $\hat{\mu}_\beta, \hat{\sigma}_\beta$ (lognormal case) and $\hat{\delta}_\beta, \hat{\lambda}_\beta$ (Weibull case) were calculated.
2. Next, the parameters of the failure time distribution were obtained by using the expressions above. The results are summarized below.

- **Lognormal case**

$$\begin{aligned} \beta &\sim lognormal(2.46192; 0.644248) \\ \Leftrightarrow T &\sim lognormal([\log 77 + 2.46192]; 0.644248) \\ \Leftrightarrow T &\sim lognormal(6.805725; 0.644248). \end{aligned}$$

- **Weibull case**

$$\begin{aligned} \beta &\sim Weibull(1.976719; 15.54349) \\ \Leftrightarrow T &\sim Weibull(1.976719; 77 \times 15.54349) \\ \Leftrightarrow T &\sim Weibull(1.976719; 1196.84873). \end{aligned}$$

Table 11.2 summarizes the results based on the two distributions. A total of 95% bootstrap confidence intervals were obtained for each one of the quantities of interest. For almost all of them it would have been possible to calculate asymptotic confidence intervals using the delta method [21, Chapter 5, p. 181]. One exception is the MTTF for which the calculations are not straightforward. Therefore, it was decided to calculate all the confidence intervals using the bootstrap (nonparametric) re-sampling method.

11.4.3 Estimation of $F_T(t)$ using the two-stage method

The steps of the analysis are given below.

1. As was done in the approximate and the analytical method, for each sampled unit, the degradation model (11.1) was fitted to the sample paths and the estimates of the model parameters were obtained using the least squares estimation method (the estimated values $\hat{\beta}_i$ are shown in Table 11.1).

2. Next, in order to use the two-stage estimation method, one would have to find an appropriate transformation $\hat{\varphi}_i = H(\hat{\beta}_i)$, with $\hat{\varphi}_i$ approximately normally distributed with asymptotic mean μ_φ and asymptotic variance-covariance σ_φ^2 . However, as was mentioned before, the analysis of probability plots constructed to the $\hat{\beta}_i$'s indicated that either a lognormal or a Weibull distribution could be used to fit those values, in particular, a *lognormal* (2.46192; 0.644248) or a *Weibull* (1.976719; 15.54349) (see results of the analytical method). Therefore, it was possible to move on to step 3, using these two distributions.
3. $N = 100,000$ simulated realizations β^* of β were generated from each one of the two distributions mentioned in step 2.
4. For each distribution, the corresponding N simulated failure times t^* were calculated by substituting each β^* into $D_f = \beta t$.
5. F_T was estimated from the simulated empirical distribution

$$\hat{F}_T(t) = [\text{Number of simulated first crossing times } t^* \leq t] / N$$

for any desired values of t .

Bootstrap 95% confidence intervals were calculated. The results are shown in Table 11.2.

11.4.4 Estimation of $F_T(t)$ using the numerical method

The numerical method was not applied to this problem since there were evidences against the basic assumption regarding normality of the random parameter. de Toledo [31] showed that the results of the numerical method are strongly affected by the violation of that assumption.

11.4.5 Comparison of the results generated by the methods based on “classical” inference

Some observations from Table 11.2 and Figures 11.4–11.6 are the following:

1. The point estimates obtained by the approximate and the analytical methods are very similar. This result was already expected since there is a relationship between the random parameter distribution (F_β) and the pseudo failure time distribution (F_T).
2. The precision of the methods may be evaluated by the confidence interval widths. These values are essentially the same for the central measures (MTTF and $t_{0.50}$) for both distributions. Some differences can be detected for $t_{0.10}$ and $R(300,000)$. For $t_{0.10}$, it seems that the two-stage method is slightly better than the other two for the Weibull distribution. On the other hand, the approximate method is the best one in the lognormal case. In terms of the $R(300,000)$, the approximate is the worst method for the two distributions considered.

11.4.6 Bayesian inference

In Section 11.4.1, pseudo lifetimes (obtained by fitting lines to each degradation curve and calculating the times when the fitted lines reach the failure threshold) were used

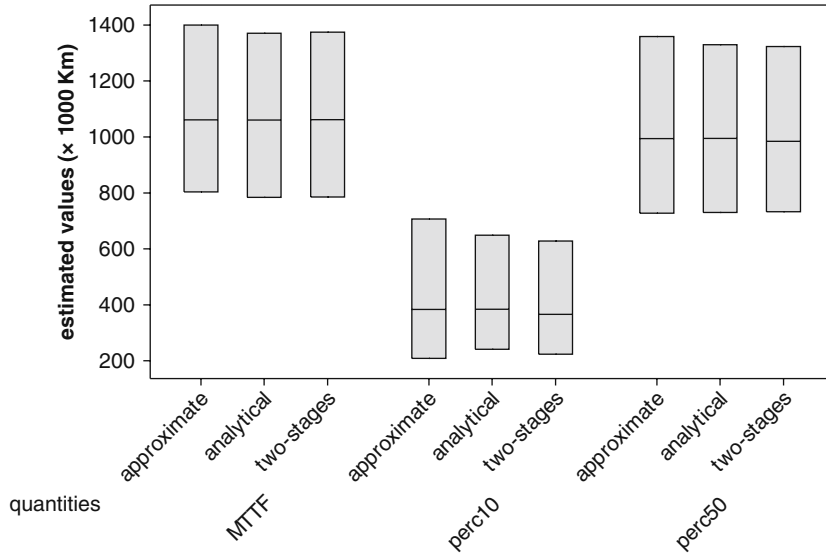


Figure 11.4. Point estimates (MTTF , $t_{0.10}$, $t_{0.50}$) and confidence intervals obtained by each method of degradation data analysis (“classical” inference). Weibull distribution

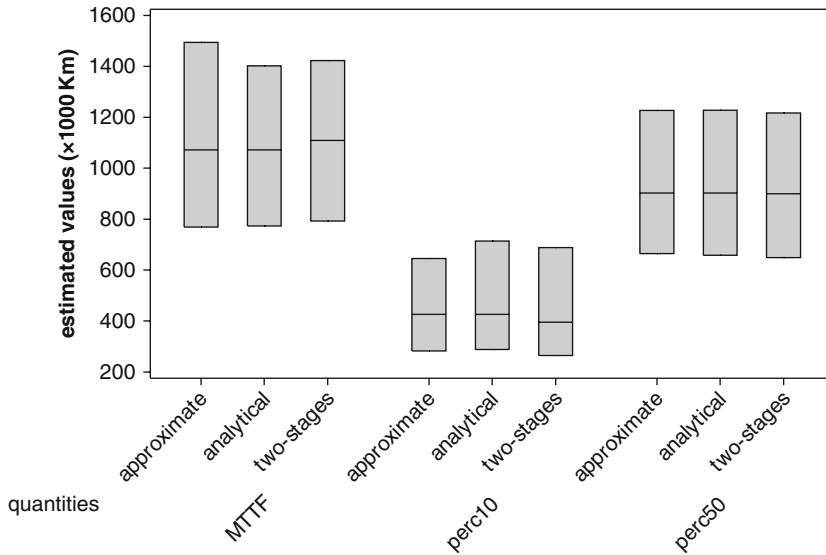


Figure 11.5. Point estimates (MTTF , $t_{0.10}$, $t_{0.50}$) and confidence intervals obtained by each method of degradation data analysis (“classical” inference). Lognormal distribution

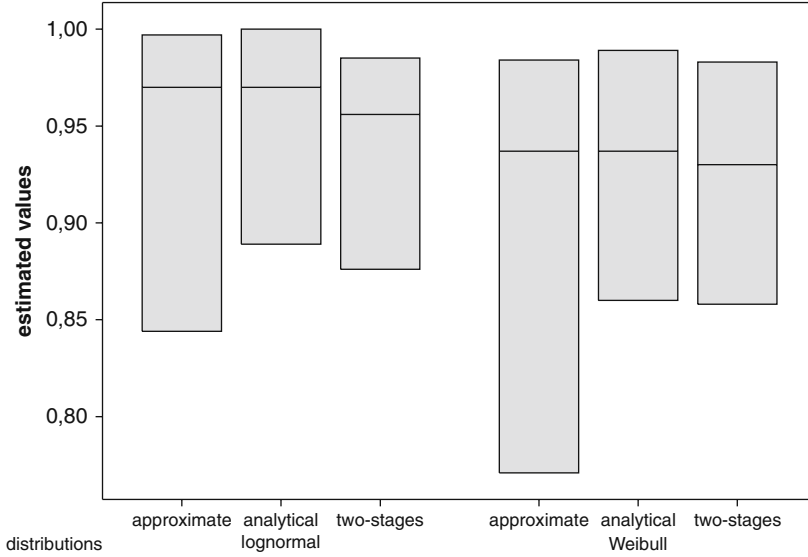


Figure 11.6. Point estimates of $R(300,000)$ and confidence intervals obtained by each method of degradation data analysis (“classical” inference). Weibull and lognormal distributions

to identify an appropriate lifetime distribution. The analysis showed that the pseudo lifetimes for the wheel degradation data are well described either by a Weibull or a lognormal distribution. In addition, analytical expressions established the relationships between the reciprocal slopes β_i and lifetimes distributions.

Consequently, in order to analyze the wheel degradation data, the following two models were considered:

- Model 1: $y_{ij} = D_i(t_{ij}) + \varepsilon_{ij} = (1/\beta_i)t_{ij}$; $\beta_i \sim Weibull(\delta_\beta, \lambda_\beta)$, therefore $y_{ij}|\beta = \beta_i \sim N((1/\beta_i)t_{ij}, \sigma_\varepsilon^2)$ for $i = 1, \dots, 14$ and $j = 1, \dots, 12$. In this case, the following flat priors were used:
 - $\delta_\beta \sim Gamma(0.01; 0.01)$,
 - $\lambda_\beta \sim Gamma(0.01; 0.01)$, and
 - $\sigma_\varepsilon^2 \sim InverseGamma(0.01; 0.01)$.

Gamma priors were chosen for δ_β and λ_β because they are positive quantities. The measurement error variance σ_ε^2 is also a positive quantity, but there has been a tradition in the Bayesian literature to use an inverse gamma prior for this parameter [19]. Consequently, the prior of the reciprocal of σ_ε^2 is also a gamma distribution. The parameters δ_β , λ_β , and σ_ε^2 are assumed independent.

- Model 2: $y_{ij} = D_i(t_{ij}) + \varepsilon_{ij} = (1/\beta_i)t_{ij}$; $\beta_i \sim lognormal(\mu_\beta, \sigma_\beta^2)$ and the following flat priors:
 - $\mu_\beta \sim Normal(0; 100)$,
 - $\sigma_\beta^2 \sim InverseGamma(0.02; 0.02)$, and
 - $\sigma_\varepsilon^2 \sim InverseGamma(0.01; 0.01)$.

These parameters are also assumed independent.

The posterior of δ_β (shape parameter), λ_β (scale parameter), σ_ε^2 , β_i ($i = 1, \dots, 14$) (Weibull case) and of μ_{beta} (location), σ_β^2 (scale), σ_ε^2 , β_i (lognormal case) were ob-

tained by MCMC. A sample of size 102,000 was considered with a burn-in period of 2,000 draws and no thinning. The burn-in period was achieved by discarding the first 2,000 samples and because there was no thinning, the next 100,000 samples were kept. Convergence was assessed by graphical methods (Gamerman and Lopes, 2006). The results for the Weibull and lognormal cases are shown in Tables 11.3 and 11.4, respectively.

Note that for the Weibull case and the quadratic loss function, the reliability of the wheels at 300,000 km is 0.92 (95% HPD region is [0.82;0.99]). In addition, 10% of the wheels will need replacement by 382.80×10^3 km of usage (95% HPD: [151.86×10^3 km; 611.34×10^3 km]). The Bayesian estimates for the other quantities are given in Table 11.3 along with 95% HPD regions and selected quantiles of the posterior distribution.

Table 11.3. Bayesian estimates of the quantities of interest and 95% HPD regions (prior: Weibull distribution)

Quantities	Mean	Median	Standard deviation	$Q^{(1)}$ 2.5%	$HPD^{(3)}$ LB	$Q^{(2)}$ 97.5%	$HPD^{(4)}$ UB
λ_β	0.01	0.01	0.00	0.00	0.04	0.03	
δ_β	1.95	1.93	0.41	1.22	1.18	2.80	2.75
σ_ε	0.99	0.99	0.06	0.88	0.88	1.11	1.11
$MTTF^{(5)}$	1097.00	1083.00	172.13	800.84	772.21	1473.13	1433.46
$t_{0.10}^{(5)}$	382.80	378.90	118.79	163.65	151.86	624.66	611.34
$t_{0.50}^{(5)}$	1011.00	1006.00	170.75	688.85	679.88	1361.00	1349.66
R(300,000)	0.92	0.93	0.05	0.80	0.82	0.98	0.99

(1) 2.5% quantile for the posterior distribution; (2) 97.5% quantile of the posterior distribution; (3) HPD region lower bound; (4) HPD region upper bound; (5) values should be multiplied by 10^3 km.

Table 11.4. Bayesian estimates of the quantities of interest and 95% HPD regions (prior: lognormal distribution)

Quantities	Mean	Median	Standard deviation	$Q^{(1)}$ 2.5%	$HPD^{(3)}$ LB	$Q^{(2)}$ 97.5%	$HPD^{(4)}$ UB
μ_θ	2.46	2.46	0.18	2.10	2.08	2.81	2.81
σ_θ	0.65	0.62	0.14	0.44	0.41	0.98	0.92
σ_ε	0.99	0.99	0.06	0.88	0.88	1.11	1.11
$MTTF^{(5)}$	1147.00	1101.00	264.13	796.86	741.67	1776.64	1647.80
$t_{0.10}^{(5)}$	405.60	405.90	94.01	220.39	216.15	589.28	584.85
$t_{0.50}^{(5)}$	916.10	901.30	163.98	636.39	606.56	1264.10	1242.41
R(300,000)	0.95	0.96	0.05	0.82	0.86	0.99	1.00

(1) 2.5% quantile for the posterior distribution; (2) 97.5% quantile of the posterior distribution; (3) HPD region lower bound; (4) HPD region upper bound; (5) values should be multiplied by 10^3 km.

In the lognormal case, the reliability of the wheels at 300,000 km is 0.95 (95% HPD region: $[0.86; 1.00]$). In addition, 10% of the wheels will need replacement by 405.60×10^3 km of usage (95% HPD: $[216.15 \times 10^3 \text{ km}; 584.85 \times 10^3 \text{ km}]$). The Bayesian estimates for the other quantities are given in Table 11.4 along with 95% HPD regions and selected quantiles of the posterior distribution. The DIC value (deviance information criterion; [28]) for the Weibull and the lognormal models were 6309.27 and 6309.50, respectively, indicating a similar performance of the two selected models.

Figures 11.7 and 11.8 show histograms of the posterior distributions of the MTTF (mean time to failure or mean covered distance); $R(300,000)$ and $t_{0.10}$ for the Weibull and lognormal models, respectively. Note that $R(300,000)$ has an asymmetrical distribution in both cases.

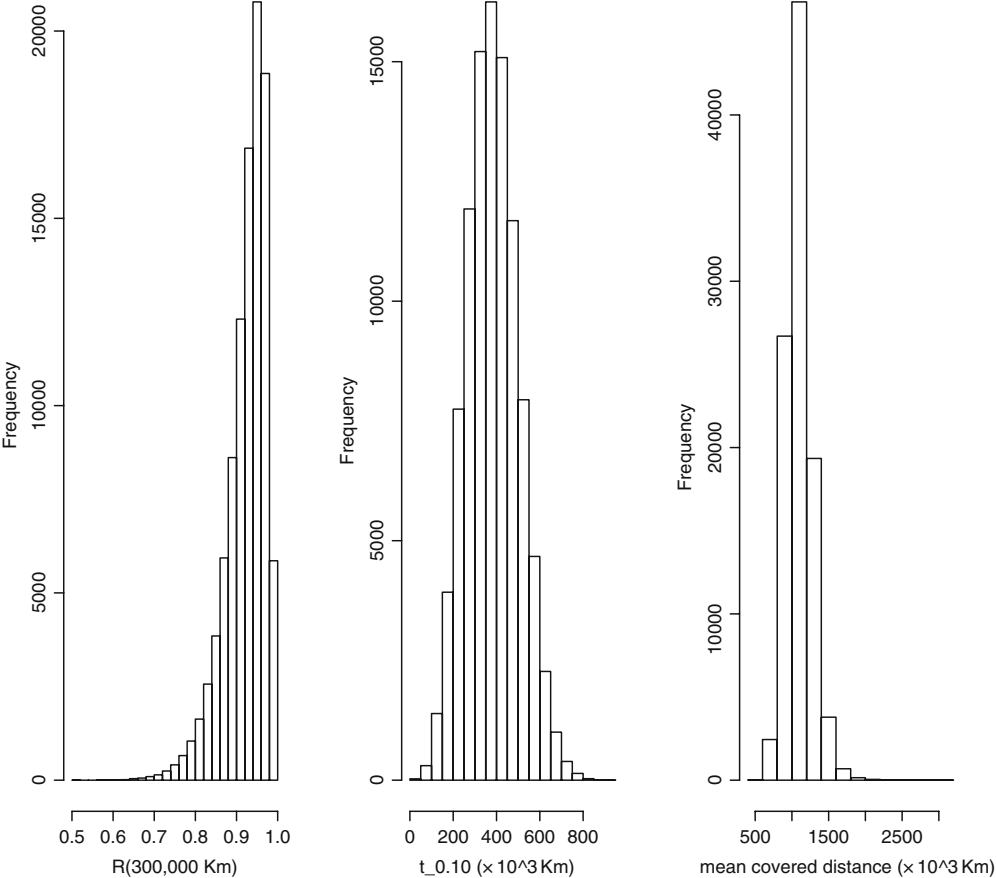


Figure 11.7. Histograms of the posterior distributions for $R(300,000)$, $t_{0.10}$, and the mean covered distance, respectively. Weibull model

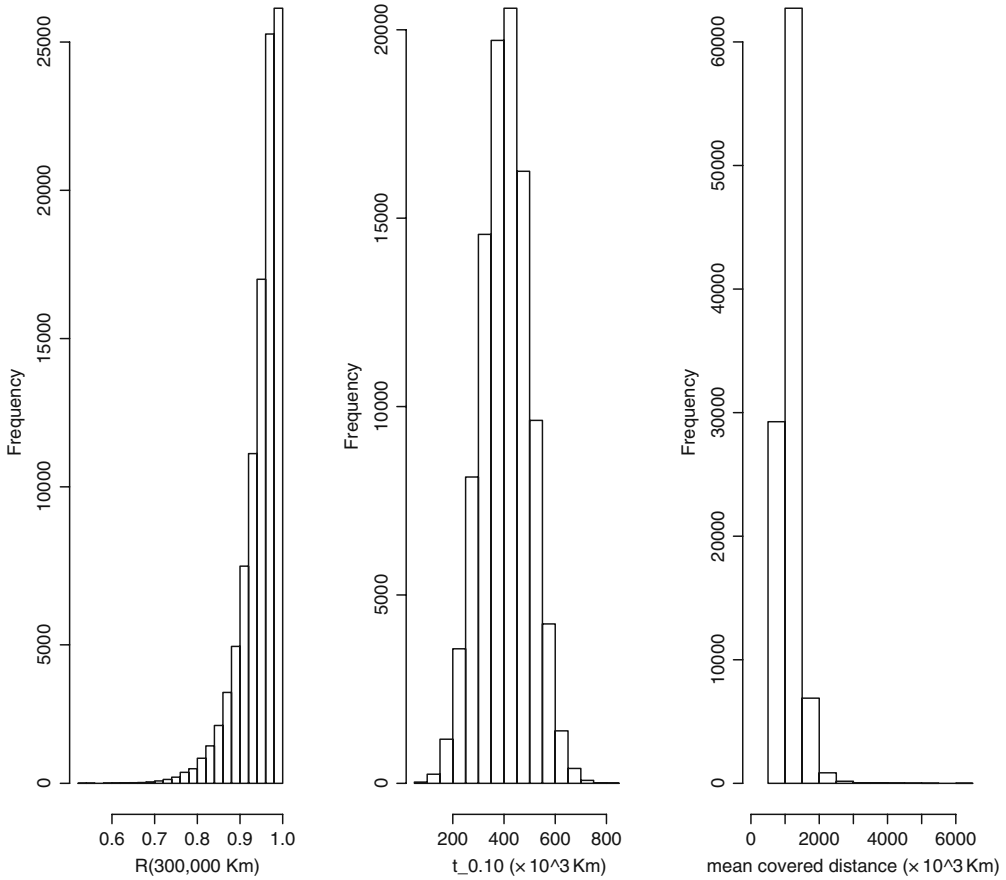


Figure 11.8. Histograms of the posterior distributions for $R(300,000)$, $t_{0.10}$ and, the mean covered distance, respectively. Lognormal model

11.5 Conclusions

In this chapter, five methods of degradation data analysis were presented. Four of them are based on the so-called classical inference. The numerical method was not applied to the real data set since a basic model assumption was not valid in that situation. The point estimates obtained by each one of the “classical” methods were very similar. In particular, due to the relationship between the random parameter distribution and the failure time distribution, it was found that the approximate and the analytical methods lead to the same results. For more complicated models (nonlinear, more than one random parameter, or even a mixed parameter model), the application of those methods might be difficult and may lead to different results. In these cases, researches

will have to use the numerical method, assuming that the vector of random parameters has a multivariate normal distribution.

On the other hand, Bayesian approach seems to be a reasonable choice especially if one needs to handle more complicated degradation models. Because reliability and lifetime distribution quantiles are functions of the model parameters, posteriors for these quantities are easily obtained from draws of the model parameter posteriors; for each such draw, simply evaluate the quantity of interest to obtain draws from that quantities' posterior.

One should be careful to compare the results of Bayesian and "classical" approaches since the concepts behind them are quite different. The former leads to a posterior distribution of the (random) quantity of interest while the latter produces a point estimate (of a fixed quantity). In "classical" approaches, confidence intervals are constructed while credible intervals are obtained in the Bayesian methods. But in practical situations like the one described in this chapter, it is necessary to report some kind of "point estimate" in order to support future technical decisions. In that case, it is fair to say that by using flat priors and the quadratic loss function, Bayesian and classical approaches lead roughly to the same results.

References

1. Casella, G. and George, E.I.: Explaining the Gibbs sampler. *The American Statistician*, **46**, no.3, 167–174 (1992).
2. Chiao, C.H. and Hamada, M.: Using degradation data from an experiment to achieve robust reliability for light emitting diodes. *Quality and Reliability Engineering International*, **12**, 89–94 (1996).
3. Chiao, C.H. and Hamada, M.: Experiments with degradation data for improving reliability and for achieving robust reliability. *Quality and Reliability Engineering International*, **17**, 333–344 (2001).
4. Chib, S. and Greenberg, E.: Understanding the Metropolis-Hasting algorithm. *The American Statistician*, **49**, no.4, 327–335 (1995).
5. Crk, V.: Reliability assessment from degradation data. In: *Proceedings of the Annual Reliability and Maintainability Symposium*, pp.155–161 (2000).
6. Doksum, K.A.: Degradation rate models for failure time and survival data. *CWI Quarterly*, **4**, 195–203 (1991).
7. Efron, B.: Bootstrap confidence intervals for a class of parametric problems. *Biometrika*, **72**, 45–58 (1985).
8. Gamerman, D. and Lopes, H.F.: Markov chain Monte Carlo. *Stochastic Simulation for Bayesian Inference*, 2nd edition. Texts in Statistical Science, Chapman & Hall, London (2006).
9. Gelfand, A.E. and Smith, A.F.M.: Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**, 398–409 (1990).
10. Geman, S. and Geman, D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721–741 (1984).
11. Hamada, M.: Using degradation data to assess reliability. *Quality Engineering*, **17**, 615–620 (2005).
12. Jiang, M. and Zhang, Y.: Dynamic modeling of degradation data. In: *Proceedings of the Annual Reliability and Maintainability Symposium*, pp. 607–611 (2002).

13. Lindstrom, M.J. and Bates, D.M.: Nonlinear mixed effects models for repeated measures data. *Biometrics*, **46**, 673–687 (1990).
14. Lu, C.J. and Meeker, W.Q.: Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, **35**, 161–174 (1993).
15. Lu, C. ., Meeker, W.Q., and Escobar, L.A.: Using degradation measures to estimate a time-to-failure distribution. *Statistica Sinica*, **6**, 531–546 (1996).
16. Lu, C.J., Park, J., and Yang, Q.: Statistical inference of a time-to-failure distribution derived from linear degradation data. *Technometrics*, **39**, 391–400 (1997).
17. Marseguerra, M., Zio, E., and Cipollone, M.: Designing optimal degradation tests via multi-objective genetic algorithms. *Reliability Engineering & System Safety*, **79**, 287–294 (2003).
18. Meeker, W.Q. and Escobar, L.A.: *Statistical Methods for Reliability Data*, John Wiley & Sons, New York (1998).
19. Migon, H.S. and Gamerman, D.: *Statistical Inference: An Integrated Approach*, Arnold, London (1999).
20. Minitab, Statistical Package, release 15.
21. Mood, Alexander A., Graybill, Franklin A., and Boes, Duane C.: *Introduction to the Theory of Statistics*, 3rd edition, McGraw Hill, London (1974).
22. Nelson, W.: Analysis of performance degradation data from accelerated tests. *IEEE Transactions on Reliability*, **30**, 149–155 (1981).
23. Nelson, W.: *Accelerated Testing: Statistical Methods, Test Plans, and Data Analysis*, John Wiley & Sons, New York (1990).
24. Pinheiro, J.C. and Bates D.C.: Approximations to the log-likelihood function in The nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, **1**, 12–35 (1995).
25. Pinheiro, J.C. and Bates D.C.: *Mixed-Effects Models in S and S-PLUS*, Springer-Verlag, New York (2000).
26. Plummer, M., Best, N., Cowles, K., and Vines, K.: Output analysis and diagnostics for MCMC. The Coda Package. Software R. Available at <http://www.fis.iarc.fr/coda/> (2005).
27. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org> (2006).
28. Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and van der Linde, A.: Bayesian measures of model complexity and fit (with discussion and rejoinder). *Journal of the Royal Statistical Society, Series B*, **64**, 583–639 (2002).
29. Spiegelhalter, D.J, Thomas, A., Best, N.: WinBUGS v.1.4 User Manual available at <http://www.mrc-bsu.cam.ac.uk/bugs/> (2000)
30. Su, C., Lu, J.C., Chen,D., and Hughes-Oliver,J.M.A.: Random coefficient degradation model with random sample size. *Lifetime Data Analysis*, **5**, 173–183 (1999).
31. de Toledo, M.L.G.: Degradation tests: A comparative study of data analysis methods. *Masters Degree Thesis*. Department of Statistics. Federal University of Minas Gerais, Brazil (2007).
32. Tsang, L.C. and Chang, D.S.: Reliability prediction using nondestructive accelerated-degradation data: Case study on power supplies. *IEEE Transactions on Reliability*, **44**, 562–566 (1995).
33. Tseng, S.T., Hamada, M., and Chiao, C.H.: Using degradation data to improve fluorescent lamp reliability. *Journal of Quality Technology*, **27**, 363–369 (1995).
34. Whitmore, G.A. and Shenkelberg, F.: Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. *Lifetime Data Analysis*, **3**, 27–45 (1997).
35. Wu, S.J. and Chang, C.T.: Optimal design of degradation tests in presence of cost constraint. *Reliability Engineering & System Safety*, **76**, 109–115 (2002).

36. Wu, S.J. and Shao, J.: Reliability analysis using the least squares method in nonlinear mixed-effect degradation models. *Statistica Sinica*, **9**, 855–877 (1999).
37. Wu, S.J. and Tsai, T.R.: Estimation of time-to-failure distribution derived from a degradation model using fuzzy clustering. *Quality and Reliability Engineering International*, **16**, 261–267 (2000).
38. Yacout, A.M., Salvatores, S., and Orechwa, Y.: Degradation analysis estimates of the time-to-failure distribution of irradiated fuel elements. *Nuclear Technology*, **113**, 177–189 (1996).
39. Yu, H.F. and Tseng, S.T.: Designing a Degradation Experiment. *Naval Research Logistics*, **46**, 689–706 (1997).
40. Yu, H.F. and Chiao, C.H.: An Optimal Designed Degradation Experiment for Reliability Improvement. *IEEE Transactions on Reliability*, **51**, 427–433 (2002).
41. Yu, H.F. and Tseng, S.T.: Designing a degradation experiment with a reciprocal Weibull degradation rate. *Quality Technology and Quantitative Management*, **1**, 47–63 (2004).

Optimal Prophylaxis Policy Under Non-monotone Degradation

S.S. Rasova and B.P. Harlamov

Institute of Problems of Mechanical Engineering (Russian Academy of Science)
St. Petersburg, Russia

Abstract: A stationary server system with observable degradation is considered. An optimization problem for choice of a time to begin a prophylactic repair of the system is being investigated. A mathematical problem is to choose a Markov time on a random process which is optimal with respect to some criterion. A necessary condition for the time to be optimal is such that it determines a unique solution for monotone random processes. For non-monotone processes this necessary condition determines a set of Markov times containing a time of the first fulfilment of the condition (trivial solution), if this set is not empty. The question arises: Is the trivial solution optimal? We show that it depends on parameters of the process and mainly on difference between the hazard rate and the rate of useful output of the system. For Markov processes the following alternative is true: either the trivial time is optimal or there exist no optimal times.

Keywords and phrases: Degradation, non-monotone, hazard rate, insurance, prophylaxis, optimal Markov time

12.1 Setting of the Problem

In this chapter we continue investigation which had begun in work [1]. We consider a server system with one channel. Its action consists of work and repair periods. The repair periods are of two types: a repair after a failure and a prophylactic repair. The prophylactic repair begins at the instant defined by an operator of the system. A rule of these prophylactic switches is an object of optimization. A criterion of optimization may be either a coefficient of readiness or some other functional depending on different evaluation of loss due to failure and that due to prophylactic stopping.

From a point of view of reliability this server system is being described with a regenerative process, generated by a sequence of i.i.d. sets of five $(X_n, \zeta_n, \tau_n, M_n, R_n)_{n=1}^{\infty}$. Elements of every set of five are mutually dependent. They are

- $X_n = (X_n(t))$ ($t \geq 0$), an one-dimensional random process with trajectories from the Skorokhod space \mathcal{D} ; actually the process is considered on a random interval

$[0, \zeta_n \wedge \tau_n)$ and is interpreted as a degradation of the system in the n th cycle of regeneration;

- ζ_n , a random time of a failure, connected with the process X_n with the help of a hazard rate $h(X_n)$, where $h(x)$ ($x \in \mathbb{R}$) is a non-decreasing positive function; hence

$$E(\zeta_1 > t | X_1) = \exp \left(- \int_0^t h(X_1(s)) ds \right),$$

we assume that $E\zeta_1 < \infty$ (recently this model is rather popular among specialists on reliability theory; see, for example, [5], where different models of hazard rate dependence on degradation are considered);

- τ_n , a time of beginning of the prophylactic repair, which is a Markov time with respect to the process X_n and the natural filtration (\mathcal{F}_t) on \mathcal{D} .
- M_n , a random duration of a repair period.
- R_n , a random value of loss due to repair and its reason.

Let $T_n = \zeta_n \wedge \tau_n + M_n$ be a duration of the n th cycle of regeneration. Hence (T_n) is a sequence of i.i.d. positive random values. Let $\sigma_n = \sum_{k=1}^n T_k$ ($n \geq 1$, $\sigma_0 = 0$) and $N_t = \sum_{n=1}^{\infty} I(\sigma_n \leq t)$ be the corresponding renewal process. A regeneration cycle begins from a work period, while degeneration varies and ends with a repair period. We assume the system has a loss during one regeneration cycle

$$f_n = - \int_0^{\zeta_n \wedge \tau_n} A(X_n(t)) dt + R_n,$$

where $A(\cdot)$ is a non-decreasing positive function; we interpret $A(X_n(t)) dt$ as a useful output during an infinitesimal interval dt .

Let us remark that M_n , R_n are not measurable with respect to the process $X_n \equiv (X_n(t))$. In what follows we use only their conditional expectations:

$$m \equiv E(M_1 | \zeta_1 \leq \tau_1), \quad n \equiv E(M_1 | \zeta_1 > \tau_1),$$

$$r \equiv E(R_1 | \zeta_1 \leq \tau_1), \quad s \equiv E(R_1 | \zeta_1 > \tau_1).$$

We suppose that $m > n$ and $r > s$.

The average loss of the system during a long period of action is as follows:

$$L \equiv \text{a.s.} \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{k=1}^{N_t} f_k \circ \theta_{\sigma_{k-1}} = \frac{Ef_1}{ET_1},$$

where θ_t is an operator of shift on the set of generalized trajectories. We have

$$Ef_1 = -E \int_0^{\zeta_1 \wedge \tau_1} A(X_1(t)) dt + rP(\zeta_1 \leq \tau_1) + sP(\zeta_1 > \tau_1),$$

$$ET_1 = E(\zeta_1 \wedge \tau_1) + mP(\zeta_1 \leq \tau_1) + nP(\zeta_1 > \tau_1).$$

In what follows we will drop index 1 in denotations of random processes and values of the first regeneration cycle if it does not lead to ambiguity, and use denotation X_t instead of $X(t)$, when it seems convenient.

Using positivity of ζ , we can obtain

$$Ef_1 = -W_\tau + r - (r - s)U_\tau, \quad ET_1 = V_\tau + m - (m - n)U_\tau,$$

where

$$\begin{aligned} U_\tau &= E \exp \left(- \int_0^\tau h(X_s) ds \right), \\ V_\tau &= E \int_0^\tau \exp \left(- \int_0^t h(X_s) ds \right) dt, \\ W_\tau &= E \int_0^\tau A(X_t) \exp \left(- \int_0^t h(X_s) ds \right) dt. \end{aligned}$$

Hence $L \equiv L_\tau = \frac{Y_\tau}{Z_\tau}$ where

$$Y_\tau = -W_\tau + r - (r - s)U_\tau, \quad Z_\tau = V_\tau + m - (m - n)U_\tau. \quad (12.1)$$

12.2 Optimization Problem

We will investigate a classical optimization problem to find τ_* such that

$$L_{\tau_*} = \min_{\tau \in \mathcal{T}} L_\tau,$$

where \mathcal{T} is the set of Markov times with respect to filtration (\mathcal{F}_t) .

Let η be some positive-bounded measurable function on \mathcal{D} . Let us consider for any $\varepsilon > 0$ the difference $L_{\tau-\varepsilon\eta} - L_\tau = (Y_{\tau-\varepsilon\eta}Z_\tau - Y_\tau Z_{\tau-\varepsilon\eta})/(Z_{\tau-\varepsilon\eta}Z_\tau)$. It is clear that if τ is a point of local minimum of the functional L , then $\liminf_{\varepsilon \rightarrow 0} (L_{\tau-\varepsilon\eta} - L_\tau)/\varepsilon \geq 0$. For $\varepsilon > 0$ we have

$$Y_{\tau-\varepsilon\eta}Z_\tau - Y_\tau Z_{\tau-\varepsilon\eta} = (Y_{\tau-\varepsilon\eta} - Y_\tau)Z_\tau + Y_\tau(Z_\tau - Z_{\tau-\varepsilon\eta})$$

and

$$\begin{aligned} Y_\tau - Y_{\tau-\varepsilon\eta} &= E \left(- \int_{\tau-\varepsilon\eta}^\tau A(X_t) \exp \left(- \int_0^t h(X_s) ds \right) dt \right. \\ &\quad \left. - (r - s) \left(\exp \left(- \int_0^\tau h(X_s) ds \right) - \exp \left(- \int_0^{\tau-\varepsilon\eta} h(X_s) ds \right) \right) \right). \end{aligned}$$

From here we have

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (Y_\tau - Y_{\tau-\varepsilon\eta}) &= E \left(-A(X_{\tau-0}) \exp \left(- \int_0^\tau h(X_s) ds \right) \eta \right. \\ &\quad \left. + (r - s) \exp \left(- \int_0^\tau h(X_s) ds \right) h(X_{\tau-0}) \eta \right). \end{aligned}$$

Similarly

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (Z_\tau - Z_{\tau-\varepsilon\eta}) = E \left(\exp \left(- \int_0^\tau h(X_s) ds \right) \eta \right. \\ \left. + (m-n) \exp \left(- \int_0^\tau h(X_s) ds \right) h(X_{\tau-0}) \eta \right)$$

and

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (Y_{\tau-\varepsilon\eta} Z_\tau - Y_\tau Z_{\tau-\varepsilon\eta}) \\ = E \left(\left(A(X_{\tau-0}) \exp \left(- \int_0^\tau h(X_s) ds \right) \eta \right. \right. \\ \left. \left. - (r-s) \exp \left(- \int_0^\tau h(X_s) ds \right) h(X_{\tau-0}) \eta \right) Z_\tau + \left(\exp \left(- \int_0^\tau h(X_s) ds \right) \eta \right. \right. \\ \left. \left. + (m-n) \exp \left(- \int_0^\tau h(X_s) ds \right) h(X_{\tau-0}) \eta \right) Y_\tau \right).$$

Since η is arbitrary we conclude that in a point of local minimum

$$(A(X_{\tau-0}) - (r-s)h(X_{\tau-0}))Z_\tau + (1 + (m-n)h(X_{\tau-0}))Y_\tau \geq 0 \quad (12.2)$$

is almost sure. On the other hand, at the local minimum point we must have

$$\liminf_{\varepsilon \rightarrow 0} (L_{\tau+\varepsilon\eta} - L_\tau) / \varepsilon \geq 0$$

and from here

$$(A(X_{\tau+0}) - (r-s)h(X_{\tau+0}))Z_\tau + (1 + (m-n)h(X_{\tau+0}))Y_\tau \leq 0.$$

Hence for processes continuous at the point τ on the set $\{\tau < \infty\}$ we obtain the equation a.s.

$$\frac{-A(X_\tau) + (r-s)h(X_\tau)}{1 + (m-n)h(X_\tau)} = \frac{Y_\tau}{Z_\tau} \quad (\tau < \infty). \quad (12.3)$$

Moreover, using inequality (12.2) one can derive that a local minimum exists any time when the left-hand side of (12.3) reaches the level, represented by the right-hand side of this equation, with up-crossing this level (crossing from below).

For a non-decreasing process X a time of the first fulfilment of this condition (trivial case) is unique. In general case let \mathcal{T}^* be the set of all Markov times τ such that $P(\tau < \infty) > 0$ and (12.3) holds. Evidently, if $\mathcal{T}^* \neq \emptyset$ it contains the trivial solution of this equation. The questions arise: Is there any trivial solution? Is it optimal, if any? How to find this solution?

In what follows we will consider continuous processes. Thus we will investigate solutions of (12.3).

12.3 Trajectories of Choice

Let us denote $P_x(A)$ a conditional probability of event A given $X_0 = x$. For a process of general type this conception requires a strict substantiation. In our case (Markov

process and regenerative process) there is no need in it, because here $P_x(X_0 = x) = 1$, and a consistent family of probability measures is determined in advance. Integrals with respect to a measure P_x will be denoted with argument x as follows: E_x , $U_\tau(x)$, $V_\tau(x)$, $W_\tau(x)$ and so on.

Let $\sigma_\Delta = \inf\{t \geq 0 : X_t \notin \Delta\}$, where $\Delta \subset \mathbb{R}$. We will use the following property of first exit times: if $\Delta_1 \subset \Delta_2$, then on the set $\sigma_{\Delta_1} < \infty$

$$\sigma_{\Delta_2} = \sigma_{\Delta_1} + \sigma_{\Delta_2} \circ \theta_{\sigma_{\Delta_1}},$$

where θ_t is an operator of shift on \mathcal{D} . For semi-infinite intervals we introduce additional denotations: $\mu_b = \sigma_{(-\infty, b)}$ and $\nu_b = \sigma_{(b, \infty)}$. Further we will use denotation $\tau_1 \dot{+} \tau_2 = \tau_1 + \tau_2 \circ \theta_{\tau_1}$, where $\tau_1, \tau_2 \in \mathcal{T}$, $\tau_1 < \infty$; and on the set $\{\tau_1 = \infty\}$, we define $\tau_1 \dot{+} \tau_2 = \infty$. The operation $\dot{+}$ is not commutative, but it is associative [2].

A Markov time τ is said to be localizing if the meaning of the process at this time is determined on the set $\{\tau < \infty\}$, i.e., all trajectories have identical values at this time if it is finite. Both the time μ_b for a continuous process with measure P_x with $x < b$ and the time ν_b for such a process with $x > b$ are examples of localizing Markov times. It is clear that every solution of (12.3) is a localizing Markov time.

Let us denote $B(X_\tau)$ the left-hand part of (12.3). Using representation $U_\tau(x) = S_\tau(x) + 1$, where

$$S_\tau(x) = E_x \int_0^\tau \exp\left(-\int_0^t h(X_s) ds\right) (-h(X_t)) dt,$$

we can rewrite (12.3) as follows

$$\begin{aligned} 0 &= B(X_\tau)Z_\tau(x) - Y_\tau(x) \\ &= B(X_\tau)(V_\tau(x) + n - (m - n)S_\tau(x)) - (-W_\tau(x) + s - (r - s)S_\tau(x)) \\ &= B(X_\tau)n - s + B(X_\tau)V_\tau(x) + W_\tau(x) - (m - n)B(X_\tau)S_\tau(x) + (r - s)S_\tau(x). \end{aligned}$$

From here

$$\begin{aligned} &s - B(X_\tau)n \\ &= E_x \int_0^\tau \exp\left(-\int_0^t h(X_s) ds\right) [B(X_\tau) + A(X_t) + (m - n)B(X_\tau)h(X_t) \\ &\quad - (r - s)h(X_t)] dt. \end{aligned}$$

The expression in square brackets can be represented as

$$(1 + (m - n)h(X_t))(B(X_\tau) - B(X_t)),$$

which is a key result for the following theorem. Denote

$$H_\tau(a|x) = E_x \int_0^\tau \exp\left(-\int_0^t h(X_s) ds\right) (1 + (m - n)h(X_t))(B(a) - B(X_t)) dt.$$

Now we can rewrite (12.3) in the form

$$s - B(X_\tau)n = H_\tau(X_\tau|x) \quad (\tau < \infty). \quad (12.4)$$

We see that a local minimum must be a localizing Markov time τ_a with $X_{\tau_a} = a$ for some a on the set $\{\tau_a < \infty\}$. Hence (12.4) can be represented like

$$s - nB(a) = H_{\tau_a}(a|x). \quad (12.5)$$

Let us remark that the left-hand side of this equation has sense any time when $P_x(\tau_a < \infty) > 0$. In this case it has a constant value on a set of a positive measure. On the other hand, the right-hand side of this equation as a degenerating random value has this meaning everywhere. Hence (12.3) can be formally extended on the whole space \mathcal{D} , i.e., we may consider the left-hand part as a determinate value. That is why (12.5) does not require restriction $\tau_a < \infty$.

Theorem 1. *Let (X_t) be a continuous process and*

(1) $h(0) = 0$, $A(0) = A_0 \geq 0$;

(2)

$$\frac{h'(x)}{1 + (m-n)h(x)} \geq \frac{A'(x)}{r-s + (m-n)A(x)} \geq 0 \quad (x > 0); \quad (12.6)$$

(3) $h(x) \rightarrow \infty$ as $x \rightarrow \infty$;

(4) $rn - sm \geq n\alpha$, where $\alpha = \limsup_{x \rightarrow \infty} A(x)/h(x)$.

Then there exists a solution of (12.3).

PROOF. From (2) we have

$$B'(x) = \frac{-A'(x)(1 + (m-n)h(x)) + h'(x)(r-s + (m-n)A(x))}{(1 + (m-n)h(x))^2} \geq 0 \quad (x > 0).$$

Hence the function $B(a)$ increases for $a > 0$. On the other hand, for $a_2 > a_1$

$$\begin{aligned} & H_{\mu_{a_2}}(a_2|x) - H_{\mu_{a_1}}(a_1|x) \\ &= E_x \int_{\mu_{a_1}}^{\mu_{a_2}} \left(- \int_0^t h(X_s) ds \right) (1 + (m-n)h(X_t))(B(a_2) - B(X_t)) dt \\ & \quad + (B(a_2) - B(a_1)) E_x \int_0^{\mu_{a_1}} \left(- \int_0^t h(X_s) ds \right) (1 + (m-n)h(X_t)) dt > 0, \end{aligned}$$

since in the first integral $X_t < a_2$. Hence the function $H_{\mu_a}(a|x)$ increases for $a > 0$. In the same manner we can prove that $H_{\nu_a}(a|x)$ is increasing on $a \in (-\infty, a)$. Moreover, from assumptions (1) and (2), it follows that

$$\begin{aligned} & \frac{1}{m-n} \int_0^x (\log(1 + (m-n)h(s)))' dt = \log(1 + (m-n)h(x)) \\ & \geq \frac{1}{m-n} \int_0^x (\log(r-s + (m-n)A(s)))' dt = \log \frac{r-s + (m-n)A(x)}{r-s + (m-n)A_0}. \end{aligned}$$

Hence $h(x) \geq (A(x) - A_0)/(r-s + (m-n)A_0)$ and there exists the limit, α , of the ratio $A(x)/h(x)$ and

$$\lim_{x \rightarrow \infty} B(x) = \frac{-\alpha + r - s}{m - n}.$$

Condition (4) implies that a graph of the right-hand side of (12.5) intersects the abscissa axis at some point b_0 or tends to this axis as $x \rightarrow \infty$ ($b_0 = \infty$). If $x < b_0$ then this graph

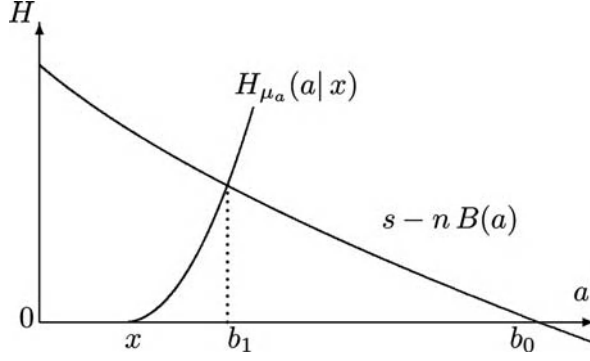


Figure 12.1. A trivial solution

has some point, b_1 , of intersection with a graph of the left-hand side (see Figure 12.1). This point determines the level of the first fulfilment of condition (12.3). A trajectory of the process intersects the level from below. So it is a trivial solution of (12.3).

If $x \geq b_0$ there is a point, c , of intersection of the graph $s - nB(a)$ ($a > 0$) with the graph $H_{\nu_a}(a|x)$ ($a \leq x$). Let $\tau_a = \nu_{c-\varepsilon} + \mu_a$ ($\varepsilon > 0$).

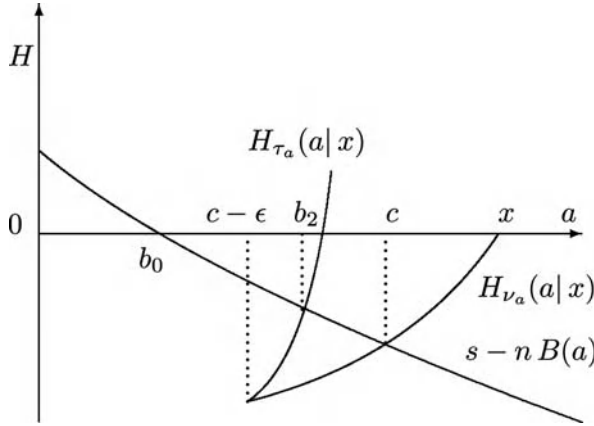


Figure 12.2. A trivial solution of the second kind

Consider a composition of graphs $H_{\nu_a}(a|x)$ $a \in [c-\varepsilon, x)$ and $H_{\tau_a}(a|x)$ $a \geq c-\varepsilon$, i.e., a union of these curves (these graphs have a common point with abscissa $c-\varepsilon$). Since the last function increases on this interval (one can prove it like that for $H_{\mu_a}(a|x)$) there exists $\varepsilon > 0$ such that the composition intersects the graph $s - nB(a)$ ($a > 0$) by its ascending link in some point b_2 . This second point of intersection determines a solution of (12.5) and at the time τ_{b_2} , the trajectory of the process intersects the level from below. In order to find this point, we have first to wait for the first exit from interval $(c-\varepsilon, \infty)$ and then from point $c-\varepsilon$ to wait for the first fulfilment of the condition (12.5). So it is a solution of (12.3). The point b_2 belongs to some neighbourhood of the point c . It can be more or less than c . It is clear that this solution is not unique. The point c , a limit of such solutions as $\varepsilon \rightarrow 0$, is said to be a trivial solution of the second kind (Figure 12.2) \square

In order to explain the procedure of looking for local minimums of L_τ , we define so-called trajectory of choice (more general than that in proof of Theorem 1).

Definition 1 For given start point x of a trajectory and a finite sequence of pairwise distinct levels $(a_k)_1^n$ ($n \geq 1$) (the last level can be $\pm\infty$) we define a sequence of iterated first exit times $(\tau_k)_1^n$, where

$$\tau_k = \begin{cases} \tau_{k-1} \dot{+} \mu_{a_k}, & a_k > a_{k-1}, \\ \tau_{k-1} \dot{+} \nu_{a_k}, & a_k < a_{k-1} \end{cases} \quad (k = 1, \dots, n)$$

($\tau_0 = 0$, $a_0 = x$). A trajectory of choice corresponding to the initial point and the sequence of levels is said to be the sequence $\Psi(a|x) = (\Psi_k(a|x))_{k=1}^n$ of functions

$$\Psi_k(a|x) = \begin{cases} H_{\tau_{k-1} \dot{+} \mu_a}(a|x), & a_{k-1} \leq a < a_k \text{ (ascending link)} \\ H_{\tau_{k-1} \dot{+} \nu_a}(a|x), & a_{k-1} \geq a > a_k \text{ (descending link)} \end{cases} \quad (k = 1, \dots, n)$$

called as links of the trajectory of choice.

Simple ascending trajectory of choice is said to be the function $H_{\mu_a}(a|x)$ ($a \geq x$).

Simple descending trajectory of choice is said to be the function $H_{\nu_a}(a|x)$ ($a \leq x$).

Due to properties $a < b \Rightarrow \mu_a \dot{+} \mu_b = \mu_b$ and $a < b \Rightarrow \nu_b \dot{+} \nu_a = \nu_a$, we can exclude the first one from every two sequential levels, determined ascending links, and the same for two sequential levels, determined descending links. Thus we obtain an alternating sequence of levels and the corresponding sequence of links.

Composition of trajectories of choice.

Let τ_b and τ be two Markov times. We have

$$\begin{aligned} & H_{\tau_b \dot{+} \tau}(a|x) \\ &= E_x \int_0^{\tau_b \dot{+} \tau} \exp \left(- \int_0^t h(X_s) ds \right) (1 + (m-n)h(X_t))(B(a) - B(X_t)) dt \\ &= H_{\tau_b}(a|x) + E_x \int_{\tau_b}^{\tau_b \dot{+} \tau} \exp \left(- \int_0^t h(X_s) ds \right) (1 + (m-n)h(X_t))(B(a) - B(X_t)) dt \\ &= H_{\tau_b}(a|x) + E_x \left(\exp \left(- \int_0^{\tau_b} h(X_s) ds \right) \right. \\ &\quad \times \left. \int_{\tau_b}^{\tau_b \dot{+} \tau} \exp \left(- \int_{\tau_b}^t h(X_s) ds \right) (1 + (m-n)h(X_t))(B(a) - B(X_t)) dt \right) \\ &= H_{\tau_b}(a|x) + E_x \left(\exp \left(- \int_0^{\tau_b} h(X_s) ds \right) \right. \\ &\quad \times \left. \left(\int_0^\tau \exp \left(- \int_0^t h(X_s) ds \right) (1 + (m-n)h(X_t))(B(a) - B(X_t)) dt \circ \theta_{\tau_b} \right); \tau_b < \infty \right). \end{aligned}$$

Hence, if $P_x(\tau_b < \infty) = 0$, then $H_{\tau_b \dot{+} \tau}(a|x) = H_{\tau_b}(a|x)$. Let τ_b be a regeneration time of the process (it means the homogeneous Markov property is fulfilled with respect to this time). Then

$$H_{\tau_b \dot{+} \tau}(a|x) = H_{\tau_b}(a|x) + E_x \left(\exp \left(- \int_0^{\tau_b} h(X_s) ds \right) H_{\tau}(a|X_{\tau_b}); \tau_b < \infty \right).$$

If, in addition, τ_b is a localizing Markov time for which $\tau_b < \infty \Rightarrow X_{\tau_b} = b$ and the following condition holds:

$$E_x \left(\exp \left(- \int_0^{\infty} h(X_s) ds \right); \tau_b = \infty \right) = 0, \quad (12.7)$$

then we obtain the formula

$$H_{\tau_b \dot{+} \tau}(a|x) = \begin{cases} H_{\tau_b}(a|x), & P_x(\tau_b < \infty) = 0, \\ H_{\tau_b}(a|x) + U_{\tau_b}(x)H_{\tau}(a|b), & P_x(\tau_b < \infty) > 0. \end{cases} \quad (12.8)$$

Comparison of links in Markov case.

For any alternating sequence of levels the trajectory of choice is a zigzag with curvilinear links. In order to apply this trajectory for us to obtain partial local minimums, we compare values of two sequential links intersecting the same level. Compare an ascending link and the succeeding descending one. Let $x < a < a_1$. We have

$$\begin{aligned} H_{\mu_a}(a|x) - H_{\mu_{a_1} \dot{+} \nu_a}(a|x) &= H_{\mu_a}(a|x) - H_{\mu_a \dot{+} \mu_{a_1} \dot{+} \nu_a}(a|x) = \\ &= -U_{\mu_a}(x)H_{\mu_{a_1} \dot{+} \nu_a}(a|a). \end{aligned}$$

Let us denote $\overline{G}_z(a) = H_{\mu_{a+z} \dot{+} \nu_a}(a|a)$. This function defines a sign of the difference. Now we compare a descending link and the succeeding ascending one. Let $x > a > a_1$. We have

$$H_{\nu_a}(a|x) - H_{\nu_{a_1} \dot{+} \mu_a}(a|x) = -U_{\nu_a}(x)H_{\nu_{a_1} \dot{+} \mu_a}(a|a).$$

The function $G_z(a) = H_{\nu_{a-z} \dot{+} \mu_a}(a|a)$ is another function with similar properties.

A good quality of these functions is their independence of preceding trajectory up to the time hitting the initial (for them) point a . We will show that these functions are independent of parameter z . Let $\alpha_z = \nu_{a-z} \dot{+} \mu_a$ and α_z^n be n th iteration of this Markov time ($\alpha_z^0 = 0$, $\alpha_z^{n+1} = \alpha_z^n \dot{+} \alpha_z$ ($n \geq 1$)).

Lemma 1. *For any $0 < z_1 < z_2$*

$$\alpha_{z_2} = \sum_{n=1}^{\infty} \alpha_{z_1}^n I_{A_n}, \quad (12.9)$$

where $A_n = \{\alpha_{z_1}^{n-1} \leq \nu_{a-z_2} < \alpha_{z_1}^n\} \in \mathcal{F}_{\alpha_{z_1}^n}$.

This property is proved in work [4]. The similar property is true for $\overline{\alpha}_z = \mu_{a+z} \dot{+} \nu_a$.

From the evident relation $\tau \dot{+} \sigma_{\Delta} \geq \sigma_{\Delta}$ ($\forall \tau \in \mathcal{T}$, $\Delta \subset \mathbb{R}$), it follows that function α_z decreases as $z \rightarrow 0$. The case when $\lim_{z \rightarrow 0} \alpha_z = \alpha_0 > 0$ and $\alpha_0^n \rightarrow \infty$ P_x -a.s. is called discrete. It is realized for piecewise increasing processes. For processes of diffusion type $\lim_{z \rightarrow 0} \alpha_z = 0$ P_x -a.s. (continuous case).

Theorem 2. For any $0 \leq z_1 < z_2$ such that $P_x(\alpha_{z_1} > 0) > 0$ the following formula is true:

$$G_{z_2}(a) = \frac{G_{z_1}(a)}{1 - E_b(\exp(-\int_0^{\alpha_{z_1}} h(X_s) ds); \alpha_{z_1} \leq \nu_{a-z_2})}.$$

Proof. Let us denote

$$f(t) = \exp\left(-\int_0^t h(X_s) ds\right) (1 + (m-n)h(X_t))(B(a) - B(X_t)).$$

According to (12.9) we have

$$\begin{aligned} G_{z_2}(a) &= E_a \int_0^{\alpha_{z_2}} f(t) dt = \sum_{n=1}^{\infty} E_a \left(\int_0^{\alpha_{z_1}^n} f(t) dt; \alpha_{z_1}^{n-1} \leq \nu_{a-z_2} < \alpha_{z_1}^n \right) \\ &= \sum_{n=1}^{\infty} E_a \left(\sum_{k=1}^n \int_{\alpha_{z_1}^{k-1}}^{\alpha_{z_1}^k} f(t) dt; \alpha_{z_1}^{n-1} \leq \nu_{a-z_2} < \alpha_{z_1}^n \right) \\ &= \sum_{k=1}^{\infty} E_a \left(\int_{\alpha_{z_1}^{k-1}}^{\alpha_{z_1}^k} f(t) dt; \alpha_{z_1}^{k-1} \leq \nu_{a-z_2} \right) \\ &= \sum_{k=1}^{\infty} E_a \left(\exp\left(-\int_0^{\alpha_{z_1}^{k-1}} h(X_s) ds\right) \left(\int_0^{\alpha_{z_1}} f(t) dt \circ \theta_{\alpha_{z_1}^{k-1}} \right); \alpha_{z_1}^{k-1} \leq \nu_{a-z_2} \right) \\ &= \sum_{k=1}^{\infty} E_a \left(\exp\left(-\int_0^{\alpha_{z_1}^{k-1}} h(X_s) ds\right); \alpha_{z_1}^{k-1} \leq \nu_{a-z_2} \right) E_a \int_0^{\alpha_{z_1}} f(t) dt \\ &= G_{z_1}(a) \sum_{k=0}^{\infty} E_a \left(\exp\left(-\int_0^{\alpha_{z_1}^k} h(X_s) ds\right); \alpha_{z_1}^k \leq \nu_{a-z_2} \right). \end{aligned}$$

On the other hand

$$\begin{aligned} &E_a \left(\exp\left(-\int_0^{\alpha_{z_1}^k} h(X_s) ds\right); \alpha_{z_1}^k \leq \nu_{a-z_2} \right) \\ &= E_a \left(\exp\left(-\int_0^{\alpha_{z_1}^{k-1}} h(X_s) ds\right) \left(\exp\left(-\int_0^{\alpha_{z_1}} h(X_s) ds\right) \circ \theta_{\alpha_{z_1}^{k-1}} \right); \right. \\ &\quad \left. \alpha_{z_1}^{k-1} \leq \nu_{a-z_2}, \alpha_{z_1}^{k-1} + \alpha_{z_1} \leq \alpha_{z_1}^{k-1} + \nu_{a-z_2} \right) \\ &= E_a \left(\exp\left(-\int_0^{\alpha_{z_1}^{k-1}} h(X_s) ds\right); \alpha_{z_1}^{k-1} \leq \nu_{a-z_2} \right) \\ &\quad \times E_a \left(\exp\left(-\int_0^{\alpha_{z_1}} h(X_s) ds\right); \alpha_{z_1} \leq \nu_{a-z_2} \right) \\ &= \left(E_a \left(\exp\left(-\int_0^{\alpha_{z_1}} h(X_s) ds\right); \alpha_{z_1} \leq \nu_{a-z_2} \right) \right)^k. \end{aligned}$$

From here

$$\begin{aligned}
G_{z_2}(a) &= G_{z_1}(a) \sum_{k=0}^{\infty} \left(E_a \left(\exp \left(- \int_0^{\alpha_{z_1}} h(X_s) ds \right); \alpha_{z_1} \leq \nu_{a-z_2} \right) \right)^k \\
&= \frac{G_{z_1}(a)}{1 - E_a \left(\exp \left(- \int_0^{\alpha_{z_1}} h(X_s) ds \right); \alpha_{z_1} \leq \nu_{a-z_2} \right)}. \quad \square
\end{aligned}$$

Thus a sign of $G_z(a)$ does not depend on z . For discrete case a character of varying sequence of links of a trajectory of choice is determined by function $G_0(a)$. For continuous case an information about this sequence consists in a derivative of the function $G_z(a)$ with respect to z at point $z = 0$. If this derivative is positive any $G_z(a)$ is positive and vice versa. For diffusion Markov processes these functions and their derivatives can be found analytically and evaluated with the help of computer.

Thus if the differences between links and their succeeding ones are positive ($G_z(a) < 0$) it means that for any alternating sequence of levels the zigzag trajectory of choice has an average moving to SO^e . If these differences are negative ($G_z(a) > 0$), the zigzag trajectory of choice has the average moving to NW (see Figure 12.3). It seems as a sailing of a yacht by the wind. Let us note, however, that a difference of position of the “yacht” on the same level at two succeeding “halses” is multiplied any time by the factor (for example, $U_{\tau_{k-1} + \mu_a}(x)$), which depends on history of the process up to the current instant ($\tau_{k-1} + \mu_a$) and decreases as a power k of some small value. It means that a distance of such a sailing along the given level cannot be infinite.

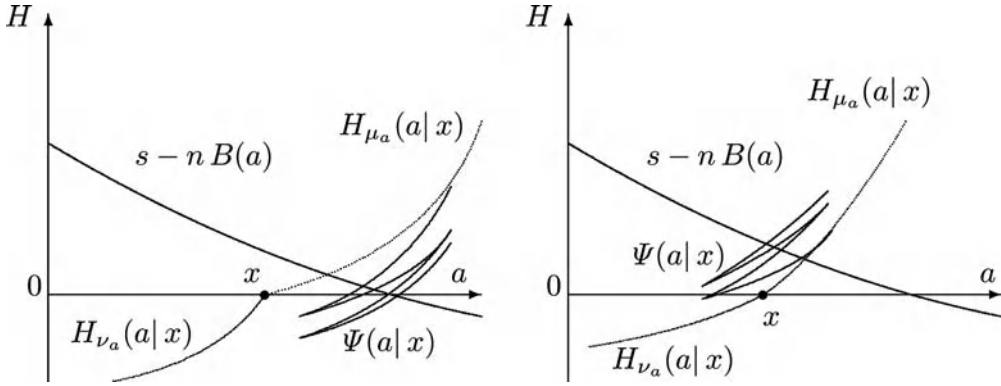


Figure 12.3. Trajectories of choice with moving on SO^e and on NW .

Any intersection of ascending links of such zigzags with a decreasing line $s - nB(a)$ determines a local minimum. Between these local minimums more preferable is such, whose abscissa of an intersection point is less than that of others.

Using trajectory of choice we can prove that if on the whole half-line $G_z(a) < 0$, then the trivial time of the first fulfilment of the condition (12.3) is the optimal Markov time. Actually, let c be a level of the trivial solution of (12.3) and there exists a solution τ of (12.3) on a level $b < c$. So we have $X_\tau = b$. Let, for example, $x < b < c$. The point $(b, s - B(b))$ is located on NW from the line (H_{μ_a}) ($x < a < c$). Denote $\alpha_{b,r} = \nu_{b-r} + \mu_b$, $\alpha_{b,r}^n = \alpha_{b,r}^{n-1} + \alpha_{b,r}$ ($n \geq 1$, $\alpha_{b,r}^0 = 0$) and $\beta_n = \mu_b + \alpha_{b,r}^n$ ($n \geq 0$). The following theorem is proved in [4]:

Theorem 3. *Let (P_x) be a consistent family of measures of a Markov process and $X_\tau = b$ on $\{\tau < \infty\}$; let $\alpha_{b,r} \rightarrow 0$ ($r \rightarrow 0$) and for any $r > 0$ $\alpha_{b,r}^n \rightarrow \infty$ ($n \rightarrow \infty$) P_b -a.s. Then $(\forall x \in \mathbb{R})$ a mixed time*

$$\tau_r = \mu_b I(\tau = \mu_b) + \sum_{n=1}^{\infty} \beta_n I(\beta_{n-1} < \tau \leq \beta_n)$$

is defined and converges to τ P_x -a.s. as $r \rightarrow 0$.

Consequently, the right-hand side of (12.4) is equal to $\lim_{r \rightarrow 0} H_{\tau_r}(b|x)$, and every $H_{\tau_r}(b|x)$ is a mixture of $H_{\beta_n}(b|x)$. On the other hand, every $H_{\beta_n}(b|x)$ can be obtained as an end point of a zigzag trajectory of choice. And we know that for $G_z(a) < 0$ ($\forall a > 0$) this zigzag trajectory goes towards SO^{st} from the line (H_{μ_a}) ($x < a < c$). It is a contradiction.

If there is some part of half-line where $G_z(a) > 0$, then different solutions of the problem on different parts of half-line can be. So the problem requires further investigation.

12.4 Degradation Process of Diffusion Type

Let us consider a continuous Markov process of diffusion type. We assume that shift and volatility parameters of the process are bounded. It implies that for any $a < x < b$ probabilities $P_x(\mu_b < \infty)$, $P_x(\nu_a < \infty)$ are positive. Let us denote

$$\begin{aligned} f_a(x) &= U_{\nu_a}(x), & g_a(x) &= V_{\nu_a}(x), & k_a(x) &= W_{\nu_a}(x) & (a < x), \\ \bar{f}_a(x) &= U_{\mu_a}(x), & \bar{g}_a(x) &= V_{\mu_a}(x), & \bar{k}_a(x) &= W_{\mu_a}(x) & (a > x). \end{aligned}$$

Then

$$\begin{aligned} G_z(a) &= B(a)(g_{a-z}(a) + f_{a-z}(a)\bar{g}_a(a-z)) + k_{a-z}(a) + f_{a-z}(a)\bar{k}_a(a-z) \\ &\quad - C(a)(1 - f_{a-z}(a)\bar{f}_a(a-z)), \end{aligned}$$

where

$$C(a) = r - s - (m - n)B(a) = \frac{r - s + (m - n)A(a)}{1 + (m - n)h(a)}.$$

Let us denote a derivative with respect to a lower index with a dot and that with respect to an argument in parentheses with a prime. Taking into account that $g_a(a) = \bar{g}_a(a) = k_a(a) = \bar{k}_a(a) = 0$ and $f_a(a) = \bar{f}_a(a) = 1$, we have

$$-\dot{G}_0(a) = B(a)(\dot{g}_a(a) + \dot{\bar{g}}_a(a)) + \dot{k}_a(a) + \dot{\bar{k}}_a(a) + C(a)(\dot{f}_a(a) + \dot{\bar{f}}_a(a)).$$

The similar expression we can obtain for $\bar{G}_z(a)$. Taking into account that

$$\begin{aligned} \dot{f}_a(a) &= -f'_a(a), & \dot{g}_a(a) &= -g'_a(a), & \dot{k}_a(a) &= -k'_a(a), \\ \dot{\bar{f}}_a(a) &= -\bar{f}'_a(a), & \dot{\bar{g}}_a(a) &= -\bar{g}'_a(a), & \dot{\bar{k}}_a(a) &= -\bar{k}'_a(a), \end{aligned}$$

we derive that $\dot{G}_0(a) = \dot{\bar{G}}_0(a)$.

Wiener process with a drift as a model of degradation.

Recently this process with a positive drift is being frequently used as a model of degeneration (see, for example, [3]).

Let (X_t) ($t \geq 0$) be a Wiener process with a constant drift $c > 0$. From well-known Feinman-Kac formula it follows that functions f_a and \bar{f}_a satisfy the differential equation

$$\frac{1}{2}f'' + cf' - h(x)f = 0, \quad (12.10)$$

functions g_a and \bar{g}_a that of the differential equation

$$\frac{1}{2}g'' + cg' - h(x)g + 1 = 0, \quad (12.11)$$

and functions k_a and \bar{k}_a that of the differential equation

$$\frac{1}{2}k'' + ck' - h(x)k + A(x) = 0. \quad (12.12)$$

These equations were solved under appropriate boundary conditions. We assume $h(x) = x^+ = \max\{x, 0\}$ and $A(x) = \text{constant}$. On Figure 12.4 we show graphs of the function $\dot{G}_0(a)$ for three values of c obtained on a computer. A rule of behaviour for an operator of the system on the whole half-line is determined uniquely only for $c = 1$. He must begin prophylactic repair at the first instant when (12.3) holds. For $c = 0.01$ and $c = 0.1$ he must estimate an initial degradation of the system and evaluate the function $\dot{G}_0(\cdot)$ in this point. If this value is positive he can wait for a failure, because a finite instant of a global minimum of the loss function does not exist. But if it is negative he can wait for the first fulfilment of condition (12.3). But his decision may be non-optimal, if the estimated point is near the boundary between positive and negative regions. Determination of such a zone of uncertainty requires further investigation. However, one should note that such an uncertainty arises only for small shifts. Our evaluation shows that from $c > 0.3$, a Wiener process from the point of view of reliability is equivalent to a non-decreasing process. For both of them the optimal stopping time is the first fulfilment of some conditions.

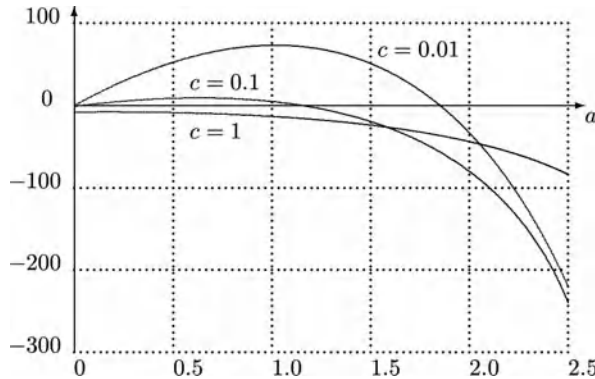


Figure 12.4. Derivative of $G_r(a)$ with respect to r as $r = 0$ for different drift

References

1. Bagdonavicius, V., Nikulin, M.S.: Statistical analysis of degradation data in dynamic environment. In: Dipartimento di matematica “Guido Castelnuovo”, 1–20 (2000)
2. Harlamov, B.P.: Continuous Semi-Markov Processes. ISTE & Wiley, London (2007)
3. Harlamov, B.P.: Optimal prophylaxis policy for systems with partly observable parameters. In: Ionescu, V., Limnios, N. (ed) Statistical and Probabilistic Models in Reliability. Birkhauser, Boston. 265–278 (1999)
4. Lehmann, A.: Degradation-threshold-shock models. In: Nikulin, M.S., Commenges, D., Huber, C. (ed) Probability, Statistics and Modelling in Public Health. Springer, Berlin Heidelberg New York. 286–298 (2006)
5. Rasova, S.S., Harlamov, B.P.: Optimal local first exit time. Zapiski seminarov POMI, **361**, 83–108 (2008) (in Russian)

Deterioration Processes with Increasing Thresholds

S. Zacks

Department of Mathematical Sciences, Binghamton University, Binghamton, NY, USA

Abstract: The present chapter derives the reliability functions and hazard functions, when the threshold for deterioration is an increasing function of time. Four cases are considered. Case I: The threshold is a step function with K jumps at known points. Case II: The threshold is a step function with K jumps, where the location of jumps are random, following a renewal process. Case III: The controlled case. The jumps occur after the first crossing of a control limit. Case IV: The threshold increases linearly. The theory is illustrated in all four cases with a Markovian deterioration process, i.e., a compound Poisson process with i.i.d. jumps following an exponential distribution.

Keywords and phrases: Compound Poisson process, deterioration process, hazard function, increasing threshold function, reliability function

13.1 Introduction

The present chapter is a fourth one in a series of papers on the reliability (availability) of systems subjected to compound Poisson deterioration processes. In the first paper (Zacks [1]) non-homogeneous compound Poisson processes were considered. Failure time distributions and their moments were studied. The second paper (Zacks [3]) discussed failure distributions associated with general compound renewal processes. The third paper (Zacks [4]) studies the availability of systems which are replaced after failure, either immediately or after a random time. In the present chapter, the failure times are the first crossing of a deterioration process $\{D(t), t \geq 0\}$ of an increasing boundary, $B(t)$, where $B(0) = \beta_0$, $0 < \beta_0 < \infty$. More specifically, let $D(t) = \sum_{n=0}^{N(t)} X_n$ denote a deterioration process, where $\{N(t)\}$ is a Poisson process and X_1, X_2, \dots are i.i.d. random variables representing the amount of damage to the system ($X_n > 0, n \geq 1$) in each shock event. The system fails at the stopping time

$$T = \inf\{t > 0 : D(t) \geq B(t)\}. \quad (13.1)$$

There are various applications of the present theory. We can think of a bridge that once in a while is strengthened. Another application is in insurance. For a given insurance

file, let $B(t)$ be the total amount of funds entering the system (premiums, etc.) up to time t and $D(t)$ are the total claims (minus deductibles) up to time t . The stopping time T is called in insurance a ruin time. The reliability of a system at time t is $R(t) = P\{T > t\}$. The corresponding hazard function is $\Lambda(t) = -d/dt R(t)/R(t)$. We consider several cases.

Case I:

The first case is of an increasing step function

$$B(t) = \sum_{n=1}^K I\{\tau_{i-1} < t \leq \tau_i\} \beta_{i-1} + I(\tau_K < t) \beta_K, \quad (13.2)$$

where $0 < \beta_0 < \beta_1 < \dots < \beta_K$, τ_i , $i = 1, 2, \dots, K$ are fixed growth times, with $\tau_0 = 0$.

Case II:

The second case holds when $\{\tau_i, i \geq 1\}$ are random times following a renewal process.

Case III:

The third case is that of controlled threshold growth. In this case

$$\tau_i = \inf\{t > 0 : D(t) \geq \gamma_i\}, \quad i = 1, 2, \dots, \quad (13.3)$$

where $\gamma_i < \beta_i$ and $\gamma_i < \gamma_{i+1}$ ($i = 1, 2, \dots$).

Notice that the control limits are below the threshold levels. The instant $D(t)$ crosses a control limit, γ ; the threshold level, β , is increased.

Case IV:

The fourth case is that of linearly increasing threshold $B(t) = \beta_0 + \gamma t$, where $0 < \beta_0 < \infty$ and $0 < \gamma < \infty$. This is the usual case in insurance ruin problems.

We start with the general theory for each case, followed by the special case of exponential damage.

13.2 Preliminaries

In the present chapter we focus attention on processes $\{D(t), t \geq 0\}$ which are compound Poisson, i.e., $D(t) = \sum_{n=0}^{N(t)} X_n$, $t \geq 0$, where $\{N(t), t \geq 0\}$ is a homogeneous Poisson process; $X_0 \equiv 0$; X_1, X_2, \dots are independent identically distributed (i.i.d.) random variables, independent of $\{N(t); t \geq 0\}$. We further assume that X_n have a common absolutely continuous distribution, F , $F(0) = 0$, with density function f . Let $H(y; t) = P\{D(t) \leq y\}$ be the distribution function of $D(t)$. H has an atom at $y = 0$, i.e., $H(0; t) = e^{-\lambda t}$, where λ is the intensity coefficient of $\{N(t), t \geq 0\}$. Moreover, for $0 < y < \infty$,

$$H(y; t) = \sum_{n=0}^{\infty} p(n; \lambda t) F^{(n)}(y), \quad (13.4)$$

where $p(n; \lambda t) = e^{-\lambda t} (\lambda t)^n / n!$, $n \geq 0$, is the probability mass function of the Poisson distribution $\text{Pois}(\lambda t)$.

$F^{(n)}(y) = \int_0^y f(x)F^{(n-1)}(y-x)dx$, all $y > 0$, $n \geq 1$, where $F^{(0)}(y) = 1$ for all $y > 0$. This is the n -fold convolution of F . The density function of $D(t)$, for $y \in (0, \infty)$ is

$$h(y; t) = \sum_{n=1}^{\infty} p(n; \lambda t) f^{(n)}(y), \quad (13.5)$$

where $f^{(n)}(y)$, $n \geq 1$, is the n -fold convolution of f . Notice that $f^{(n)}(y) = 0$ for all $y \leq 0$ and for $y > 0$, $f^{(n)}(y) = \int_0^y f(x)f^{(n-1)}(y-x)dx$, $n \geq 2$, with $f^{(1)}(y) = f(y)$, $y > 0$.

The process $\{D(t), t \geq 0\}$ is a strongly Markovian jump process, where the epochs of jump are the arrival times of the Poisson renewal points. Let T be a stopping time defined in (13.1). Define the defective density

$$g(y; t) = \frac{d}{dy} P\{D(t) \leq y, T > t\}, \quad 0 < y < B(t). \quad (13.6)$$

Since $B(t) \geq \beta_0$ for all t ,

$$g(y; t) = h(y; t) \quad \text{for all } 0 < y \leq \beta_0. \quad (13.7)$$

The system reliability when the threshold is $\{B(t), t \geq 0\}$ is

$$R(t) = P\{T > t\} = e^{-\lambda t} + \int_0^{B(t)} g(y; t) dy. \quad (13.8)$$

The corresponding hazard function is

$$\begin{aligned} \Lambda(t) &= -\frac{d}{dt} R(t)/R(t) \\ &= \left[\lambda e^{-\lambda t} + B'(t)g(B(t); t) + \int_0^{B(t)} \left(\frac{\partial}{\partial t} g(y; t) \right) dy \right] / R(t), \end{aligned} \quad (13.9)$$

where $B'(t) = \frac{d}{dt} B(t)$. Notice that if $B(t)$ is the step function (13.2), then $B'(t) = 0$ for all $t \neq \tau_i$. A specific hazard function for this case will be developed in the next section.

13.3 The Reliability and Hazard in Case I

In Case I the threshold is a step function. In the case of $K = 0$ the threshold is a constant β_0 , and then the reliability function is

$$R(t; \beta_0) = H(\beta_0; t). \quad (13.10)$$

This is obviously a decreasing function of t on $(0, \infty)$. The hazard function is then

$$\begin{aligned} \Lambda(t; \beta_0) &= -\frac{d}{dt} H(\beta_0; t) / H(\beta_0; t) \\ &= \lambda \sum_{n=0}^{\infty} p(n; \lambda t) (F^{(n)}(\beta_0) - F^{(n+1)}(\beta_0)) / H(\beta_0; t). \end{aligned} \quad (13.11)$$

Generally, when $K \geq 1$

$$B(t; \tau, \beta) = \sum_{i=1}^K I(\tau_{i-1} < t \leq \tau_i) \beta_{i-1} + I(\tau_K < t) \beta_K. \quad (13.12)$$

Define recursively the defective density $g(t; \tau, \beta)$ as follows:

$$g(x; t) = h(x; t) I(x \leq \beta_0), \quad \text{if } 0 < t \leq \tau_1. \quad (13.13)$$

For $i \geq 2$,

$$\begin{aligned} g(x; t) &= \sum_{i=2}^{\infty} I(\tau_{i-1} < t \leq \tau_i) \int_0^{x \wedge (\beta_{i-2})} g(y; \tau_{i-1}) \cdot \\ &\quad \cdot h(x - y; t - \tau_{i-1}) dy \cdot I(0 < x \leq \beta_{i-1}). \end{aligned} \quad (13.14)$$

Notice that if $x \leq \beta_0$, then $g(x; t) = h(x; t)$ for all $t > 0$.

The reliability function is given by

$$\begin{aligned} R_K(t; \tau, \beta) &= I(t \leq \tau_1) H(\beta_0; t) \\ &\quad + \sum_{i=1}^{K-1} I(\tau_i < t \leq \tau_{i+1}) \left[e^{-\lambda \tau_i} H(\beta_i; t - \tau_i) \right. \\ &\quad \left. + \int_0^{\beta_{i-1}} g(x; \tau_i) H(\beta_i - x; t - \tau_i) dx \right] \\ &\quad + I(t > \tau_K) \left[e^{-\lambda \tau_K} H(\beta_K; t - \tau_K) \right. \\ &\quad \left. + \int_0^{\beta_{K-1}} g(x; \tau_K) H(\beta_K - x; t - \tau_K) dx \right]. \end{aligned} \quad (13.15)$$

The hazard function is

$$\Lambda_K(t; \tau, \beta) = \frac{-\frac{d}{dt} R_K(t; \tau, \beta)}{R_K(t; \tau, \beta)}.$$

Let

$$\begin{aligned} D_0(t; \gamma) &= -\frac{d}{dt} H(\gamma; t) \\ &= \lambda \sum_{n=0}^{\infty} p(n; \lambda t) [F^{(n)}(\gamma) - F^{(n+1)}(\gamma)]. \end{aligned} \quad (13.16)$$

Thus we obtain $-\frac{d}{dt} R_K(t; \tau, \beta)$ by substituting in (13.15) the function $D_0(t, \beta_0)$ corresponding to $H(\beta_0, t)$; the function $D_0(t - \tau_i; \beta_i)$ corresponding to $H(\beta_i; t - \tau_i)$; etc.

13.4 The Reliability and Hazard in Case II

In Case II, the K jump points of $B(t)$ are randomly assigned according to some renewal process. More specifically, let U_1, U_2, \dots, U_K be i.i.d. positive random variables representing the times elapsing between the jumps of $B(t)$, i.e., $\tau_1 = U_1$, $\tau_2 = U_1 + U_2$, \dots , $\tau_K = U_1 + \dots + U_K$. Let $\psi(u)$ denote the density of U on $(0, \infty)$. The reliability function is $R_K^*(t; \beta_0, \beta_1, \dots, \beta_K)$. This function is defined recursively as follows:

(i) For $K = 1$, let

$$\begin{aligned} R_1^*(t; \beta_0, \beta_1) &= P(U > t)H(\beta_0; t) \\ &+ \int_0^t \psi(u) \left[e^{-\lambda u} H(\beta_1; t - u) \right. \\ &\left. + \int_0^{\beta_0} h(y; u) H(\beta_1 - y; t - u) dy \right] du. \end{aligned} \quad (13.17)$$

For each $n = 2, \dots, K$, define

$$\begin{aligned} R_n^*(t; \beta_0, \beta_1, \dots, \beta_n) &= P(u > t)H(\beta_0; t) \\ &+ \int_0^t \psi(u) \left[e^{-\lambda u} R_{n-1}^*(t - u; \beta_1, \dots, \beta_n) \right. \\ &\left. + \int_0^{\beta_0} h(y; u) R_{n-1}^*(t - u; \beta_1 - y, \dots, \beta_n - y) dy \right] du. \end{aligned} \quad (13.18)$$

General solution of this integral difference equation is complicated, we will develop later some special cases.

The hazard function corresponding to $R_1^*(t; \beta_0, \beta_1)$ is

$$\begin{aligned} A_1(t; \beta_0, \beta_1) &= \left\{ P(U > t)D_0(t; \beta_0) \right. \\ &+ \int_0^t \psi(u) \left[e^{-\lambda u} D_0(t - u; \beta_1) \right. \\ &+ \left. \int_0^{\beta_0} h(y; u) D_0(t - u; \beta_1 - y) dy \right] du \left. \right\} \\ &\div R_1^*(t; \beta_0, \beta_1). \end{aligned} \quad (13.19)$$

For $n = 1, \dots, K$, define

$$\begin{aligned} D_n(t; \beta_0, \dots, \beta_n) &= P(U > t)D_0(t; \beta_0) \\ &+ \int_0^t \psi(u) \left[e^{-\lambda u} D_{n-1}(t - u; \beta_1, \dots, \beta_n) \right. \\ &\left. + \int_0^{\beta_0} h(y; u) D_{n-1}(t - u; \beta_1 - y, \dots, \beta_n - y) dy \right] du. \end{aligned} \quad (13.20)$$

Then, the hazard function of the system is

$$\lambda_K(t; \beta_0, \dots, \beta_K) = \frac{D_K(t; \beta_0, \dots, \beta_K)}{R_K^*(t; \beta_0, \dots, \beta_K)}. \quad (13.21)$$

13.5 Reliability and Hazard in Case III

In Case III the jump from β_0 to a larger level is at the first time the process $\{D(t), t \geq 0\}$ crosses a horizontal control line at level γ_0 , $\gamma_0 < \beta_0$. The jump point is at

$$T_c = \inf\{t > 0 : D(t) \geq \gamma_0\}. \quad (13.22)$$

Let $R(T_c) = D(T_c) - \gamma_0$. If $D(T_c) \geq \beta_0$ the system fails at time T_c . If $D(T_c) < \beta_0$ the system survives at T_c and the threshold increases immediately to β_1 . The joint density of $(T_c, R(T_c))$ is

$$p(t, r; \gamma_0) = \lambda e^{-\lambda t} f(\gamma_0 + r) + \lambda \int_0^{\gamma_0} h(y; t) f(\gamma_0 + r - y) dy. \quad (13.23)$$

The marginal density of T_c is

$$\begin{aligned} \psi_c(t; \gamma_0) &= \int_0^\infty p(t, r; \gamma_0) dr \\ &= \lambda e^{-\lambda t} \bar{F}(\gamma_0) + \lambda \int_0^{\gamma_0} h(y; t) \bar{F}(\gamma_0 - y) dy, \end{aligned} \quad (13.24)$$

where $\bar{F}(x) = 1 - F(x)$. We formulate now the reliability function for the simpler case of one increase of the threshold, i.e.,

$$B_c(t) = I(t \leq T_c) \beta_0 + I(t > T_c) \beta_1, \quad (13.25)$$

where $0 < \beta_0 < \beta_1$. The system fails as soon as $D(t) \geq B_c(t)$. Thus, the conditional reliability given T_c is

$$R_c(t | T_c) = I(t \leq T_c) + I(t > T_c, D(T_c) < \beta_0) H(\beta_1 - D(T_c); t - T_c). \quad (13.26)$$

Hence, the reliability function is

$$R_c(t; \gamma_0, \beta_0, \beta_1) = H(\gamma_0; t) + \int_0^t \int_0^{\beta_0 - \gamma_0} p(s, r; \gamma_0) H(\beta_1 - \gamma_0 - r; t - s) dr ds. \quad (13.27)$$

The corresponding hazard function is

$$\begin{aligned} \Lambda_c(t; \gamma_0, \beta_0, \beta_1) &= \left\{ D_0(t; \gamma_0) + \int_0^t \int_0^{\beta_0 - \gamma_0} p(s, r; \gamma_0) D_0(t - s; \beta_1 - \gamma_0 - r) dr ds \right\} \\ &\div R_c(t; \gamma_0, \beta_0, \beta_1). \end{aligned} \quad (13.28)$$

13.6 Reliability and Hazard in Case IV

In Case IV the threshold boundary is

$$B(t) = \beta + \gamma t, \quad 0 \leq t < \infty; \quad 0 < \gamma < \infty, \quad 0 < \beta < \infty. \quad (13.29)$$

Without loss of generality we can assume that $\gamma = 1$. Indeed, if $\gamma \neq 1$ make the transformation $\lambda' = \lambda \cdot \gamma$ and $t' = t/\gamma$. Then $p(n; \lambda' t) = p(n; \lambda t)$ for all $n \geq 0$. Consider the stopping time

$$T(\beta) = \inf\{t > 0 : D(t) \geq \beta + t\}. \quad (13.30)$$

This is the failure time of the system. As in (13.8),

$$R_L(t; \beta) = e^{-\lambda t} + \int_0^{\beta+t} g(x; t, \beta) dx. \quad (13.31)$$

As shown in Zacks [2], the defective density $g(x; t, \beta)$ is given in terms of the density $h(x; t)$ as

$$\begin{aligned} g(x; t, \beta) = & h(x; t)I(x \leq \beta) \\ & + I(\beta < x \leq \beta + t) \left[h(x; t) - e^{-\lambda(\beta+t-x)} h(x; x - \beta) \right. \\ & \left. - (\beta + t - x) \int_0^{x-\beta} h(u + \beta; u) \frac{1}{t-u} h(x - \beta - u; t - u) du \right]. \end{aligned} \quad (13.32)$$

Accordingly, the reliability function is

$$\begin{aligned} R_L(t; \beta) = & H(\beta + t; t) - \int_{\beta}^{\beta+t} e^{-\lambda(\beta+t-x)} \cdot \\ & \cdot h(x; x - \beta) dx - \int_{\beta}^{\beta+t} (\beta + t - x) \int_0^{x-\beta} h(u + \beta; u) \cdot \\ & \cdot \frac{1}{t-u} h(x - \beta - u; t - u) du dx. \end{aligned} \quad (13.33)$$

Simple change of variables yields the formula

$$R_L(t; \beta) = H(\beta + t; t) - \int_0^t h(\beta + t - y; t - y) \left[e^{-\lambda y} + y \int_0^1 (1 - z) h(yz; y) dz \right] dy. \quad (13.34)$$

Let

$$M(y) = e^{-\lambda y} + y \int_0^1 (1 - z) h(yz; y) dz. \quad (13.35)$$

We can write then

$$R_L(t; \beta) = H(\beta + t; t) - \int_0^t M(y) h(\beta + t - y; t - y) dy. \quad (13.36)$$

The corresponding hazard function is

$$A(t; \beta) = \left\{ D_0(t; \beta + t) - h(\beta + t; t) + \int_0^t M(y) \left(\frac{\partial}{\partial t} h(\beta + t - y; t - y) \right) dy \right\} \div R(t; \beta). \quad (13.37)$$

13.7 Exponential Deterioration

In the present section we focus attention on the special case where the amount of deterioration in each shock has an exponential distribution, i.e., $F(x) = 1 - e^{-\mu(x)}$, $x \geq 0$. In this case

$$h(x; t) = \mu \sum_{n=1}^{\infty} p(n; \lambda t) \cdot p(n-1; \mu x) \quad (13.38)$$

and

$$\begin{aligned} H(x; t) &= e^{-\lambda t} + \sum_{n=1}^{\infty} p(n; \lambda t) (1 - P(n-1; \mu x)) \\ &= 1 - \sum_{n=1}^{\infty} p(n; \lambda t) P(n-1; \mu x) \\ &= \sum_{j=0}^{\infty} p(j; \mu x) P(j; \lambda t). \end{aligned} \quad (13.39)$$

13.7.1 Case I with $K = 1$

In the case of $K = 1$ the threshold function is

$$B(t; \tau_1, \beta_0, \beta_1) = \beta_0 I(t \leq \tau) + \beta_1 I(t > \tau_1). \quad (13.40)$$

Let $\zeta_0 = \mu\beta_0$ and $\zeta_1 = \mu\beta_1$. The reliability function is then

$$\begin{aligned} R_1(t; \tau_1, \beta_0, \beta_1) &= I(t \leq \tau_1) \sum_{n=0}^{\infty} p(n; \zeta_0) P(n; \lambda t) \\ &\quad + I(t > \tau_1) \left[e^{-\lambda\tau_1 - \zeta_1} + \sum_{l=1}^{\infty} p(l; \zeta_1) \left(e^{-\lambda\tau_1} P(l; \lambda(t - \tau_1)) \right. \right. \\ &\quad \left. \left. + \sum_{n=1}^l p(n; \lambda\tau_1) P(l-n; \lambda(t - \tau_1)) I_{\beta_0/\beta_1}(n, l-n+1) \right) \right], \end{aligned} \quad (13.41)$$

where $I_x(\nu_1, \nu_2)$ denotes the incomplete beta function ratio at x , $0 < x < 1$. This is the c.d.f. of $\text{beta}(\nu_1, \nu_2)$.

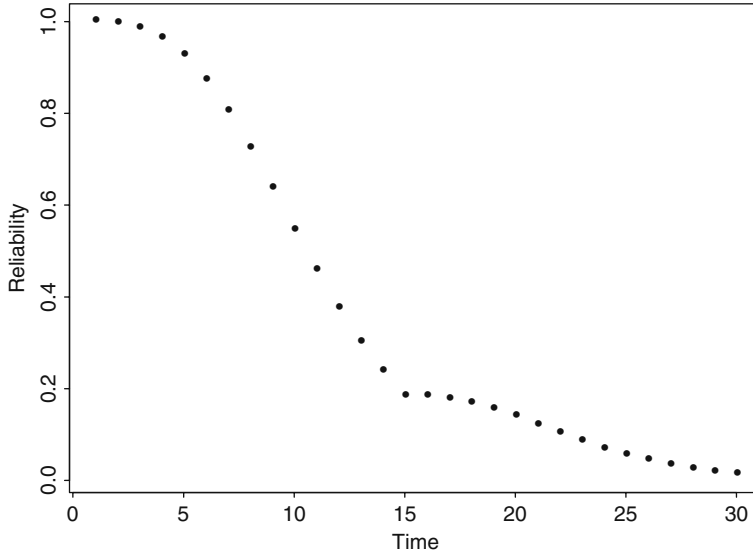


Figure 13.1. The reliability function $R_1(t; 15, 10, 15)$

In Figure 13.1 we show such a reliability function for the case of $\tau_1 = 15$, $\beta_0 = 10$, $\beta_1 = 15$, $\mu = 2$, and $\lambda = 1$.

The hazard function is

$$A_1(t; \tau_1, \beta_0, \beta_1) = \frac{-\frac{d}{dt} R_1(t; \tau_1, \beta_0, \beta_1)}{R_1(t; \tau_1, \beta_0, \beta_1)}. \quad (13.42)$$

Notice that $-\frac{d}{dt} P(n; \lambda t) = \lambda p(n; \lambda t)$. Thus, according to (13.41),

$$\begin{aligned} -\frac{d}{dt} R_1(t; \tau_1, \beta_0, \beta_1) &= I(t \leq \tau_1) \lambda \sum_{n=0}^{\infty} p(n; \zeta_0) \cdot \\ &\cdot p(n; \lambda t) + I(t > \tau_1) \left[\lambda \sum_{l=1}^{\infty} p(l; \zeta_1) \left(e^{-\lambda \tau_1} p(l; \lambda(t - \tau_1)) \right. \right. \\ &\left. \left. + \sum_{n=1}^l p(n; \lambda \tau_1) p(l - n; \lambda(t - \tau_1)) I_{\beta_0/\beta_1}(n, l - n + 1) \right) \right]. \end{aligned} \quad (13.43)$$

The hazard function is then the ratio of (13.43) over (13.41). In Figure 13.2 we present the hazard function corresponding to the reliability function of Figure 13.1.

13.7.2 Case I with $K = 2$

In the present case,

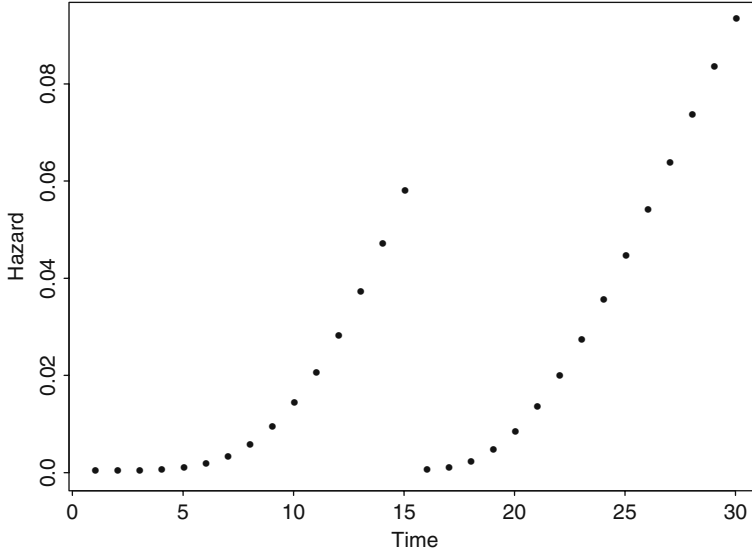


Figure 13.2. The hazard function $A_1(t; 15, 10, 15)$ with $\mu = 2$ and $\lambda = 1$

$$\begin{aligned}
 R_2(t; \tau_1, \tau_2, \beta_0, \beta_1, \beta_2) &= I(t \leq \tau_2) R_1(t; \tau_1, \beta_0, \beta_1) \\
 &\quad + I(t > \tau_2) \left[e^{-\lambda \tau_2} H(\beta_2; t) \right. \\
 &\quad \left. + \int_0^{\beta_1} g_2(x; \tau_1, \tau_2, \beta_0) H(\beta_2 - x; t) dx \right],
 \end{aligned} \tag{13.44}$$

where

$$g_2(x; \tau_1, \tau_2, \beta_0) = \int_0^{x \wedge \beta_0} h(y; \tau_1) h(x - y; \tau_2 - \tau_1) dy. \tag{13.45}$$

Notice that if $x \leq \beta_0$, then $g_2(x; \tau_1) = h(x; \tau_1)$. For $\beta_0 < x \leq \beta_1$ we get

$$\begin{aligned}
 g_2(x; \tau_1, \tau_2, \beta_0) &= \int_0^{\beta_0} h(y; \tau_1) h(x - y; \tau_2 - \tau_1) dy \\
 &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} p(n; \lambda \tau_1) p(m; \lambda(\tau_2 - \tau_1)) \int_0^{\beta_0} f^{(n)}(y) f^{(m)}(x - y) dy \\
 &= \mu \sum_{l=2}^{\infty} p(l - 1; \mu x) e^{-\lambda \tau_2} \frac{\lambda^l}{l!} \sum_{n=1}^{l-1} \binom{l}{n} \tau_1^n (\tau_2 - \tau_1)^{l-n} I_{\beta_0/x}(n, l - n).
 \end{aligned} \tag{13.46}$$

In Figure 13.3 we show the reliability function $R_2(t; \tau_1, \tau_2, \beta_0, \beta_1, \beta_2)$ with $\lambda = 3$, $\mu = 2$, $\tau_1 = 10$, $\tau_2 = 15$, $\beta_0 = 10$, $\beta_1 = 15$, and $\beta_2 = 20$.

13.7.3 Case II with $K = 1$

We assume here that the time of change τ_1 has an exponential distribution, i.e., $P(U > t) = e^{-kt}$ and $\psi(U) = ke^{-ku}$, $u > 0$. According to (13.17) the reliability function is

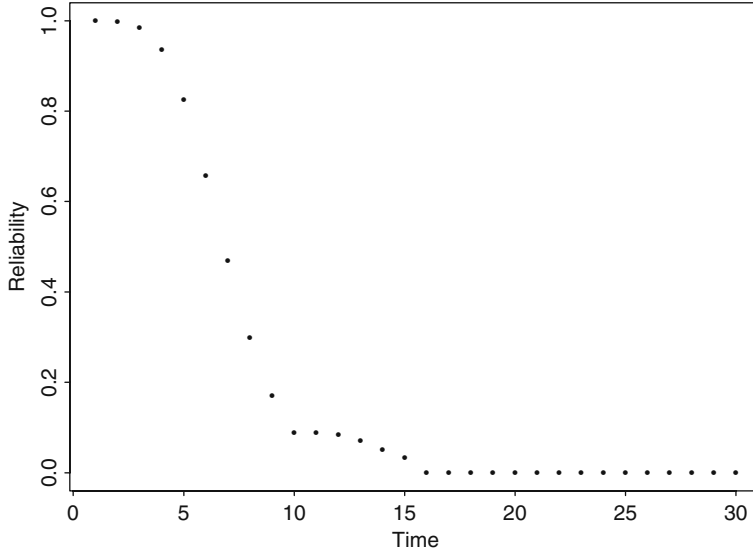


Figure 13.3. Reliability function $R_2(t; \tau_1, \tau_2, \beta)$ with $K = 2$

$$\begin{aligned}
 R_1^*(t; k, \beta_0, \beta_1) &= e^{-kt} \sum_{n=0}^{\infty} p(n; \zeta_0) P(n; \lambda t) \\
 &+ k \int_0^t e^{-(k+\lambda)u} \sum_{n=0}^{\infty} p(n; \zeta_1) P(n; \lambda(t-u)) du \\
 &= k \int_0^t e^{-ku} \int_0^{\beta_1} \mu \sum_{n=1}^{\infty} p(n; \lambda u) p(n-1; \mu y) \cdot \\
 &\quad \cdot \sum_{m=0}^{\infty} p(m; \mu(\beta_1 - y)) P(m; \lambda(t-u)) dy du.
 \end{aligned} \tag{13.47}$$

Define the function

$$M^*(i, j; \alpha) = \alpha \int_0^1 e^{-\alpha z} z^{i-1} (1-z)^{j-1} dz. \tag{13.48}$$

With this function we obtain, after some manipulations,

$$\begin{aligned}
 R_1^*(t; k, \beta_0, \beta_1) &= e^{-kt} \sum_{n=0}^{\infty} p(n; \zeta_0) P(n; \lambda t) \\
 &+ \sum_{n=0}^{\infty} p(n; \lambda t) (1 - P(n-1; \zeta_1)) M^*(1, n+1; kt) \\
 &+ \sum_{n=1}^{\infty} p(n; \zeta_1) \sum_{m=0}^{n-1} \sum_{j=0}^m p(n+j; \lambda t) M^*(n-m+1, j+1; kt).
 \end{aligned} \tag{13.49}$$

In Figure 13.4 we present the reliability function $R_1^*(t; k, \beta_0, \beta_1)$ for $k = 0.1$ (points) and $k = 0.5$ (line), $\lambda = \mu = 2$, $\beta_0 = 10$, $\beta_1 = 15$.

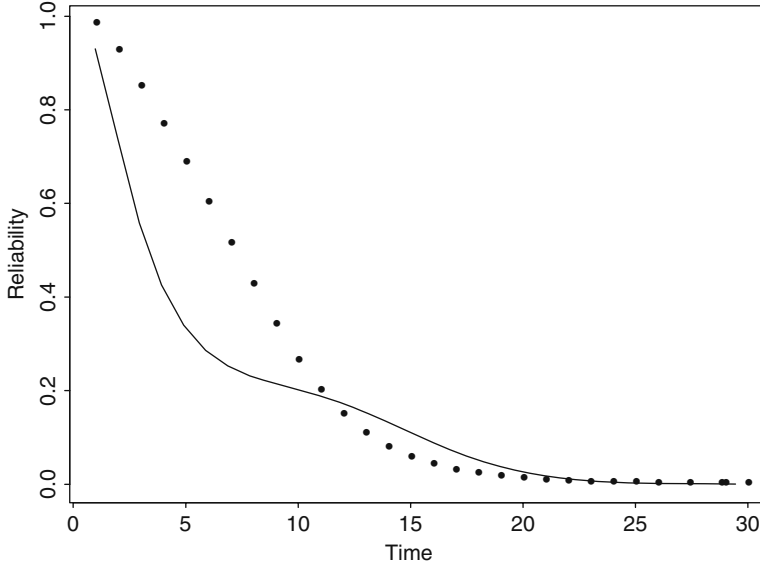


Figure 13.4. The reliability function $R_1^*(t; k, \beta_0, \beta_1)$

13.7.4 Case III, $K = 1$

When $X \sim \exp(\mu)$, the distribution of $R(T_c)$ is $\exp(\mu)$ independent of T_c . Thus, the joint p.d.f. of $(T_c, R(T_c))$ is

$$p(s, r; \gamma_0) = \psi(s; \gamma_0) \mu e^{-\mu r}, \quad 0 < s < \infty, \quad 0 < r < \infty, \quad (13.50)$$

where the p.d.f. of T_c is

$$\psi(s; \gamma_0) = \lambda \sum_{n=0}^{\infty} p(n; \mu \gamma_0) p(n; \lambda s). \quad (13.51)$$

Thus, according to (13.27) the reliability function is

$$\begin{aligned} R_c(t; \gamma_0, \beta_0, \beta_1) &= \sum_{n=0}^{\infty} p(n; \mu \gamma_0) P(n; \lambda t) \\ &+ \int_0^t \psi(s; \gamma_0) \int_0^{\beta_0 - \gamma_0} \mu e^{-\mu r} \sum_{n=0}^{\infty} p(n; \mu \beta_1 - \mu \gamma_0 - \mu r) \cdot \\ &\cdot P(n; \lambda(t - s)) dr ds. \end{aligned} \quad (13.52)$$

Explicitly, we get

$$\begin{aligned}
R_c(t; \gamma_0, \beta_0, \beta_1) &= \sum_{n=0}^{\infty} p(n; \mu\gamma_0) P(n; \lambda t) \\
&+ \lambda t \sum_{l=1}^{\infty} \sum_{m=0}^{l-1} p(l-m; \mu(\beta_0 - \gamma_0)) \left(\frac{\beta_0 - \gamma_0}{\beta_1 - \gamma_1} \right)^{l-m} \cdot \\
&\cdot p(m; \mu\gamma_0) \sum_{j=0}^{l-m} p(m+j; \lambda t) / (m+j+1).
\end{aligned} \tag{13.53}$$

In Figure 13.5 we present $R_c(t; \gamma_0, \beta_0, \beta_1)$ for the values of $\lambda = 1$, $\mu = 2$, $\gamma_0 = 10$, $\beta_0 = 15$, and $\beta_1 = 20$.

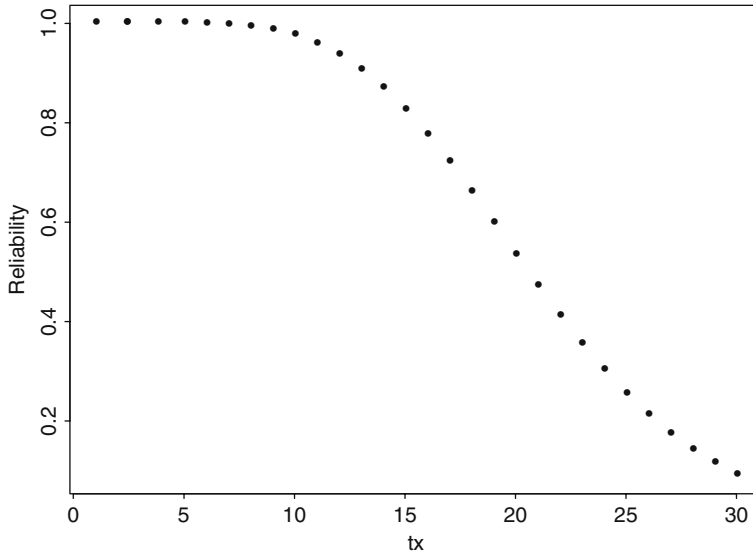


Figure 13.5. The reliability function $R_c(t; \gamma_0, \beta_0, \beta_1)$

13.7.5 Case IV: Linear threshold

According to (13.36)

$$\begin{aligned}
R_L(t; \beta) &= \sum_{n=0}^{\infty} p(n; \mu(\beta + t)) P(n; \lambda t) \\
&- \sum_{n=1}^{\infty} p(n; \lambda t) p(n-1; \mu(\beta + t)) \mu^{n+1} t \cdot \\
&\cdot \int_0^1 \left(1 - \frac{t}{t + \beta} z \right)^{n-1} (1-z)^n e^{(\lambda + \mu)tz} M(tz) dz.
\end{aligned}$$

In Figure 13.6 we present this reliability function for $\lambda = 1$, $\mu = 2$, and $\beta = 10$.

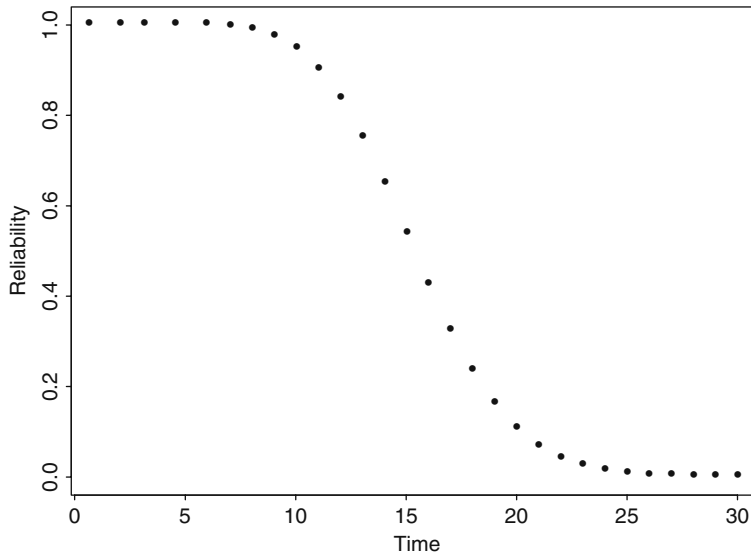


Figure 13.6. The reliability function $R_L(t; \beta)$

References

1. Zacks, S. (2004). Distribution of failure times associated with non-homogeneous compound Poisson damage process. *A Festschrift for Herman Rubin*, Institute of Mathematical Statistics, Lecture Notes-Monograph Series **45**, pp. 396–407.
2. Zacks, S. (2005). Some recent results on the distributions of stopping times of compound Poisson processes with linear boundaries. *Journal of Statistical Planning and Inference*, **130**, 95–109.
3. Zacks, S. (2006). Failure distributions associated with general renewal damage processes. *Probability, Statistics and Modeling in Public Health*, N. Nikulin, D. Commenges and C. Huber (eds.), pp. 466–475, Springer, New York.
4. Zacks, S. (2010). The availability and hazard of a system under a cumulative damage process with replacements. *Methodology and Computing in Applied Probability*, **12**, 1 (to appear).

Failure Time Models Based on Degradation Processes

Axel Lehmann

Hochschule für Technik, Wirtschaft und Kultur Leipzig, Department of Computer Science, Mathematics and Natural Sciences, University of Applied Sciences, Leipzig, Germany

Abstract: Many failure mechanisms can be traced to underlying degradation processes and stochastically changing covariates. Degradation and covariate data can provide much reliability information additional to failure time data and are particularly useful when the application of traditional reliability models is limited due to rare failures of highly reliable items or due to items operating in dynamic environments. This chapter surveys some approaches to model the relationship between failure time data and degradation and covariate data. These models which reflect the dependency between system state and system reliability include threshold models and hazard-based models. In particular, we consider the class of degradation–threshold–shock models in which failure is due to the competing causes of degradation and trauma. For this class of reliability models we compute the survival function of the resulting failure time and derive the likelihood function for the joint observation of failure times and degradation data at discrete times. We consider a special class of degradation–threshold–shock models where degradation is modeled by a process with stationary independent increments and related to external covariates through a random timescale and extend this model class to repairable items by a marked point process approach. The proposed model class provides a rich conceptual framework for the study of degradation–failure issues.

Keywords and phrases: Covariate process, degradation process, degradation–threshold–shock model, failure time, gamma process, Lévy process, marked point process, maximum likelihood estimation, moment estimator, repairable item, semiparametric estimation, timescale, traumatic event, Wiener process

14.1 Introduction

Many failure mechanisms in engineering, medical, social, and economic settings can be traced to an underlying degradation process and stochastically changing covariates that may influence degradation and failure. Most items under study degrade physically over time and a measurable physical deterioration almost always precedes failure.

Examples of items that degrade over time are devices subject to wear in reliability and the immune system of HIV-infected individuals in biostatistics. In an engineering setting, degradation is the irreversible accumulation of damage throughout life that leads to failure (cf. [6]) and may involve corrosion, material fatigue, wearing out, and fracturing. In the biosciences, degradation may be characterized by markers of health status and quality of life data. Frequently, an item is regarded as failed and switched off when degradation first reaches a critical threshold level. Moreover, in most practical applications items or systems operate in heterogeneous environments and loads, environmental stresses, and other dynamically changing environmental factors may influence their failure rate.

Traditionally, statistical reliability estimation is mostly based on failure time data. In practical applications, collecting samples of lifetimes can be time consuming and costly because the systems and devices under study are often highly reliable and expensive. An increasing emphasis on product reliability and short product-development cycles in industry results in few or no failures during reliability tests. Accelerated life tests can be used but the acceleration methods may not properly imitate the actual failure process and, even with acceleration, only few failures may occur. In reliability analysis, biostatistics, and econometrics longitudinal data are frequently available additionally to time-to-event data like failure and survival time data. In many cases, collecting degradation data may be difficult and costly but degradation data can provide much more reliability information than traditional failure time data. When additional longitudinal data are available, an alternative approach to classical reliability and survival analysis is the modeling of degradation and environmental factors and of their failure-generating mechanisms by stochastic processes. This stochastic-process-based approach shows great flexibility and can give rise to new or alternative time-to-failure distributions defined by the degradation model. It provides additional information to failure time observations and is particularly useful when the application of traditional reliability models based on failure and survival data only is limited due to rare failures of highly reliable items or due to items operating in dynamic environments.

14.1.1 Models for degradation data

In engineering applications, degradation data can be obtained by direct measurements of physical degradation such as crack growth and tire wear. If actual physical degradation cannot be observed directly, surrogates like a deteriorating product performance parameter (e.g., power output), usage or exposure measures, and accumulated maintenance costs are usually referred to as degradation data, as well. In the biosciences, degradation data can be measurements of a biomarker like CD4 cell counts or quality of life data. As degradation is a complex random mechanism it is best represented by a stochastic process $X = \{X(t) : t \geq 0\}$. In most applications, the monotonicity of degradation paths is a reasonable assumption because most real degradation processes are increasing. However, there are situations in which degradation need not be increasing. For instance, fatigue cracks sometimes show a tendency to heal and CD4 cell counts may fluctuate. A further cause for non-monotone degradation paths is the fact that degradation is measured with error and the missing monotonicity is due to the error term whereas the actual degradation is monotone. Degradation processes are frequently described by general path models and by processes with stationary and independent increments. More sophisticated models for degradation processes are Markov diffusion

processes (e.g., the Ornstein–Uhlenbeck process), semi-Markov processes, and marked point processes (cf. [30]).

General path models

In applied literature degradation processes are frequently described by a general path model $X(t) = g(t, A)$, i.e., by a stochastic process that depends on a finite dimensional random variable $A = (A_1, \dots, A_r)$ via a deterministic function g . In [24, 25, 8], a Paris growth curve path model was used for alloy fatigue crack size data. Moreover, general path degradation models with linear, convex, and concave functions g were applied to tire wear, electronic component degradation, silicon solar cell degradation, and disk storage media (cf. [7, 24–26]). General path degradation models have been considered in more detail by Meeker and Escobar [25].

Lévy process models

A Lévy process is a stochastic process with stationary and independent increments. The application of such processes as degradation models can be based on the fact that in systems subject to shocks, the order of damage causing shocks is often irrelevant. So, the increments of the accumulated damage process are stationary and exchangeable random variables which is somewhat weaker than stationary and independent increments. The most widely used Lévy processes for degradation modeling are compound Poisson processes, the Wiener process with drift, and the gamma process. When degradation evolves as an accumulation of shocks, each of them resulting in a random degradation increment, a compound Poisson process is an appropriate model. It is given by $X(t) = \sum_{i=1}^{N(t)} X_i$ where N is a Poisson counting process jumping at the times of shocks and $\{X_1, X_2, \dots\}$ is a sequence of i.i.d. random variables that represent the random increments of degradation caused by the shocks and that are independent of N . The Wiener process with drift is a Gaussian process given by $X(t) = x_0 + \mu t + \sigma W(t)$ where W denotes a standard Brownian motion, x_0 is some initial degradation level, and μ and σ are the drift and the variance coefficient, respectively. The Wiener process is the only Lévy process with continuous sample paths; however, its paths are not monotone. The gamma process has increasing sample paths but it is a pure jump process. Simulated data of a Wiener process and a gamma process DTS model with the drift and the variance parameter of the simulated processes being equal to 1 are shown in Figures 14.1 and 14.2.

A Lévy process Y has linear mean and variance functions. To cover non-linear degradation behavior a timescale $\tau = \tau(t)$ can be used, i.e., a positive, increasing function of real time with $\tau(0) = 0$. The timescale drives the speed of deterioration and describes slowing or accelerating degradation in real time. It measures the physical progress of degradation and is referred to as operational time. In Section 14.3 we consider several examples of timescales.

To model the influence on failure of an item's dynamic operating environment, the timescale may depend on a covariate process Z , for instance on different loads, stress levels, or usage measures, i.e., $\tau_Z(t) = \tau(t, Z)$. An example is the model of additive accumulation of damage:

$$\tau_Z(t) = \tau \left(\int_0^t \exp(\beta^T Z(s)) ds \right).$$

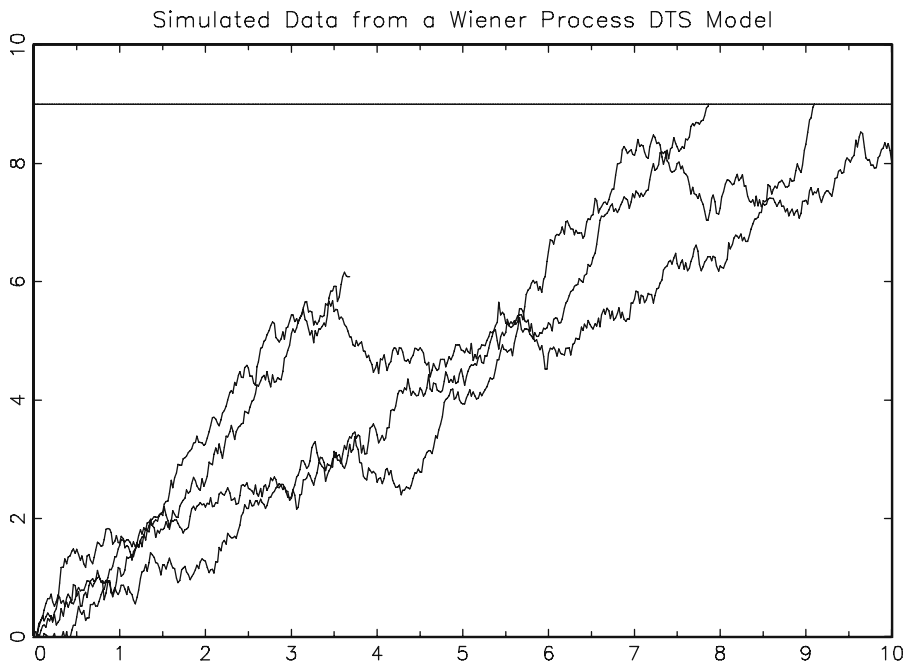


Figure 14.1. Wiener process degradation model

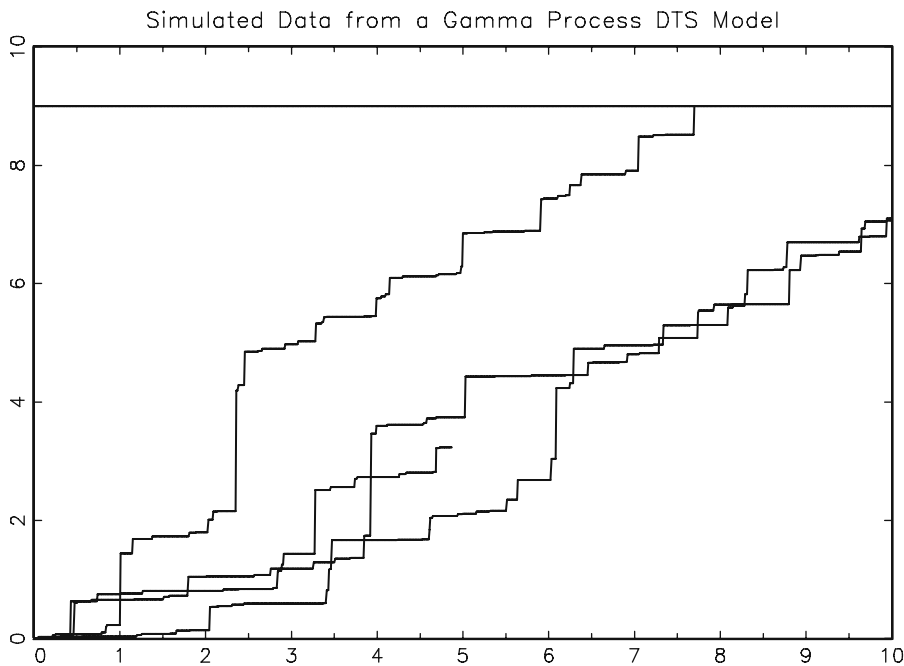


Figure 14.2. Gamma process degradation model

The degradation model is given by

$$X(t) = X_0 + Y(\tau_Z(t)) =: X_0 + Y^Z(t),$$

where Y^Z is the time-scaled Lévy process Y and X_0 is some possibly random initial degradation level independent of (Y, Z) . Conditioned on Z , X is a process with independent but not stationary increments. For example, let the baseline degradation process Y be a Wiener process with drift $Y(t) = \mu t + \sigma W(t)$ so that the time-scaled degradation process X is given by $X(t) = X_0 + \mu \tau_Z(t) + \sigma W(\tau_Z(t))$. This model is appropriate if the actual degradation process $\mu \tau_Z(t)$ is latent but observable with some additive error term which has a variance proportional to the actual degradation process.

Wiener process degradation models have found application in [10, 11]. Whitmore [36] applied a Wiener process model with error to degradation data of transistor gain. Wiener processes with a non-random timescale were used as degradation models for heating cables and carbon-film resistors in [37, 27]. A bivariate Wiener process degradation model was considered in [38] and in [18] and applied to equipment degradation data in aluminum production and to biomarker data of AIDS clinical trials. In these models, one process is a latent or unobservable degradation process and the second one represents a marker process that is correlated with the degradation process and tracks its progress.

Degradation models based on the gamma process were considered in [35, 30] and with application to crack growth data in [17]. In [3] a gamma process with a random timescale was applied to tire wear, and in [28], both geometric Brownian motion and gamma processes were considered in a time-scaled degradation model.

14.1.2 Models relating degradation and failure

Considerable interest has focused on joint models for the distribution of the failure time T and some related covariate process X of internal degradation or external environmental data. Two model classes, threshold models and shock models, predominate in the engineering context (cf. [9]). In a threshold model, X is *failure defining* by $T = \inf\{t \geq 0 : X(t) \geq X^*\}$, where X^* is a critical threshold level which is possibly random. This kind of failure, where an item is regarded as failed and usually switched off when its degradation level first reaches a critical threshold, is also called *soft failure* (cf. [25]) or *nontraumatic failure*. In a shock model, the item or system is subjected to external shocks which may be survived or which lead to a failure that is referred to as *hard failure* or *traumatic failure*. The shocks usually occur according to a Poisson process whose intensity depends on degradation and environmental factors. In a shock model X is *failure rate determining*, which means that the conditional failure rate of T

$$\lambda(t, X_0^t) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} P(t \leq T < t + \Delta \mid X_0^t, T \geq t) \quad (14.1)$$

as a function of t and the path $X_0^t = \{X(s) : 0 \leq s \leq t\}$ of the covariate process up to time t determine the corresponding conditional survival function via the *exponential formula*

$$P(T > t \mid X_0^t) = \exp\left(-\int_0^t \lambda(s, X_0^s) ds\right). \quad (14.2)$$

However, Yashin and Arjas [40] pointed out that this formula does not necessarily hold without some further conditions on the behavior of the conditional distribution function

$F^X(t) := P(T \leq t | X_0^t)$ of the failure time T . The problem, whether the exponential formula (14.2) still holds in the presence of time-dependent stochastic covariates, was addressed by Flournoy [13] and by Kalbfleisch and Prentice [15] and considered on a descriptive level with interesting examples in [31, 12]. It was treated on a technical level using the martingale approach and partly answered by Yashin and Arjas [40], Aven and Jensen [1], and Lehmann [21]. In general F^X is a submartingale (with respect to the filtration induced by X) and its paths may be rather irregular, i.e. they need not be either monotone or continuous. Yashin and Arjas [40] showed that the extended exponential formula

$$P(T > t | X_0^t) = \exp(-\Lambda^c(t)) \prod_{s \leq t} (1 - \Delta\Lambda(s)), \quad (14.3)$$

where Λ denotes the integrated conditional failure rate of T with continuous part Λ^c and discrete part $\Delta\Lambda$, holds if F^X is increasing and predictable, and that the exponential formula (14.2) holds if F^X is increasing and absolutely continuous. Sufficient conditions for the monotonicity, predictability, and continuity of F^X were given by Lehmann [21]. These conditions are closely related to the notion of external and strongly external covariates, the latter being defined in [21]. In the following section we consider a class of reliability models which unifies the approaches of hard and soft failures.

14.2 Degradation–Threshold–Shock Models

14.2.1 The general model

Two relevant stochastic models relating failure to degradation and other covariates have evolved in the theoretical and applied literature, *threshold models* and *shock models*. A threshold model supposes that the item or system fails whenever its degradation level reaches a certain critical deterministic or random threshold. In a shock model the item or system is subjected to external shocks which may be survived or which lead to failure. The shocks usually occur according to a Poisson process whose intensity depends on degradation and environmental factors. It appears that Lemoine and Wenocur [23] may have been the first to combine both approaches by considering two competing causes of failure: degradation reaching a threshold, and occurrence of a traumatic event like a shock of large magnitude severe enough to destroy the item. So, the failure time of an item is the minimum of the moment when degradation first reaches a critical threshold and the moment when a censoring traumatic event occurs.

We call this class of reliability models, which combines the ideas of hard and soft failures by considering them as competing causes of failure, *degradation–threshold–shock models* (DTS models). Suppose that the degradation level of some item is described by a stochastic process X . An item is regarded as failed when the degradation process reaches a critical threshold level X^* which is possibly random but independent of X . This time is called the *nontraumatic* failure time $D = \inf\{t \geq 0 : X(t) \geq X^*\}$ caused by degradation. Additionally to such soft failures which are immediately related to degradation, an item may also fail when a traumatic event like a shock of large magnitude occurs although the degradation process has not yet reached the threshold. Such

traumatic failure time C is assumed to be the first point of a doubly stochastic Poisson process $\Psi = \{\Psi(t) : t \in \mathbb{R}_+\}$ with a stochastic intensity $\kappa(t, X(t))$ that may depend on time t and on the degradation level $X(t)$. Thus, $C = \inf\{t \geq 0 : \Psi(t) = 1\}$ has the stochastic failure rate $\kappa(t, X(t))$ and a survival function that satisfies, conditioned on the degradation path $X_0^t = \{X(s) : 0 \leq s \leq t\}$, the exponential formula

$$P(C > t | X_0^t) = \exp\left(-\int_0^t \kappa(s, X(s)) ds\right). \quad (14.4)$$

The survival function of the observable failure time

$$T = \min(D, C)$$

is given by the expectation

$$P(T > t) = E\left[\mathbf{1}(D > t) \exp\left(-\int_0^t \kappa(s, X(s)) ds\right)\right],$$

where $\mathbf{1}(D > t)$ denotes the indicator function of the event $\{D > t\}$. To find an expression of the survival function and the failure rate of T in a DTS model we use a theorem given in [41]:

Theorem 1 (Yashin, Manton). *Let ζ and ξ be stochastic processes influencing a failure rate $\alpha(t, \zeta, \xi)$ and satisfying measurability conditions such that, for $t \geq 0$*

$$E \int_0^t \alpha(u, \zeta, \xi) du < \infty,$$

and let T be related to ζ_0^t and ξ_0^t by

$$P(T > t | \zeta_0^t, \xi_0^t) = \exp\left(-\int_0^t \alpha(u, \zeta, \xi) du\right). \quad (14.5)$$

If the trajectories of ζ are observed up to t , then

$$P(T > t | \zeta_0^t) = \exp\left(-\int_0^t \bar{\alpha}(u, \zeta_0^u) du\right),$$

where

$$\bar{\alpha}(t, \zeta_0^t) = E[\alpha(t, \zeta, \xi) | \zeta_0^t, T > t].$$

The random failure rate $\alpha(t, \zeta, \xi)$ may depend on either the current values $\zeta(t)$ and $\xi(t)$ or on the trajectories ζ_0^t and ξ_0^t up to t . If the covariate process ζ does not appear in (14.5), i.e., $\alpha = \alpha(t, \xi)$, the statement of Theorem 1 obviously reads

$$P(T > t) = \exp\left(-\int_0^t E[\alpha(u, \xi) | T > u] du\right) \quad (14.6)$$

(see [39]). Although we observe that

$$P(T > t | X_0^t) = E[\mathbf{1}(D > t) \mathbf{1}(C > t) | X_0^t] = \mathbf{1}(D > t) P(C > t | X_0^t)$$

is not of the form (14.5), Theorem 1 can be used to compute $P(T > t)$. The proof of the following theorem was given in [22].

Theorem 2. *Let the traumatic failure time C have a stochastic failure rate $\kappa(t, X(t))$ satisfying $\mathbb{E} \int_0^t \kappa(s, X(s)) ds < \infty$ for all $t \geq 0$. Let X^* be independent of X and assume that, given $X^* = x^*$, the nontraumatic failure time D has the conditional failure rate $\lambda(t, x^*)$ with $\mathbb{E} \int_0^t \lambda(s, X^*) ds < \infty$ for all $t \geq 0$. Then*

$$P(C > t) = \exp \left(- \int_0^t \check{\kappa}(s) ds \right),$$

and

$$P(D > t) = \exp \left(- \int_0^t \bar{\lambda}(s) ds \right),$$

where the failure rates $\check{\kappa}$ and $\bar{\lambda}$ are given by $\check{\kappa}(t) = \mathbb{E}[\kappa(t, X(t)) | C > t]$ and $\bar{\lambda}(t) = \mathbb{E}[\lambda(t, X^*) | D > t]$. The survival function of T can be expressed as

$$P(T > t) = \exp \left(- \int_0^t (\bar{\kappa}(s) + \bar{\lambda}(s)) ds \right),$$

where $\bar{\kappa}(t) = \mathbb{E}[\kappa(t, X(t)) | T > t]$ is the failure rate of a traumatic event if a nontraumatic event has not yet occurred.

If D has a failure rate $\lambda(t)$ which is independent of X^* and if C has a deterministic failure rate $\kappa(t)$, then the survival function of T simplifies to

$$P(T > t) = \exp \left(- \int_0^t (\kappa(s) + \lambda(s)) ds \right).$$

The class of DTS models contains two important subclasses, *degradation-threshold models* (DT models) and *degradation-shock models* (DS models).

14.2.2 Degradation-threshold models

In a degradation-threshold model only nontraumatic failures can occur, i.e., the traumatic event intensity κ is equal to zero and the failure time $T = D = \inf\{t \geq 0 : X(t) \geq X^*\}$ is the first passage time of X to the threshold X^* . A simple example with a nonrandom threshold x^* is a linear path model

$$X(t) = x + \Theta t,$$

with random slope Θ and initial degradation level $x < x^*$. If Θ has a reciprocal Weibull distribution, i.e. $\Theta^{-1} \sim \text{Wei}(\beta, \theta)$, then $T \sim \text{Wei}(\beta, \theta(x^* - x))$ is *Weibull* distributed. If degradation is modeled by a one-dimensional Wiener process with drift

$$X(t) = x + \mu t + \sigma W(t),$$

where W denotes a standard Brownian motion and if $x^* > x$ is non-random, then it is well known that $T \sim \text{IG} \left(\frac{x^* - x}{\mu}, \frac{(x^* - x)^2}{\sigma^2} \right)$ is *Inverse Gaussian*. In general, T has an upside bathtub failure rate, but it has an essentially increasing failure rate (IFR) if $(x^* - x)/\sigma \gg 1$ and a decreasing failure rate (DRF) if $(x^* - x)/\sigma \ll 1$ (see Figure 14.3).

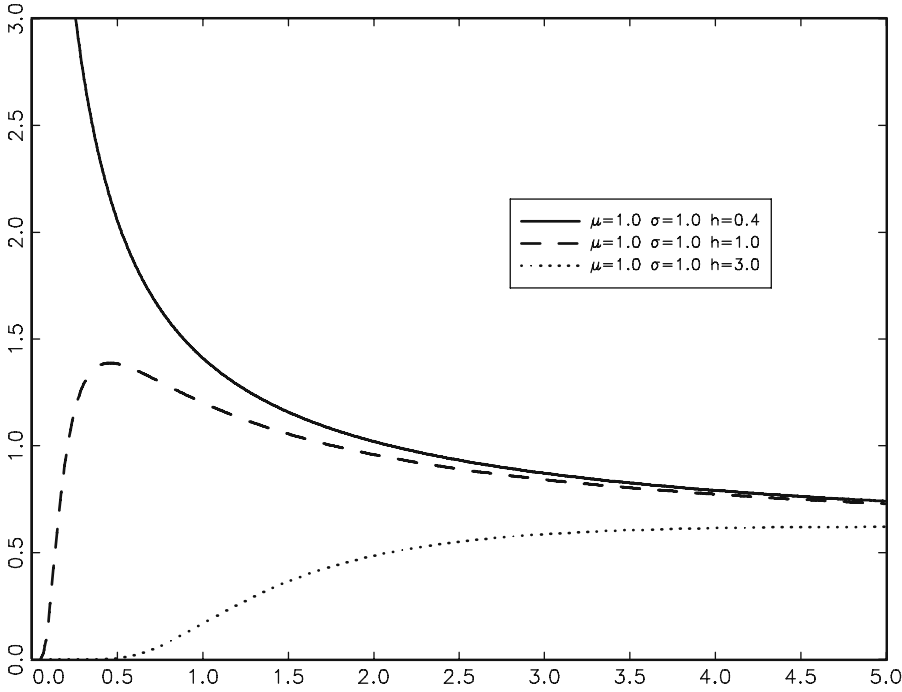


Figure 14.3. Failure rates of Inverse Gaussian distribution ($x^* = h$)

By total probability it can be easily shown that $T \sim \text{Ga}\left(\frac{\alpha}{2}, \frac{\mu^2}{2\sigma^2}\right)$ is *gamma* distributed for $x = 0$ and a gamma distributed threshold $X^* \sim \text{Ga}\left(\alpha, \frac{\mu}{\sigma^2}\right)$ being independent of W .

For an increasing degradation process X the survival function of T is given by $P(T > t) = P(X(t) < x^*)$. If, for instance, X is a homogeneous Poisson process with rate $\lambda > 0$, then $T \sim \text{Ga}([x^*], \lambda)$ follows a *gamma* distribution, which has an increasing failure rate. If X is an increasing jump process, then T has always increasing failure rate average (cf. [29]):

Theorem 3 (Shaked, Shantikumar). *If X is an increasing jump process, then T has increasing failure rate average (IFRA). If X is an increasing Lévy process with a Lévy measure ν which has a decreasing density, then T has increasing failure rate (IFR).*

14.2.3 Degradation–shock models

The class of degradation–shock models is characterized by the absence of a critical threshold X^* which can be described formally by $X^* = \infty$. Here failure is defined when the item stops working and the failure time is given by $T = C$, i.e., only traumatic failures can occur. Hence, by Theorem 2 we have

$$\begin{aligned} P(T > t) &= E \left[\exp \left(- \int_0^t \kappa(s, X(s)) \, ds \right) \right] \\ &= \exp \left(- \int_0^t \bar{\kappa}(s) \, ds \right), \end{aligned}$$

with $\bar{\kappa}(t) = \check{\kappa}(t) = \mathbb{E}[\kappa(s, X(s)) | C > t]$. For a positive increasing Lévy degradation process X with Lévy measure ν and drift rate μ , Kebir [16] proved

$$\begin{aligned} \mathbb{P}(T > t) &= \mathbb{E} \left[\exp \left(- \int_0^t X(s) \, ds \right) \right] \\ &= \exp \left(- \frac{\mu t^2}{2} - \int_0^t \int_0^\infty [1 - e^{-sx}] \nu(dx) \, ds \right), \end{aligned}$$

i.e., T has the increasing failure rate $\bar{\kappa}(t) = \mu t + \int_0^\infty [1 - e^{-tx}] \nu(dx)$. Note that for $\gamma > 0$ the process γX is a positive increasing Lévy process as well with drift rate $\gamma\mu$ and Lévy measure $\nu(dx/\gamma)$. Applying Kebir's formula to an intensity $\kappa(t, X(t)) = \gamma X(t)$ that depends proportionally on degradation and to a homogeneous Poisson process X with rate $\lambda > 0$, which has a drift rate $\mu = 0$ and a Lévy measure $\nu = \lambda \varepsilon_1$, where ε_x denotes the Dirac measure at x , we see that T follows a *Makeham* distribution with the survival function

$$\mathbb{P}(T > t) = \exp(-\lambda\gamma t + \lambda(1 - e^{-\gamma t}))$$

and the failure rate $\bar{\kappa}(t) = \lambda\gamma(1 - e^{-\gamma t})$.

For a DS model with degradation modeled by a Wiener process with drift $X(t) = x + \mu t + \frac{\sigma}{\sqrt{2}} W(t)$ and a quadratic intensity $\kappa(t, X(t)) = (X(t))^2$, Wenocur [34] computed the survival function of T as

$$\begin{aligned} \mathbb{P}(T > t) &= \mathbb{E} \left[\exp \left(- \int_0^t (X(s))^2 \, ds \right) \right] \\ &= \exp \left\{ - \frac{\mu^2}{\sigma^2} t + \left(\frac{\mu^2}{\sigma^3} - \frac{x^2}{\sigma} \right) \tanh(\sigma t) + 2 \frac{x\mu}{\sigma^2} (\operatorname{sech}(\sigma t) - 1) \right\} \sqrt{\operatorname{sech}(\sigma t)}. \end{aligned} \quad (14.7)$$

Some failure rates of T are shown in Figure 14.4. If σ tends to zero the distribution of T converges to a “generalized” *Weibull* distribution with form parameter three:

$$\mathbb{P}(T > t) \xrightarrow{\sigma \downarrow 0} \exp \left(-(\mu^2/3)t^3 - x\mu t^2 - x^2 t \right).$$

14.2.4 Likelihood function

A crucial result for the derivation of the likelihood function in a DTS model is an expression of the joint density of degradation levels $X(t_1), \dots, X(t_k)$ at planned inspection times t_1, \dots, t_k conditioned on the event that no failure has occurred up to some moment $t \geq t_k$. To derive such an expression we assume that $X(t)$ possesses a density $f_{X(t)}$ with respect to some dominating measure ν , usually the Lebesgue measure or the counting measure. However, in the following we will write dx instead of $\nu(dx)$ regardless of the nature of ν . Further, we assume that for all $t \geq 0$ and $\mathbf{t}_k = (t_0, \dots, t_k) \in \mathbb{R}^{k+1}$ with $0 \leq t_0 < \dots < t_k \leq t$ and $\mathbf{x}_k = (x_0, \dots, x_k) \in \mathbb{R}^{k+1}$ the conditional joint density $g(t, \mathbf{t}_k, \mathbf{x}_k; x^*)$ with

$$\mathbb{P}(D > t, \mathbf{X}_k \in d\mathbf{x}_k | X^* = x^*) = g(t, \mathbf{t}_k, \mathbf{x}_k; x^*) d\mathbf{x}_k \quad (14.8)$$

and $\mathbf{X}_k = (X(t_0), \dots, X(t_k))$ is known. Of course, g must satisfy $g(t, \mathbf{t}_k, \mathbf{x}_k; x^*) = 0$ if $\min(x_0, \dots, x_k) \geq x^*$. If the paths of X are increasing, then obviously

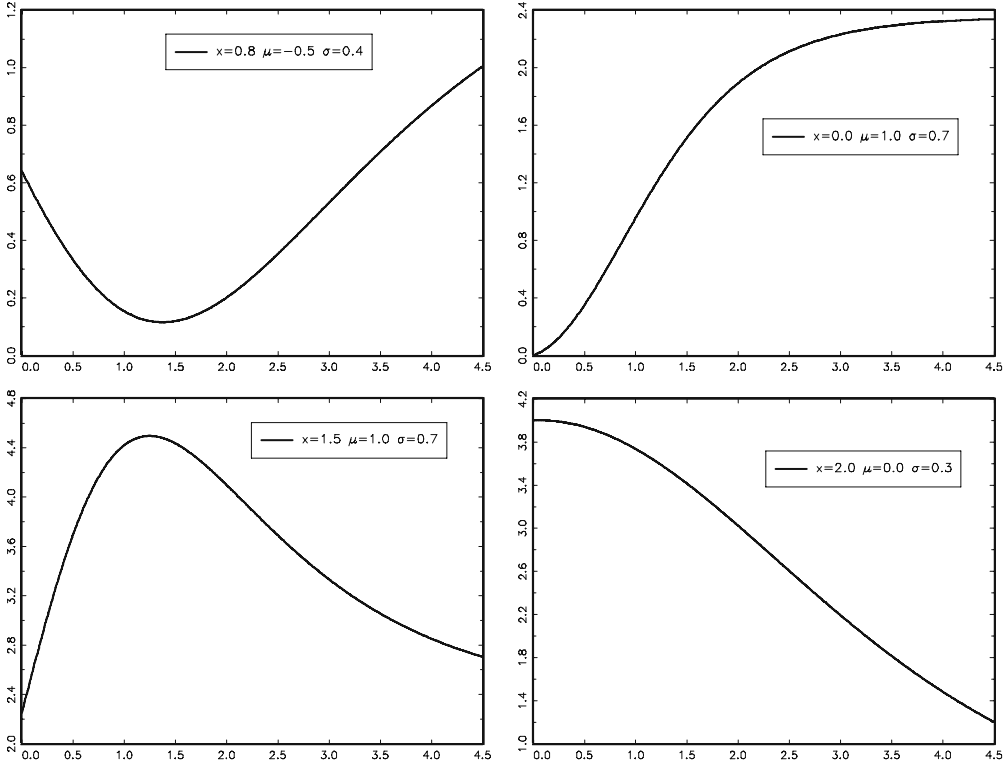


Figure 14.4. Failure rates of the distribution (14.7)

$$g(t, \mathbf{t}_k, \mathbf{x}_k; x^*) = f_{\mathbf{X}_k}(\mathbf{x}_k) P(X(t) < x^* | \mathbf{X}_k = \mathbf{x}_k)$$

for $x_0 \leq \dots \leq x_k < x^*$ and $g(t, \mathbf{t}_k, \mathbf{x}_k; x^*) = 0$ otherwise.

Modeling degradation by a Wiener process with drift $X(t) = x + \mu t + \sigma W(t)$ with drift coefficient μ , variance coefficient σ , and initial value $X(0) = 0$, we find

$$g(t, \mathbf{t}_k, \mathbf{x}_k; x^*) = \prod_{j=1}^k g_0(t_j - t_{j-1}, x_j - x_{j-1}; x^*) \bar{F}_0(t - t_k, x^* - x_k)$$

for $\max(x_1, \dots, x_k) < x^*$, where $t_0 = x_0 = 0$ and

$$g_0(t, x; x^*) = \mathbf{1}(x^* > x) \frac{1}{\sigma\sqrt{t}} \phi\left(\frac{x - \mu t}{\sigma\sqrt{t}}\right) \left[1 - \exp\left(-\frac{2x^*(x^* - x)}{\sigma^2 t}\right)\right]$$

for all $t > 0$ (cf. [14]). The function \bar{F}_0 is the survival function of an *Inverse Gaussian* distribution, i.e.,

$$\bar{F}_0(t, x) = N\left[\frac{x - \mu t}{\sigma\sqrt{t}}\right] - e^{(2\mu\sigma^{-2}x)} N\left[\frac{-x - \mu t}{\sigma\sqrt{t}}\right],$$

where ϕ and N denote the pdf and the cdf of a standard normal random variable. The proofs of Theorems 4 and 5 are given in [22]:

Theorem 4. Let the traumatic failure time C be given by (14.4) with a stochastic failure rate $\kappa(t, X(t))$ satisfying $\mathbb{E} \int_0^t \kappa(s, X(s)) \, ds < \infty$ for all $t \geq 0$. Let X^* be independent of X and assume (14.8) for all $t \geq 0$, $\mathbf{t}_k = (t_0, \dots, t_k) \in \mathbb{R}^{k+1}$ with $0 \leq t_0 < \dots < t_k \leq t$ and $\mathbf{x}_k = (x_0, \dots, x_k) \in \mathbb{R}^{k+1}$. Then

$$\mathbb{P}(T > t, \mathbf{X}_k \in d\mathbf{x}_k) = \exp\left(-\int_0^t \bar{\kappa}(s, \mathbf{x}_{k(s)}) \, ds\right) g(t, \mathbf{t}_k, \mathbf{x}_k) \, d\mathbf{x}_k,$$

where

$$g(t, \mathbf{t}_k, \mathbf{x}_k) = \mathbb{E}[g(t, \mathbf{t}_k, \mathbf{x}_k; X^*)]$$

and

$$\bar{\kappa}(s, \mathbf{x}_{k(s)}) = \mathbb{E}[\kappa(s, X(s)) \mid T > s, \mathbf{X}_{k(s)} = \mathbf{x}_{k(s)}]$$

with $k(s) = \max\{j \geq 0 : t_j \leq s\}$ for $0 \leq s \leq t$.

The following theorem gives the likelihood function in a DTS model for the case of n independent items with identically distributed degradation processes X_i and stochastic failure rates $\kappa(t, X_i(t))$, based on the observation of their degradation levels $X_i(t_{ij})$, their failure times T_i , and their failure modes (hard or soft failure).

Suppose that n independent items are observed in $[0, t^*]$ with identically distributed degradation processes X_i ; failure times D_i, C_i, T_i ; and thresholds X_i^* . We assume that X_i^* and X_i are independent for all $i = 1, \dots, n$. The i th item is observed at planned inspection times $0 \leq t_{i0} < t_{i1} < \dots$ until t^* . Let $x_{ij} = X_i(t_{ij})$ denote the observed degradation levels. If a failure occurs in $[0, t^*]$, the observation of X_i will be stopped after this event. That means, in each interval $(t_{ij-1}, t_{ij}]$ we observe either a failure at $t_i \in (t_{ij-1}, t_{ij}]$ and possibly the degradation level $X_i(t_i)$ or we observe the degradation level $x_{ij} = X_i(t_{ij})$ at t_{ij} under the condition that degradation has not yet exceeded the threshold. For the i th item let

$$l_i = l_i(t^*) = \max\{j \geq 0 : t_{ij} < \min(T_i, t^*)\},$$

i.e., $l_i + 1$ is the number of observed degradation levels without failure in $[0, t^*)$. Further, let $\tilde{T}_i = \min(T_i, t^*)$ be the observable censored failure time and

$$V_i = V_i(t^*) = \begin{cases} 0, & \text{if } T_i > t^* \\ 1, & \text{if } D_i < C_i, D_i \leq t^* \\ -1, & \text{if } C_i \leq \min(D_i, t^*) \end{cases},$$

an observable failure mode indicator. The values $V_i = 1$ and $V_i = -1$ indicate a nontraumatic failure and a traumatic failure in $[0, t^*]$, respectively, whereas the value $V_i = 0$ indicates that no failure has occurred in $[0, t^*]$. The degradation level $x_i = X_i(\tilde{t}_i)$ at the observed censored failure time \tilde{t}_i is assumed to be observable in case of a nontraumatic failure but may be observable or not observable in all other cases. Set $k_i = l_i + 1$, $t_{ik_i} = \tilde{t}_i$, and $x_{ik_i} = x_i$ if $X_i(\tilde{t}_i)$ is observed and $k_i = l_i$ otherwise. Hence the data for the i th item in $[0, t^*]$ is

$$(\tilde{t}_i = \tilde{T}_i, v_i = V_i, \mathbf{x}_{ik} = \mathbf{X}_{ik})$$

with $\mathbf{X}_{ik} = (X_i(t_{i0}), \dots, X_i(t_{ik_i}))$. By $f_D(t \mid \mathbf{t}_l, \mathbf{x}_l; x^*)$ we denote the conditional density of the nontraumatic failure time given $\{X^* = x^*\}$ and given $l + 1$ observations of the degradation process X without reaching the threshold up to $t_l < t$:

$$f_D(t | \mathbf{t}_l, \mathbf{x}_l; x^*) dt = P(D \in dt | D > t_l, \mathbf{X}_l = \mathbf{x}_l; X^* = x^*)$$

and by f_{X^*} the density of the random threshold X^* .

Theorem 5. *The likelihood function for the above observation of n independent items in $[0, t^*]$ is given by*

$$\begin{aligned} L_{t^*}(\tilde{t}_i, v_i, \mathbf{X}_{ik}) &= \prod_{i=1}^n \left\{ \left(g(\tilde{t}_i, \mathbf{t}_{ik}, \mathbf{x}_{ik}) \right)^{\mathbf{1}(v_i < 1)} \right. \\ &\quad \times \left(f_D(\tilde{t}_i | \mathbf{t}_{il}, \mathbf{x}_{il}; x_i) g(t_{il_i}, \mathbf{t}_{il}, \mathbf{x}_{il}; x_i) f_{X^*}(x_i) \right)^{\mathbf{1}(v_i = 1)} \\ &\quad \left. \times \bar{\kappa}(\tilde{t}_i, \mathbf{x}_{ik})^{\mathbf{1}(v_i = -1)} \exp \left(- \int_0^{\tilde{t}_i} \bar{\kappa}(s, \mathbf{x}_{ik(s)}) ds \right) \right\}, \end{aligned} \quad (14.9)$$

$$\begin{aligned} \text{where} \quad f_D(t | \mathbf{t}_l, \mathbf{x}_l; x^*) dt &= P(D \in dt | D > t_l, \mathbf{X}_l = \mathbf{x}_l; X^* = x^*), \\ g(t, \mathbf{t}_k, \mathbf{x}_k; x^*) d\mathbf{x}_k &= P(D > t, \mathbf{X}_k \in d\mathbf{x}_k | X^* = x^*), \\ g(t, \mathbf{t}_k, \mathbf{x}_k) &= E[g(t, \mathbf{t}_k, \mathbf{x}_k; X^*)], \\ f_{X^*}(x) dx &= P(X^* \in dx), \\ \bar{\kappa}(s, \mathbf{x}_{k(s)}) &= E[\kappa(s, X(s)) | T > s, \mathbf{X}_{k(s)} = \mathbf{x}_{k(s)}], \end{aligned}$$

with $k(s) = \max\{j \geq 0 : t_j \leq s\}$.

Based on this complex likelihood structure maximum likelihood estimators of model parameters have to be found numerically in general. Explicit estimators of the degradation parameters in a special DT model based on the Wiener process were given in [19]. Frequently, the conditional failure rate $\bar{\kappa}$ cannot be obtained in closed form and the evaluation of the integral in the likelihood function (14.9) is computationally difficult. Using a simple numerical integration scheme based on the trapezoidal rule

$$\begin{aligned} \int_0^{\tilde{t}_i} \bar{\kappa}(s, \mathbf{x}_{ik(s)}) ds &\approx \sum_{j=1}^{k_i} \frac{1}{2} (\bar{\kappa}(t_{ij-1}, \mathbf{x}_{ij-1}) + \bar{\kappa}(t_{ij}, \mathbf{x}_{ij})) (t_{ij} - t_{ij-1}) \\ &= \sum_{j=1}^{k_i} \frac{1}{2} (\kappa(t_{ij-1}, x_{ij-1}) + \kappa(t_{ij}, x_{ij})) (t_{ij} - t_{ij-1}), \end{aligned}$$

with $x_{ij} = X_i(t_{ij})$, we get even rid of the cumbersome functions $\bar{\kappa}$, since at each t_{ij} we have

$$\bar{\kappa}(t_{ij}, \mathbf{x}_{ij}) = E[\kappa(t_{ij}, X_i(t_{ij})) | T_i > t_{ij}, \mathbf{X}_{ij} = \mathbf{x}_{ij}] = \kappa(t_{ij}, X_i(t_{ij})).$$

By the same argument we find $\bar{\kappa}(\tilde{t}_i, \mathbf{x}_{ik}) = \kappa(\tilde{t}_i, X_i(\tilde{t}_i))$ if the level $X_i(\tilde{t}_i)$ is observed. Thus, the approximated likelihood function is separated into two parts which may be maximized independently of each other: the first two lines of the right-hand side of (14.9) which depend only on the parameters of the degradation process and the threshold distribution and the third line of (14.9) which depends only on the parameters of the stochastic failure rate κ .

14.2.5 Estimation of the survival function and the failure rate of T

Let the intensity $\kappa(t, X(t); \theta)$ be parameterized by some parameter θ and let $\hat{\theta}$ be the ML estimate of θ . Further, we assume that the degradation levels of all items are observed at the same inspection times $t_{ij} = t_j$ ($i = 1, \dots, n$). Both the survival function of the observable failure time T

$$P(T > t; \theta) = E \left[\mathbf{1}(D > t) \exp \left(- \int_0^t \kappa(s, X(s); \theta) ds \right) \right]$$

and the failure subrate

$$\bar{\kappa}(t; \theta) = E[\kappa(t, X(t); \theta) | T > t]$$

can be expressed as expectations and estimated at the inspection times t_j by the following moment estimators:

$$\begin{aligned} P(\widehat{T > t_j}; \theta) &= \frac{1}{\#\{i : C_i > t_j\}} \sum_{i: T_i > t_j} \exp \left(- \int_0^{t_j} \kappa(s, X(s); \hat{\theta}) ds \right) \\ \bar{\kappa}(\widehat{t_j}; \theta) &= \frac{1}{\#\{i : T_i > t_j\}} \sum_{i: T_i > t_j} \kappa(t_j, X_i(t_j); \hat{\theta}) \end{aligned}$$

14.3 DTS Models with Covariates

Additionally to the degradation process one can consider an external covariate process $Z = \{Z(t) : t \in \mathbb{R}_+\}$ which describes the dynamic environment and may influence degradation and the intensity κ of traumatic events. An external covariate is a covariate for which the history of an item is determined conditionally independently of the failure time process (cf. [15], Section 5.3, [21]). In this sense, the degradation process X in a DTS model is not an external covariate. Since such covariate processes like loads, stresses, or usage measures can often be completely observed, one is interested in the conditional distribution of degradation and failure time given the covariate history $Z_0^t = \{Z(s) : 0 \leq s \leq t\}$ up to some time t . If the failure rate $\lambda(t, Z_0^t, X^*)$ of D depends on the covariate Z and the threshold X^* and if the intensity of traumatic events $\kappa(t, Z(t), X(t))$ depends on the environment and on the degradation level, the survival function of T given in Theorem 2, but now conditioned on Z_0^t , is given by

$$P(T > t | Z_0^t) = \exp \left(- \int_0^t (\bar{\kappa}(s, Z_0^s) + \bar{\lambda}(s, Z_0^s)) ds \right),$$

with conditional failure rates $\bar{\kappa}(t, Z_0^t) = E[\kappa(t, Z(t), X(t)) | Z_0^t, T > t]$, and $\bar{\lambda}(t, Z_0^t) = E[\lambda(t, Z_0^t, X^*) | Z_0^t, D > t]$.

In the following DTS model degradation is modeled by a *Lévy process* with a random timescale which depends on the environmental covariate Z . To be more precise, let Y be a process with stationary independent increments, e.g., a *Wiener process*, a *gamma process*, or a *Poisson process*. Then, Y has linear mean and variance functions $E[Y(t)] = \mu t$ and $\text{Var}[Y(t)] = \sigma^2 t$ with $\mu = E[Y(1)]$ and $\sigma^2 = \text{Var}[Y(1)]$. To cover non-linear degradation behavior we use a timescale $\tau = \tau(t)$, i.e., a positive, increasing, and

continuously differentiable function of real time t with $\tau(0) = 0$. It describes slowing or accelerating degradation in real time and is referred to as operational time. For instance, Bagdonavičius and Nikulin [3] consider a linear timescale with an initial acceleration period

$$\tau(t) = \gamma_1 t + \gamma_2(1 - \exp(-\gamma_3 t)), \quad \gamma_i > 0,$$

for tire protector wear and Whitmore and Schenkelberg [37]

$$\tau(t) = 1 - \exp(-\gamma_1 t^{\gamma_2}) \quad \text{and} \quad \tau(t) = t^{\gamma_1}, \quad \gamma_i > 0,$$

in the context of self-regulating heating cables. The choice of τ depends on whether degradation is unbounded or approaches a saturation point. A further possible form of the timescale is a Box-Cox transformation

$$\tau(t) = \begin{cases} \gamma_0((t+1)^{\gamma_1} - 1)/\gamma_1, & \gamma_1 > 0 \\ \gamma_0 \log(t+1) & , \quad \gamma_1 = 0 \end{cases}.$$

To model the influence on failure of an item's dynamic operating environment, the timescale may depend on the covariate process Z , for instance on different stress levels, i.e., $\tau_Z(t) = \tau(t, Z_0^t)$. An example is the model of *additive accumulation of damage* (AAD model, cf. [2]):

$$\tau_Z(t) = \tau \left(\int_0^t \exp(\beta^T Z(s)) ds \right).$$

Our degradation model is given by

$$X(t) = X_0 + Y(\tau_Z(t)) =: X_0 + Y^Z(t),$$

where Y^Z is the time-scaled process Y and X_0 is some possibly random initial degradation level independent of (Y, Z) . Conditioned on $\{Z = z(\cdot)\}$, $Y^{z(\cdot)}$ is a process with independent but not stationary increments. Obviously

$$\begin{aligned} \mathbb{E}[X(t) | Z_0^t] &= \mathbb{E}[X_0] + \mu \tau_Z(t), \\ \text{Var}[X(t) | Z_0^t] &= \text{Var}[X_0] + \sigma^2 \tau_Z(t). \end{aligned}$$

An item is regarded as failed when the degradation process reaches a critical threshold level X^* or when a traumatic event like a shock of large magnitude occurs. The traumatic event time is modeled as the first point of a doubly stochastic Poisson process $\Psi = \{\Psi(t) : t \in \mathbb{R}_+\}$ with an intensity $\kappa(t, Z(t), X(t))$ that depends on the environment and on the degradation level, i.e., the intensity κ of a traumatic event at time t depends on t , on the covariate value $Z(t)$, and on the degradation level $X(t) = Y(\tau(t, Z_0^t))$. Simulated data of a Wiener process and a gamma process DTS model with an exponential timescale are shown in Figures 14.5 and 14.6. The simulation parameters were $x^* = 9.0$ and $\kappa(t, X(t)) = 0.02 X(t)$ and the drift and the variance parameters of the simulated Wiener and gamma processes were equal to 1. To derive the conditional survival function of T , conditioned on a path Z_0^t , set $Y^* = X^* - X_0$ and let $T_Y(y^*) = \inf\{t \geq 0 : Y(t) \geq y^*\}$ be the first passage time of Y to the non-random threshold y^* with a survival function $\bar{F}_0(t, y^*)$. Assume that $T_Y(y^*)$ has a density $\mathbb{P}(T_Y(y^*) \in dt) = f_0(t, y^*)dt$ and a failure rate $\alpha_0(t, y^*) = f_0(t, y^*)/\bar{F}_0(t, y^*)$ for $y^* > 0$ and that the density at t of those trajectories of Y that have not exceeded y^* is $\mathbb{P}(Y(t) \in dy, T_Y(y^*) > t) = g_0(t, y, y^*)dy$. The proof of the following theorem is given in [22].

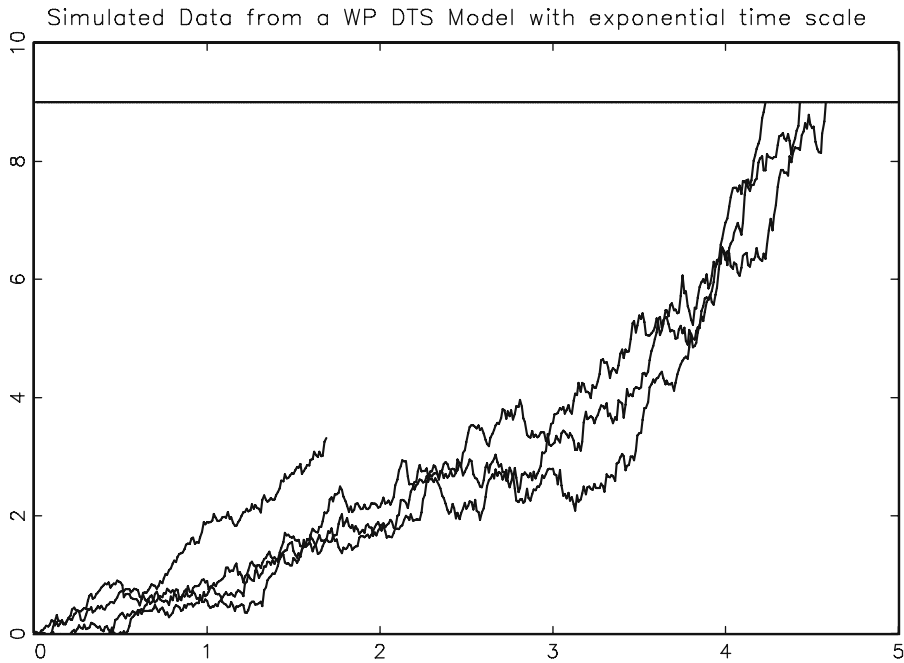


Figure 14.5. Wiener process DTS model with exponential timescale

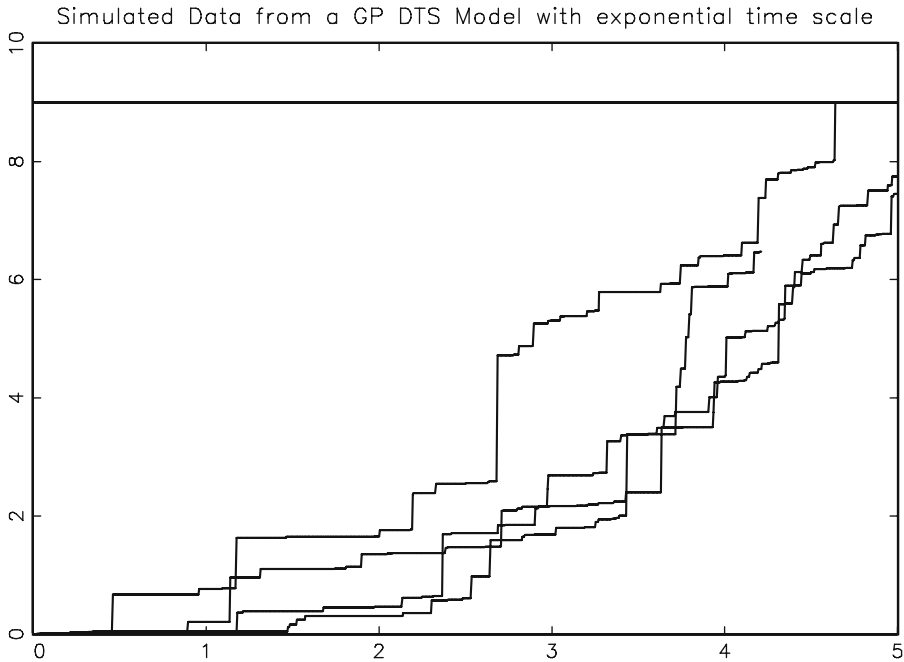


Figure 14.6. Gamma process DTS model with exponential timescale

Theorem 6. Let Y be independent of (Z, Y^*) with $P(Y^* > 0) = 1$. Conditioned on Z , the failure time T has for $t \geq 0$ a survival function given by

$$P(T > t | Z_0^t) = \exp \left(- \int_0^t (\bar{\kappa}(s, Z_0^s) + \bar{\lambda}(s, Z_0^s)) ds \right)$$

and a failure rate given by

$$\alpha(s, Z_0^s) = \bar{\kappa}(s, Z_0^s) + \bar{\lambda}(s, Z_0^s),$$

where $\bar{\lambda}(t, Z_0^t) = \tau_Z'(t) E[\alpha_0(\tau_Z(t), Y^*) | Z_0^t, D > t]$ is the failure rate of a nontraumatic event and $\bar{\kappa}(t, Z_0^t) = E[\kappa(t, Z(t), X(t)) | Z_0^t, T > t]$ is the failure rate of a traumatic event if a nontraumatic event has not yet occurred.

In particular, the failure rate α simplifies to

$$\alpha(t, Z_0^t) = \kappa(t, Z_0^t) + \tau_Z'(t) \alpha_0(\tau_Z(t), y^*)$$

if the intensity κ depends only on Z and not on X and if the threshold level $y^* = x^* - x_0$ is known. The independence condition of Y and Z means that the degradational process Y^Z depends on the external covariate Z only through the timescale τ_Z .

Example. Let $Y(t) = \mu t + \sigma W(t)$ be a Wiener process with drift. In this case the first passage time $T_Y(y^*) \sim \text{IG}\left(\frac{y^*}{\mu}, \frac{y^{*2}}{\sigma^2}\right)$ is Inverse Gaussian distributed with

$$\alpha_0(t, y^*) = \frac{\frac{y^*}{\sigma\sqrt{t^3}} \phi\left(\frac{y^* - \mu t}{\sigma\sqrt{t}}\right)}{N\left[\frac{y^* - \mu t}{\sigma\sqrt{t}}\right] - e^{(2\mu\sigma^{-2}y^*)} N\left[\frac{-y^* - \mu t}{\sigma\sqrt{t}}\right]}$$

and

$$g_0(t, y, y^*) = \frac{1}{\sigma\sqrt{t}} \phi\left(\frac{y - \mu t}{\sigma\sqrt{t}}\right) \left[1 - \exp\left(-\frac{2y^*(y^* - y)}{\sigma^2 t}\right) \right],$$

for $t > 0$ and $y < y^*$. The resulting degradation model

$$X(t) = X_0 + \mu\tau_Z(t) + \sigma W(\tau_Z(t))$$

is appropriate if the actual degradation process $\mu\tau_Z(t)$ is latent but observable with some additive error term which has a variance proportional to the actual degradation process. The shape of the stochastic failure rate $\bar{\lambda}$ depends heavily on the derivative $\tau_Z'(t)$ of the timescale. For example, it is bathtub shaped for suitably chosen timescales.

14.3.1 Maximum likelihood estimation

We consider a Wiener process DT model with a nonrandom threshold $x^* = h > 0$ and a timescale $\tau(t) = \tau(t; \gamma)$ parameterized by some parameter γ . Suppose that n independent items are put on test at real times t_{i0} . For each item we assume that the external covariate process Z_i is observed continuously in the interval $[t_{i0}, t_{im_i}]$ and that the initial degradation level $Y_{i0} = X_{i0}$ and the degradation levels $Y_{ij} = Y_i(\tau_Z(t_{ij}))$

($j > 0$) are observed at $m_i + 1$ inspection times $t_{i0} < t_{i1} < t_{i2} < \dots < t_{im_i} < \infty$, but only until a failure of the item has occurred. The failure times T_i are assumed to be observable at any time $t \in (t_{i0}, t_{im_i}]$.

Denoting the observation times in operational time by $\tau_{ij} = \tau_Z(t_{ij})$ where $\tau_{i0} = 0$ we observe in each operational time interval $(\tau_{ij-1}, \tau_{ij}]$ either a realization of the lifetime $T_i^{op} = \tau_Z(T_i)$ or the degradation level Y_{ij} at τ_{ij} under the condition that the process Y_i has not yet reached the threshold h until τ_{ij} . Conditioned on the covariate paths, the likelihood function of such a model is given by Theorem 5 as

$$L(\mathcal{X}; \theta) = \prod_{i=1}^n \left\{ \prod_{j=1}^{M_i} \left(\frac{1}{\sigma \sqrt{\Delta \tau_{ij}}} \phi \left(\frac{\Delta Y_{ij} - \mu \Delta \tau_{ij}}{\sigma \sqrt{\Delta \tau_{ij}}} \right) \times \left[1 - \exp \left(-\frac{2(h - Y_{ij-1})(h - Y_{ij})}{\sigma^2 \Delta \tau_{ij}} \right) \right] \right) \times \left(\tau'_{Z_i}(T_i) \frac{h - Y_{iM_i}}{\sigma \sqrt{(T_i^{op} - \tau_{iM_i})^3}} \phi \left(\frac{h - Y_{iM_i} - \mu(T_i^{op} - \tau_{iM_i})}{\sigma \sqrt{T_i^{op} - \tau_{iM_i}}} \right) \right)^{\mathbf{1}\{T_i^{op} < \tau_{im_i}\}} \right\}, \quad (14.10)$$

with ϕ denoting the standard normal density and \mathcal{X} the observed data, $\theta = (\mu, \sigma^2, \beta, \gamma)$, $M_i = \max\{k : \tau_{ik} < T_i^{op}\}$, and $\Delta Y_{ij} = Y_{ij} - Y_{ij-1}$, ($i = 1, \dots, n$; $j = 1, \dots, m_i$).

We use maximum likelihood estimation to find estimates of the parameters μ , σ^2 , β , and γ . In general, the log-likelihood equations have to be solved numerically. Only for the drift parameter μ we obtain explicitly the estimate

$$\hat{\mu} = \frac{\sum_{i=1}^n (Y_{iM_i} - Y_{i0} + \mathbf{1}\{T_i^{op} < \tau_{im_i}\} (h - Y_{iM_i}))}{\sum_{i=1}^n (\tau_{iM_i} - \tau_{i0} + \mathbf{1}\{T_i^{op} < \tau_{im_i}\} (T_i^{op} - \tau_{iM_i}))}$$

which depends, however, on the timescale parameters. It can be easily interpreted as the total observed degradation divided by the total time on test. For the variance parameter σ^2 we find the implicit estimate

$$\widehat{\sigma^2} = \frac{1}{M^*} \sum_{i=1}^n \left\{ \sum_{j=1}^{M_i} \frac{(\Delta Y_{ij} - \hat{\mu} \Delta \tau_{ij})^2}{\Delta \tau_{ij}} + 2\widehat{\sigma^2} Q_{ij} (1 - \coth Q_{ij}) + \mathbf{1}\{T_i^{op} < \tau_{im_i}\} \frac{(h - Y_{iM_i} - \hat{\mu}(T_i^{op} - \tau_{iM_i}))^2}{T_i^{op} - \tau_{iM_i}} \right\}$$

with

$$Q_{ij} = (h - Y_{ij-1})(h - Y_{ij}) / \widehat{\sigma^2} \Delta \tau_{ij} > 0$$

and

$$M^* = \sum_{i=1}^n (M_i + \mathbf{1}\{T_i^{op} < \tau_{im_i}\}).$$

Using the simple approximation

$$Q_{ij}(1 - \coth Q_{ij}) \approx \begin{cases} Q_{ij} - 1, & Q_{ij} < 1 \\ 0, & Q_{ij} \geq 1 \end{cases}$$

we obtain an approximate explicit estimate for σ^2 .

ML estimates of the timescale parameters have to be found numerically by maximization of the likelihood function $L(\mathcal{X}; \hat{\mu}, \hat{\sigma}^2, \beta, \gamma)$ or by the simplified least square approach $\sum_{i=1}^n \sum_{j=1}^{M_i} (\Delta Y_{ij} - \hat{\mu} \Delta \tau_{ij})^2 \xrightarrow{\beta, \gamma} \min$. If the timescale is not given in parameterized form, a semiparametric approach can be used, based on a piecewise linear approximation of $Y_{Z_i}(t)$ and conditional means $E[Y_{Z_i}(t) | (Z_i)_0^t, T_i > t]$ (see Section 14.3.2).

If the threshold level h is unknown, it can be regarded as an additional parameter that has to be estimated. An inspection of the likelihood function (14.10) shows that the ML estimate \hat{h} will not exist if in all realizations failures could not be observed, since in this case we have $\arg \sup_{h>0} L(\mathcal{X}; \theta) = +\infty$ and the maximum of the likelihood function with respect to h will be attained at infinity at the edge of the parameter region. Maximum likelihood estimation of the parameters μ , σ^2 , and h for a Wiener degradation model without timescale transformation using simulated data is illustrated in Chapter 9 by Kahle and Lehmann. This simulation study shows that the mean-squared errors of the ML estimates of μ , σ^2 , and h based on the likelihood function (14.10) are smaller than those of ML estimates purely based on failure time data or on degradation data.

The above threshold model can be extended to repairable items in the following way. Upon each failure and possibly at the regular inspection times of degradation measurements the item is repaired in negligible time. At an inspection time degradation is measured and, if necessary, the item is preventively repaired. Each repair action sets back the degradation level to a value between its previous level and the level of a new item. The latter case is equivalent to a perfect repair, the first case to a minimal repair, and any level in between corresponds to an imperfect repair.

14.3.2 Semiparametric estimation

Suppose that the mean degradation function τ in the above parameterized model is completely unknown and that the covariate process Z is constant in time or a step function. In this case we can use an approach proposed by Bagdonavičius and Nikulin [3] which is presented here in adapted form.

Denote by $P(t, Z_0^t; \beta)$ the inverse of the increasing function

$$Q(t, Z_0^t; \beta) = \int_0^t q(Z(s); \beta) ds.$$

Then, the process $Y(\tau_Z(P(t, Z_0^t; \beta)))$ is a process with independent increments and

$$E[Y(\tau_Z(P(t, Z_0^t; \beta)))] = \mu \tau(t), \quad \text{Var}[Y(\tau_Z(P(t, Z_0^t; \beta)))] = \sigma^2 \tau(t). \quad (14.11)$$

Now, we drop the parameter μ as only $\mu \tau(t)$ can be estimated by this approach. If $Z(t) \equiv z$ is constant in time, then

$$Q(t, Z_0^t; \beta) = q(z; \beta) t, \quad P(t, Z_0^t; \beta) = (q(z; \beta))^{-1} t =: p(z; \beta) t.$$

If $Z(t)$ is a step function of the form

$$Z(t) = z_j, \quad t \in [s_{j-1}, s_j), \quad j = 1, \dots, J,$$

with $0 = s_0 < s_1 < \dots < s_J$, then

$$\begin{aligned} Q(t, Z_0^t; \beta) &= s_{j-1}^*(\beta) + q(z_j; \beta) (t - s_{j-1}), & t \in [s_{j-1}, s_j], \\ P(t, Z_0^t; \beta) &= s_{j-1} + p(z_j; \beta) (t - s_{j-1}^*(\beta)), & t \in [s_{j-1}^*(\beta), s_j^*(\beta)], \end{aligned}$$

where

$$s_j^*(\beta) = \sum_{k=1}^j q(z_k; \beta) (s_k - s_{k-1}), \quad j = 1, \dots, J, \quad s_0^*(\beta) = 0.$$

Consider the following piecewise linear approximation $\tilde{Y}_i(t)$ of the process $Y_i^{z_i(\cdot)}(t) = Y_i(\tau(Q(t, z_i(\cdot))))$, both defined on $[0, t_{im_i}]$:

$$\tilde{Y}_i(t) = \sum_{j=1}^{m_i} \left[Y_i^{z_i(\cdot)}(t_{i,j-1}) + \frac{t - t_{i,j-1}}{t_{ij} - t_{i,j-1}} \Delta y_{ij} \right] \mathbf{1}_{[t_{i,j-1}, t_{ij}]}(t).$$

Setting $P_i(t; \beta) = P(t, z_i(\cdot); \beta)$ we conclude from (14.11) that for all $i = 1, \dots, n$

$$\mathbb{E} [\tilde{Y}_i(P_i(t; \beta))] \approx \tau(t), \quad \text{Var} [\tilde{Y}_i(P_i(t; \beta))] \approx \sigma^2 \tau(t),$$

if the differences Δt_{ij} are small. In this case the mean $\tilde{\tau}(t; \beta)$ of the values $\tilde{Y}_i(P_i(t; \beta))$, which, of course, still depends on β , should be a reasonable estimator of the unknown function $\tau(t)$ at t . Note that the processes $\tilde{Y}_i(P_i(t; \beta))$ are censored at the points $t_i^*(\beta) = P_i(t_{im_i}; \beta)$. Consider the ordered sequence

$$t_{(1)}^*(\beta) < \dots < t_{(n)}^*(\beta)$$

of these points. Then the estimator $\tilde{\tau}(t; \beta)$ is given by

$$\tilde{\tau}(t; \beta) = \frac{1}{n} \sum_{i=1}^n \tilde{Y}_i(P_i(t; \beta))$$

for $t \in [0, t_{(1)}^*(\beta))$ and

$$\begin{aligned} \tilde{\tau}(t; \beta) &= \tilde{\tau}(t_{(j-1)}^*(\beta); \beta) \\ &+ \frac{1}{n - j + 1} \sum_{i: t_i^*(\beta) > t_{(j-1)}^*(\beta)} \left(\tilde{Y}_i(P_i(t; \beta)) - \tilde{Y}_i(P_i(t_{(j-1)}^*(\beta); \beta)) \right) \end{aligned}$$

for $t \in [t_{(j-1)}^*(\beta), t_{(j)}^*(\beta))$.

Plugging $\tilde{\tau}(t; \beta)$ in the likelihood function (14.10) and computing the maximum likelihood estimator $\hat{\theta}^*$ of $\theta^* = (\mu, \sigma^2, \beta)$ we can estimate $\tau(t)$ by $\tilde{\tau}(t; \hat{\beta})$.

14.4 A DTS Model for Repairable Items

To extend the DTS model with covariates from the last section to repairable items we use a marked point process approach (cf. [20]). Upon each failure and possibly at regular inspection times $0 = b_0 < b_1 < \dots$, an item will be repaired in negligible time depending on the current degradation level. Each repair action reduces the degradation level from the current level $X(T_n)$ to $(1 - R_n)X(T_n)$, where $(R_n)_{n \geq 1}$ is a sequence of $[0, 1]$ -valued random repair degrees. A repair degree zero denotes a minimal repair without reduction of the degradation level, or no repair, if a repair action is not necessary. A repair degree one is equivalent to a perfect repair which sets back the degradation level to zero, which is assumed to be the level of a new item. Any repair degree between zero and one corresponds to an imperfect repair.

The degradation failure marked point process $\Phi = ((T_n, V_n))_{n \geq 1}$ consists of event times T_n and marks V_n that contain additional information. The event times T_n and the process $X(t)$ on $(T_n, T_{n+1}]$ are defined recursively in the following way. First set $B_0 = C_0 = D_0 = T_0 = R_0 = 0$ and $X(0) = X_0$, where X_0 is some random initial degradation level. If X is already defined on $(0, T_n]$ for $n \geq 0$ set

$$X(t) = (1 - R_n)X(T_n) + Y^Z(t) - Y^Z(T_n)$$

on $(T_n, T_{n+1}]$ where $T_{n+1} = B_{n+1} \wedge C_{n+1} \wedge D_{n+1}$ with

$$\begin{aligned} B_{n+1} &= \min\{b_l : b_l > T_n\}, \\ C_{n+1} &= \inf\{t > T_n : \Delta\Psi(t) > 0\}, \\ D_{n+1} &= \inf\{t > T_n : X(t) \geq X^*\}, \end{aligned}$$

i.e., B_{n+1} is the next inspection time, C_{n+1} denotes the next traumatic event time, and D_{n+1} refers to the next nontraumatic failure time after T_n . Ψ is modeled as a doubly stochastic Poisson process with an intensity $\kappa(t, Z(t), X(t))$ depending on the environment and on the degradation level. Thus, for $T_n < t \leq T_{n+1}$ the resulting degradation process X is given as

$$X(t) = \pi_{1,n}X_0 + \sum_{i=1}^n \pi_{i,n}\Delta Y_i + Y^Z(t) - Y^Z(T_n),$$

where $\pi_{i,n} = \prod_{j=i}^n (1 - R_j)$ and $\Delta Y_i = Y^Z(T_i) - Y^Z(T_{i-1})$.

The marks $V_n = (V_n^1, V_n^2, V_n^3)^\top$ contain additional information on the event modes, the degradational increments between event times, and the results of repair actions. They are defined by $V_0 = (0, 0, X_0)^\top$ and

$$V_n^1 = \delta_n = \begin{cases} 0, & \text{if } B_n < C_n \wedge D_n \\ 1, & \text{if } D_n < C_n, D_n \leq B_n, \\ -1, & \text{if } C_n \leq B_n \wedge D_n \end{cases}$$

$$V_n^2 = \Delta Y_n = Y^Z(T_n) - Y^Z(T_{n-1}),$$

and

$$V_n^3 = X(T_n+) = (1 - R_n)X(T_n)$$

for $n \geq 1$. A value of 0 of the first mark V_n^1 stands for a regular inspection at T_n , a value of 1 for a nontraumatic failure at T_n , and a value of -1 for a traumatic failure at T_n .

We suppose that the repair degrees $(R_n)_{n \geq 1}$ are conditionally independent given $(T_n, X(T_n), V_n^1)_{n \geq 1}$ and that $R_n > 0$ if $V_n^1 = 1$, i.e., in case of a nontraumatic failure. The point process Φ is adapted to the global history \mathcal{F}_t of information on Φ up to time t and on Z . The conditional survival function of T_{n+1} given \mathcal{F}_{T_n} , i.e., given all observable information up to T_n , is given in the following theorem for the case of a nonrandom threshold x^* . The proof of the following theorems was given in [22]:

Theorem 7. *Let $\Phi = ((T_n, V_n))_{n \geq 1}$ be the marked point process in the DTS model and assume that Y is independent of (Z, X_0, R) and $P(X_0 < x^*) = 1$. Then, on (T_n, ∞) , the conditional survival function of T_{n+1} is given by*

$$P(T_{n+1} > t | \mathcal{F}_{T_n}) = \mathbf{1}(B_{n+1} \geq t) \exp \left(- \int_{T_n}^t (\bar{\kappa}(s | \mathcal{F}_{T_n}) + \lambda(s | \mathcal{F}_{T_n})) ds \right)$$

with

$$\bar{\kappa}(t | \mathcal{F}_{T_n}) = E[\kappa(t, Z(t), X(t)) | \mathcal{F}_{T_n}, T_{n+1} > t]$$

and

$$\lambda(t | \mathcal{F}_{T_n}) = \tau'_Z(t) \alpha_0(\tau_Z(t) - \tau_Z(T_n), x^* - X(T_n+)).$$

Frequently, the failure rate $\bar{\kappa}(t | \mathcal{F}_{T_n})$ cannot be evaluated in closed form. If the inspection time intervals (b_n, b_{n+1}) are short and the degradation process is slowly varying, a reasonable simplification of the model is the assumption that the intensity $\kappa(t, Z(t), X(t))$ of traumatic events depends in (b_n, b_{n+1}) only on the degradation level at b_n . Then the failure rate $\bar{\kappa}$ simplifies in the following way.

Corollary 1. *If, on $(T_n, B_{n+1}]$,*

$$\kappa(t, Z(t), X(t)) = \kappa(t, Z(t), X(T_n+)),$$

then $\bar{\kappa}(t | \mathcal{F}_{T_n}) = \kappa(t, Z(t), X(T_n+))$ on $(T_n, B_{n+1}]$.

14.5 Application of DTS Models

In applied literature degradation processes are frequently described by a general path model, i.e., by a stochastic process that depends only on a finite dimensional random variable (cf. [25]), or by a univariate process with stationary independent increments.

Degradation-threshold models with general path degradation modeling were applied to crack growth, tire wear, electronic component degradation, silicon solar cell degradation, and disk storage media (cf. [24–26]). DT models with Gaussian degradation processes and random timescales have found application to fatigue crack data, degradation of heating cables, carbon fibers, and biomarker data (cf. [11, 37, 27]).

Degradation–shock models based on gamma and Wiener degradation processes were considered by Wenocur [34, 35] and by Singpurwalla and Youngren [32]. Bagdonavičius and Nikulin [3] considered maximum likelihood and semiparametric estimation in a DTS model with a time-scaled gamma degradation process and applied this model to tire wear.

In the context of DTS models, Bagdonavičius et al. [4] consider a linear path model with time-dependent covariates and multiple traumatic event modes. Non-parametric estimators, in general path DTS models with multiple traumatic failure modes and renewals, were analyzed by Bagdonavičius et al. [5].

A detailed review on stochastic-process-based reliability models including DTS models was given by Singpurwalla [30]. Tsiatis and Davidian [33] gave an overview on joint modeling of longitudinal and failure data in biostatistics.

The DTS model can be easily extended to the case that m different modes of traumatic events are considered such that traumatic failures of mode i occur due to a point process Ψ_i . If all these point processes are doubly stochastic Poisson processes conditionally independent given the degradation path $X(\cdot)$ and the covariate path $Z(\cdot)$ and adapted to appropriate filtrations with intensities $\kappa_i(t, X(t))$ or $\kappa_i(t, Z(t), X(t))$, then Theorems 2, 4, 5, 6, and 7 remain valid if we replace $\bar{\kappa}(\cdot)$ by $\sum_{i=1}^m \bar{\kappa}_i(\cdot)$.

References

1. Aven, T. and Jensen, U. (1999). *Stochastic Models in Reliability*. Springer, New York.
2. Bagdonavičius, V. (1978). Testing the hypothesis of the additive accumulation of damage, *Probability Theory and its Applications*, **23**, 403–408.
3. Bagdonavičius, V. and Nikulin, M.S. (2001). Estimation in degradation models with explanatory variables, *Lifetime Data Analysis*, **7**, 85–103.
4. Bagdonavičius, V., Bikelis, A., and Kazakečius, V. (2004). Statistical analysis of linear degradation and failure time data with multiple failure modes. *Lifetime Data Analysis*, **10**, 65–81.
5. Bagdonavičius, V., Bikelis, A., Kazakečius, V., and Nikulin, M.S. (2006). Non-parametric estimation in degradation-renewal-failure models, In: *Probability, Statistics and Modelling in Public Health*, M.S. Nikulin, D. Commenges, and Huber, C. (eds.), Springer Science+Business Media.
6. Bogdanoff, J.L. and Kozin, F. (1985). *Probabilistic Methods of Cumulative Damage*, John Wiley & Sons, New York.
7. Carey, M.B. and Koenig, R.H. (1991). Reliability assessment based on accelerated degradation: A case study, *IEEE Transactions on Reliability*, **40**, 499–506.
8. Couallier, V. (2006). Some recent results on joint degradation and failure time modelling. In: *Probability, Statistics and Modelling in Public Health*, M.S. Nikulin, D. Commenges, and Huber, C. (eds.), Springer Science+Business Media.
9. Cox, D.R. (1999). Some remarks on failure-times, surrogate markers, degradation, wear and the quality of life, *Lifetime Data Analysis*, **5**, 307–314.
10. Doksum, K.A. and Hoyland, A. (1992). Models for variable-stress accelerated life testing experiment based on a Wiener process and the inverse Gaussian distribution, *Technometrics*, **34**, 74–82.
11. Doksum, K.A. and Normand, S.L.T. (1995). Gaussian models for degradation processes – part I: Methods for the analysis of biomarker data, *Lifetime Data Analysis*, **1**, 131–144.

12. Finkelstein, M.S. (2004). On the exponential formula of reliability, *IEEE Transactions on Reliability*, **53**, 265–269.
13. Flournoy, N. (1980). On the survivor and hazard functions. Paper read at the 1980 Statistical Meetings, Houston, Texas: Abstract published by the American Statistical Association.
14. Kahle, W. and Lehmann, A. (1998). Parameter estimation in damage processes: Dependent observations of damage increments and first passage time. In: *Advances in Stochastic Models for Reliability, Quality and Safety*, W. Kahle, E. von Collani, F. Franz, and U. Jensen, (eds.), Birkhauser, Boston.
15. Kalbfleisch, J.D. and Prentice, R.S. (1980). *The Statistical Analysis of Failure Time Data*, John Wiley & Sons, New York.
16. Kebir, Y. (1991). On hazard rate processes, *Naval Research Logistics Quarterly*, **38**, 865–876.
17. Lawless, J. and Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure *Lifetime Data Analysis*, **10**, 213–227.
18. Lee, M.L.T., DeGruttola, V., and Schoenfeld, D. (2000). A model for markers and latent health status, *Journal of the Royal Statistical Society, Series B*, **62**, 747–762.
19. Lehmann, A. (2001). A Wiener process based model for failure and degradation data in dynamic environments, *Dresdner Schriften zur Mathemat. Stochastik*, **4/2001**, 35–40.
20. Lehmann, A. (2004). On a degradation-failure model for repairable items. In: *Semiparametric Models and its Applications for Reliability, Survival Analysis and Quality of Life*, M.S. Nikulin, N. Balakrishnan, N. Limnios, and M. Mesbah (eds.), Birkhauser, Boston.
21. Lehmann, A. (2005). Joint models for degradation and failure time data, In: *Proceedings of the International Workshop Statistical Modelling and Inference in Life Sciences*, September 1–4. Potsdam.
22. Lehmann, A. (2009). Joint modeling of degradation and failure time data, *Journal of Statistical Planning and Inference*, **139**, 1693–1706.
23. Lemoine, A.L. and Wenocur, M.L. (1985). On failure modeling, *Naval Research Logistics Quarterly*, **32**, 479–508.
24. Lu, C.J. and Meeker, W.Q. (1993). Using degradation measures to estimate a time-to-failure distribution, *Technometrics*, **35**, 161–174.
25. Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*, John Wiley & Sons, New York.
26. Meeker, W.Q., Escobar, L.A., and Lu, C.J. (1998). Accelerated degradation test: Modeling and analysis, *Technometrics*, **40**, 89–99.
27. Padgett, W.J. and Tomlinson, M.A. (2004). Inference from accelerated degradation and failure data based on Gaussian process models, *Lifetime Data Analysis*, **10**, 191–206.
28. Padgett, W.J. and Tomlinson, M.A. (2005). Accelerated degradation models for failure based on geometric Brownian motion and gamma processes, *Lifetime Data Analysis*, **11**, 511–527.
29. Shaked, M. and Shantikumar, J.G. (1988). On the first-passage times of pure jump processes, *Journal of Applied Probability*, **25**, 501–509.
30. Singpurwalla, N.D. (1995). Survival in dynamic environments, *Statistical Science*, **10**, 86–103.
31. Singpurwalla, N.D. and Wilson, S.P. (1995). The exponential formula of reliability and survival: Does it always hold? *Lifetime Data Analysis*, **1**, 187–194.
32. Singpurwalla, N.D. and Youngren, M.A. (1998). Multivariate distributions induced by dynamic environments, *Scandinavian Journal of Statistics*, **20**, 251–261.
33. Tsiatis, A.A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: An overview, *Statistics Sinica*, **14**, 809–834.
34. Wenocur, M.L. (1986). Brownian motion with quadratic killing and some implications, *Journal of Applied Probability*, **23**, 893–903.

35. Wenocur, M.L. (1989). A reliability model based on the gamma process and its analytical theory, *Advances in Applied Probability*, **21**, 899–918.
36. Whitmore, G.A. (1995). Estimation degradation by a Wiener diffusion process subject to measurement error, *Lifetime Data Analysis*, **1**, 307–319.
37. Whitmore, G.A. and Schenkelberg, F. (1997). Modelling accelerated degradation data using Wiener diffusion with a time scale transformation, *Lifetime Data Analysis*, **3**, 27–45.
38. Whitmore, G.A., Crowder, M.I., and Lawless, J. (1998). Failure inference from a marker process based on a bivariate Wiener model, *Lifetime Data Analysis*, **4**, 229–251.
39. Yashin, A.I. (1985). Dynamics in survival analysis: conditional Gaussian property versus Cameron-Martin formula. In: *Statistics and Control of Stochastic Processes*, N.V. Krylov, R.S. Liptser, and A.A. Novikov (eds.), Springer, New York.
40. Yashin, A.I. and Arjas, E. (1988). A note on random intensities and conditional survival functions, *Journal of Applied Probability*, **25**, 630–635.
41. Yashin, A.I. and Manton, G.M. (1997). Effects of unobserved and partially observed covariate processes on system failure: A review a models and estimation strategies, *Statistical Science*, **12**, 20–34.

Degradation and Fuzzy Information

R. Viertl

Institute of Statistics and Probability Theory, Vienna University of Technology, Wien, Austria

Abstract: In lifetime analysis different kinds of uncertainty are present. Besides variability there is also imprecision of measurement results. This kind of uncertainty is best described by fuzzy models. Besides fuzziness of degradation parameters there is also fuzziness of probability distributions. Using fuzzy models in combination with statistical models and degradation functions more realistic lifetime analysis is possible.

Keywords and phrases: Accelerated life testing, degradation models, fuzzy information, fuzzy numbers, fuzzy probability distributions, lifetimes, lifetime analysis

15.1 Introduction

There are different models to describe material degradation in the context of lifetime analysis. More on degradation models can be found in (Meeker [1]). Let $q(t)$ be a function which describes the degradation of a unit depending on time $t \geq 0$. The end of the lifetime is reached if the quality parameter $q(t)$ is less than a given critical value q_{crit} , i.e.,

$$\text{lifetime} = \inf \{t \geq 0 : q(t) \leq q_{\text{crit}}\}. \quad (15.1)$$

In Figure 15.1 the situation is depicted.

The distribution function $G(\cdot)$ of the stochastic quantity T – describing the lifetime – is obtained from the distribution function $F_0(\cdot)$ of the initial value of the quality parameter $q(0)$.

For monotonic degradation function $q(\cdot)$ the values $G(t)$ are given by the following equation:

$$G(t) = \Pr \{T \leq t\} = \Pr \{q(t) \leq q_{\text{crit}}\} \quad \forall t \geq 0. \quad (15.2)$$

The event $\{T \leq t\}$ is equivalent to $\{q(0) \leq q_{0,t}\}$, where $q_{0,t}$ is the value of $q(0)$ for which

$$t = \inf \{t \geq 0 : q(t) \leq q_{\text{crit}}\}. \quad (15.3)$$

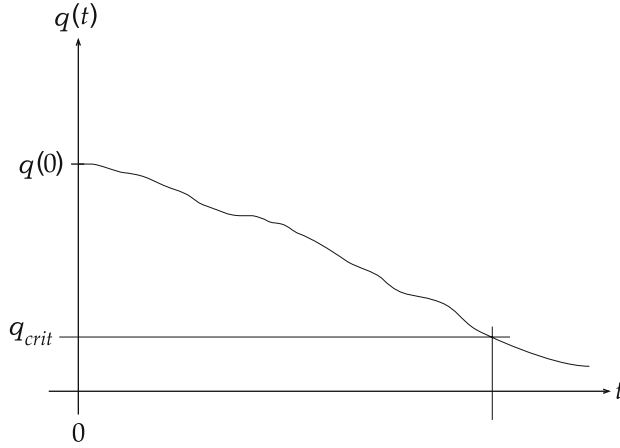


Figure 15.1. Degradation of a quality parameter $q(t)$

There are different kinds of fuzziness in this situation. First, the values $q(0)$ of the initial quality parameter are to be measured. If they are continuous the measurement results are not precise numbers, but more or less fuzzy. Another kind of fuzziness is the uncertainty of the probability distribution $F_0(\cdot)$.

15.2 Material Degradation

Physical or chemical processes are governed by differential equations depending on time t . In microelectronics the following differential equation is used:

$$\frac{dq}{dt} = \frac{-C(S)}{\Psi(q)}, \quad (15.4)$$

with $C(S) > 0$, $\Psi(q) > 0$, $C(S)$ a constant depending on thermal stress S .

The solution of Equation (15.4) is given by

$$\int_0^{q(t)} \Psi(q) dq = C(S) \cdot (t_0 - t). \quad (15.5)$$

By the structure of the differential equation (15.4) the solutions $q(\cdot)$ are nonincreasing functions which are non-crossing for different initial values $q(0)$. This is depicted in Figure 15.2.

The initial value $q(0)$ of the quality parameter whose corresponding lifetime is t is denoted by $q_{0,t}$. This value is obtained from the following equation:

$$t = \frac{1}{C(S)} \int_{q_{\text{crit}}}^{q_{0,t}} \Psi(q) dq \quad (15.6)$$

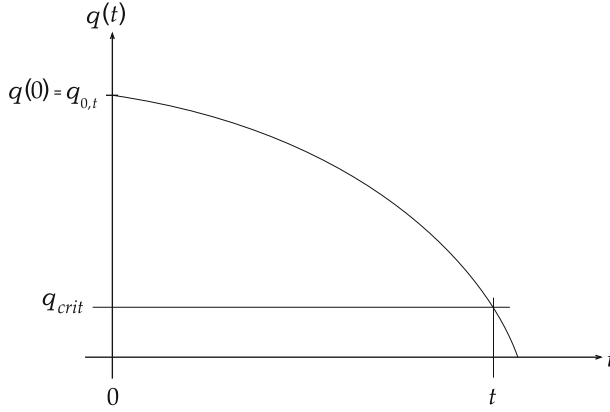


Figure 15.2. Solutions of Equation (15.4)

Based on this the distribution function $G(\cdot)$ of the lifetime T is obtained as in the introduction by

$$G(t) = F_0(q_{0,t}) \quad \forall t \geq 0. \quad (15.7)$$

15.3 Fuzzy Initial Conditions

The initial quality parameter $q(0)$ is usually a continuous quantity which has to be measured. The result of such measurements are not precise numbers but more or less fuzzy. This imprecision is best described by so-called *fuzzy numbers* x^* with corresponding *characterizing function* $\xi(\cdot)$, i.e., a real function of one real variable x obeying the following:

- (C1) $\xi : \mathbb{R} \longrightarrow [0; 1]$
- (C2) $\text{supp}[\xi]$ is bounded
- (C3) $\forall \delta \in (0; 1]$ the so-called δ -cut

$$C_\delta[\xi] := \{x \in \mathbb{R} : \xi(x) \geq \delta\} \quad (15.8)$$

is nonempty and a compact interval $[a_\delta; b_\delta]$

For more details compare (Viertl [4]).

Based on the fuzziness of the initial value $q(0)$, this is modeled by a fuzzy number $q^*(0)$. Therefore also the generalized solution of the differential equation is a *fuzzy-valued function* $q^*(\cdot)$ whose so-called δ -level functions $\underline{q}_\delta(\cdot)$ and $\bar{q}_\delta(\cdot)$ are defined in the following way.

Denoting the δ -cuts of the initial condition by $[\underline{q}_\delta(0); \bar{q}_\delta(0)]$, the δ -level functions are the solutions of the differential equation for the corresponding initial conditions $\underline{q}_\delta(0)$ and $\bar{q}_\delta(0)$, respectively. The fuzzy-valued function $q^*(\cdot)$ is graphically depicted using some δ -level functions. In Figure 15.3 an example is given.

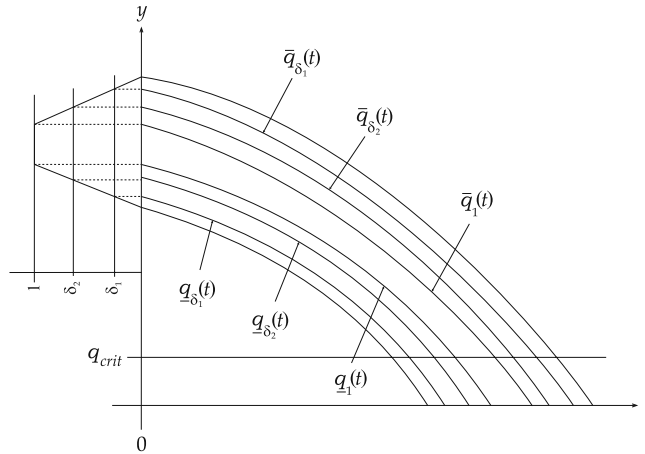


Figure 15.3. Fuzzy-valued function

The lifetime is also a fuzzy quantity t^* whose characterizing function $\eta(\cdot)$ is obtained via its δ -cuts $[\underline{\eta}_\delta; \bar{\eta}_\delta]$ which are determined by the δ -level functions $\underline{q}_\delta(\cdot)$ and $\bar{q}_\delta(\cdot)$ in the following way.

$\bar{\eta}_\delta$ is the abscissa of the intersection of $\bar{q}_\delta(t)$ with the constant function $y(t) = q_{\text{crit}}$, and $\underline{\eta}_\delta$ is the abscissa of the intersection of $\underline{q}_\delta(t)$ with $y(t) = q_{\text{crit}}$.

By the so-called *representation lemma* for characterizing functions (compare (Viertl [3])) the values $\eta(t)$ of the characterizing function $\eta(\cdot)$ of the fuzzy lifetime are given by

$$\eta(t) = \sup \left\{ \delta \cdot \mathbf{1}_{[\underline{\eta}_\delta; \bar{\eta}_\delta]}(t) : \delta \in [0; 1] \right\} \quad \forall t \in \mathbb{R}. \quad (15.9)$$

Remark 1. The support of $\eta(\cdot)$ is a subset of \mathbb{R}_+ .

15.4 Fuzzy Distribution of $q(0)$

In order to calculate the distribution function $G(\cdot)$ of the lifetime, the distribution function $F_0(\cdot)$ of the initial value $q(0)$ of the quality parameter is used as explained in Section 15.1.

In applications it is frequently unjustified to assume a precise probability distribution $F_0(\cdot)$. A more general approach from soft modeling is the so-called *fuzzy probability distributions* P^* which assign fuzzy numbers to the events of a measurable space (M, \mathcal{E}) in the following way:

- (1*) $P^* : \mathcal{E} \rightarrow \mathcal{F}([0; 1])$ the set of all fuzzy numbers with support in $[0; 1]$
- (2*) $P^*(\emptyset) = \mathbf{1}_{\{0\}}(\cdot)$ and $P^*(M) = \mathbf{1}_{\{1\}}(\cdot)$

(3*) For pairwise disjoint events A_1, \dots, A_m

$$\left. \begin{aligned} \overline{P}_\delta \left(\bigcup_{i=1}^m A_i \right) &\leq \sum_{i=1}^m \overline{P}_\delta (A_i) \\ \underline{P}_\delta \left(\bigcup_{i=1}^m A_i \right) &\geq \sum_{i=1}^m \underline{P}_\delta (A_i) \end{aligned} \right\} \quad \forall \delta \in [0; 1], \quad (15.10)$$

where $C_\delta [P^*(A)] = [\underline{P}_\delta(A); \overline{P}_\delta(A)] \quad \forall \delta \in (0; 1]$.

For generalized continuous distributions so-called *fuzzy probability densities* $f^*(\cdot)$ are considered, i.e., $f^* : \mathbb{R} \rightarrow \mathcal{F}(\mathbb{R}_+)$, where $\mathcal{F}(\mathbb{R}_+)$ is the set of positive fuzzy numbers, with δ -level functions $\underline{f}_\delta(\cdot)$ and $\overline{f}_\delta(\cdot)$ which are integrable functions with finite integrals such that there exists a classical probability density $f(\cdot)$ in \mathbb{R} obeying

$$\underline{f}_1(x) \leq f(x) \leq \overline{f}_1(x) \quad \forall x \in \mathbb{R}. \quad (15.11)$$

The fuzzy probability of a Borel set $B \subseteq \mathbb{R}$ is defined to be the fuzzy number $P^*(B)$ whose δ -cuts $[\underline{P}_\delta(B); \overline{P}_\delta(B)]$ are given by

$$\left. \begin{aligned} \overline{P}_\delta(B) &= \sup \left\{ \int_B g(x) dx : g(\cdot) \in \mathcal{D}_\delta \right\} \\ \underline{P}_\delta(B) &= \inf \left\{ \int_B g(x) dx : g(\cdot) \in \mathcal{D}_\delta \right\} \end{aligned} \right\} \quad \forall \delta \in (0; 1], \quad (15.12)$$

where \mathcal{D}_δ is the set of all classical probability densities $g(\cdot)$ obeying

$$\underline{f}_\delta(x) \leq g(x) \leq \overline{f}_\delta(x) \quad \forall x \in \mathbb{R}. \quad (15.13)$$

Remark 2. By the above definition a fuzzy probability distribution is obtained.

Based on fuzzy densities $f_0^*(\cdot)$ for $q(0)$ a fuzzy lifetime distribution is obtained in the following way.

Let $q(\cdot)$ be the function describing the degradation of a quality parameter. Application of the calculation for the distribution function $G(\cdot)$ of the lifetime yields a fuzzy probability distribution for the lifetime T .

$$\Pr \{T \leq t\} = \Pr \{q(0) \leq q_{0,t}\} = \int_{-\infty}^{q_{0,t}} f_0^*(y) dy. \quad (15.14)$$

The integral is a fuzzy number whose δ -cuts $[\underline{x}_\delta; \overline{x}_\delta]$ are given, using the above notation, by

$$\left. \begin{aligned} \overline{x}_\delta &= \sup \left\{ \int_{-\infty}^{q_{0,t}} g(x) dx : g \in \mathcal{D}_\delta \right\} \\ \underline{x}_\delta &= \inf \left\{ \int_{-\infty}^{q_{0,t}} g(x) dx : g \in \mathcal{D}_\delta \right\} \end{aligned} \right\} \quad \forall \delta \in (0; 1]. \quad (15.15)$$

By this definition a fuzzy probability distribution P^* on $([0; \infty), \mathcal{B}([0; \infty)))$ is defined which is the corresponding lifetime distribution.

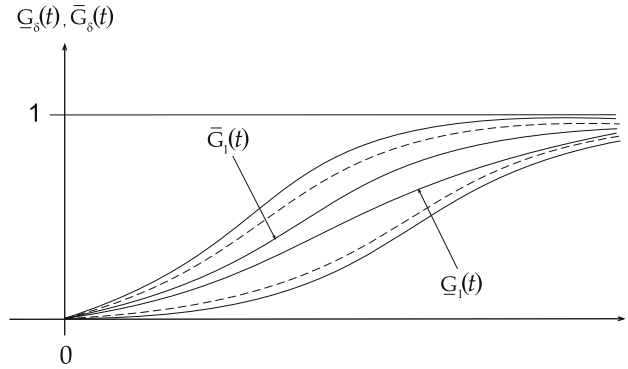


Figure 15.4. Fuzzy distribution function

The corresponding generalized distribution function $G^*(\cdot)$ is defined by its δ -cuts $\underline{G}_\delta(t) = \underline{x}_\delta$ and $\overline{G}_\delta(t) = \overline{x}_\delta \quad \forall \delta \in (0; 1]$. This so-called *fuzzy distribution function* $G^*(\cdot)$ can be graphically displayed using several δ -level functions. An example is given in Figure 15.4.

15.5 Application to Accelerated Life Testing

The differential equation (15.4) from Section 15.2 contains a temperature-dependent constant $C(S)$, where S denotes the temperature which is the stress in the context of accelerated life testing. The solution $q(\cdot)$ of Equation (15.4) depends on the stress S . This can be used for accelerated life tests. Some results are contained in (Viertl [2]). In connection with fuzzy lifetimes this is an interesting field for future research.

References

1. Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*, John Wiley & Sons, New York.
2. Viertl, R. (1981). A note on a lifetime model, *Microelectronics and Reliability*, **21**.
3. Viertl, R. (2006). Univariate statistical analysis with fuzzy data, *Computational Statistics and Data Analysis*, **51**.
4. Viertl, R. (2009). On reliability estimation based on fuzzy life-time data, *Journal of Statistical Planning and Inference*, **139**.

A New Perspective on Damage Accumulation, Marker Processes, and Weibull's Distribution

Nozer D. Singpurwalla

The George Washington University, Washington, DC, USA

Abstract: This chapter is in two interrelated parts. The two parts share the common feature that items fail when one random variable crosses a fixed or random threshold. Section II of this chapter pertains to Wallodi Weibull's model for material failure and the genesis of the famous distribution that bears his name. Section I of this chapter is more generic, because it is germane to the failure of biological entities as well. Its key feature is that a stochastic process viewpoint is adopted, and the hitting times of the process to a random threshold drive the failure phenomenon. By contrast Weibull's classic work does not take a process point of view, but is grounded in fracture mechanics; thus it has a stronger connection with engineering science. Weibull's work has some errors and unrealistic assumptions. However, it provides a platform for a synergistic enhancement of the first part of this chapter. This enhancement(s) remain to be explored.

Keywords and phrases: Crack growth, cumulative damage, degradation, deterioration, extreme value theory, failure rate, fatigue, hazard potential, marker processes, reliability, stochastic processes, weakest link principle, Wiener process

Preliminaries

Mechanical failure is the disintegration of a material specimen. It is manifested by the length of a nascent crack exceeding a threshold. A crack is the breakage of bonds which hold together a material's particles. Most cracks are microscopic and thus difficult to observe. The *strength* of a material is an abstract notion for which crack length is a proxy. Specifically, strength is inversely proportional to crack length. *Stress* is the force per unit area applied to a material specimen. But since the notion of force is unclear, stress like strength is also an abstract notion. An item fails when the stress it experiences exceeds its strength. This latter feature characterizes Weibull's work on the failure of material specimens, save for the fact that Weibull assumes a fixed stress but a random strength. This feature also provides a platform for a more universal conceptualization of the failure of items: biological, electronical, mechanical, or structural. This

is articulated in Section I of this chapter which can be seen as a general architecture for the articulation of the failure process and the biological and physical markers that it spawns.

Section I – Damage Accumulation and Marker Processes

Summary of Section I

We adopt here the viewpoint that damage is an abstract notion. An item fails when its damage exceeds a threshold. Damage cannot be observed but the surrogates (or markers) it spawns, such as crack growth, CD4 cell counts, and wear, can be. With the above in mind, we offer a probabilistic architecture that makes precise the notion of damage and its intuitive import. We look at damage as a cumulative hazard and describe its evolution as a non-decreasing stochastic process. The observable markers are also modeled by a stochastic process that is cross-correlated with the damage process. Thus a bivariate stochastic process with one component that is non-decreasing and the other that may fluctuate around some mean is a suitable model for encapsulating the phenomenon of damage and its markers. The non-decreasing process leads to the fluctuating observable process, and an item fails when the former hits a threshold. We argue that the threshold is random with an exponential(1) distribution. We call this threshold the *hazard potential* of an item.

16.1 The Hazard Potential

Let T denote the lifetime of a unit; $h(t)$ the hazard rate of $P(T \geq t)$, $t \geq 0$, and $H(t) = \int_0^t h(u) du$ be the cumulative hazard. It is easy to see that

$$P(T \geq t; h(t), t \geq 0) = \exp(-H(t)) = P(X \geq H(t)),$$

where X has an exponential(1) distribution. The random variable X is called the *hazard potential* of the item; it represents an unknown “resource” that the unit is endowed with at inception. $H(t)$ is a measure of the amount of resource consumed by time t and $h(t)$ the rate at which the resource is consumed at t . An item fails when $H(t)$ hits the random threshold X (cf. Lu and Meeker [2], Singpurwalla [4]). If the rate at which the item’s resource gets consumed is random, then $h(t)$ is described by a stochastic process, and so is $\{H(t); t \geq 0\}$. This latter process is non-decreasing. The unit fails when the process $\{H(t); t \geq 0\}$ hits a barrier X , where X is a unit exponential distribution. Candidate stochastic processes for $\{H(t); t \geq 0\}$ are alluded to in Singpurwalla [4]. Since $H(t)$ is non-decreasing in t , $t \geq 0$, it is a candidate for the damage process. Furthermore, since the conventional view claims that an item fails when the damage hits a threshold, the cumulative hazard and damage reflect a parallel feature. Like (cumulative) damage, the cumulative hazard is not observable. However, the cumulative hazard influences the time to failure.

16.2 A Stochastic Process Model for Damage and Its Markers

Because markers are linked with damage, any model for the damage process should be accompanied by a description for the markers as well. A simple way to proceed is to suppose that there is only one marker so that a bivariate stochastic process $\{H(t), Z(t); t \geq 0\}$ would be a suitable description of the damage and its marker. As stated before, the process $\{H(t); t \geq 0\}$ is non-decreasing in t . However, the process $\{Z(t); t \geq 0\}$ need not be so. Indeed, markers such as crack growth fluctuate around some trend, and thus one is free to choose any suitable model for $\{Z(t); t \geq 0\}$. A Wiener process appears to be the model of choice.

To summarize, our proposed model for the (cumulative) damage and its associated marker is a bivariate stochastic process $\{H(t), Z(t); t \geq 0\}$, with $H(t)$ non-decreasing in t and $Z(t)$ free to fluctuate around some constant or trend. We term such a process a *degradation process*. Since $H(t)$ spawns $Z(t)$, the two processes $\{H(t); t \geq 0\}$ and $\{Z(t); t \geq 0\}$ need to be linked; that is, they need to be *cross-correlated*. One way to achieve this linkage is to describe $\{Z(t); t \geq 0\}$ by a Wiener process and the unobservable (cumulative) damage process by a *Wiener Maximum Process*, namely

$$H(t) = \sup_{0 < s \leq t} \{Z(s); s \geq 0\}. \quad (16.1)$$

This strategy has been proposed in Singpurwalla [6]. The item fails when $H(t)$ hits the (random) threshold X . Whereas the model of (16.2) could be a suitable starting point, there is a caveat that needs to be addressed. Specifically, since $Z(t)$ is spawned by $H(t)$, the latter is the cause of the former. This means that $H(t)$ must lead $Z(t)$, and so any linkage between the two processes in question should incorporate a time lag. The model of (16.2) does not do this because here $H(t)$ is determined retrospective to $Z(t)$ and therefore lags $Z(t)$, instead of the other way around. Thus connecting $H(t)$ and $Z(t)$ with the observable $Z(t)$ lagging the unobservable $H(t)$ needs to be introduced. This is a possible topic for future research, and we invite potential researchers to explore the matter. An overview of some modeling strategies that have been proposed for the damage process as well as the marker process when each are treated separately can be found in Chapters 7 and 8 of Singpurwalla [5].

Section II – Weibull's Weibull Distribution

Summary of Section II

We argue that if some elementary principles of fracture mechanics are incorporated in the development of failure models for strength of materials, then a (routine) use of the Weibull distribution is questionable. We suggest possible alternatives. However, the main purpose of Section II is to offer an avenue via which the architecture of Section I can be enhanced.

16.3 Introduction

We use probabilistic methods for modeling material failure because of uncertainties about (i) applied stress (loading), (ii) material properties, and (iii) theories about the physics of failure. Statistical fracture mechanics combines the deterministic theories of failure with models for parameter uncertainty to produce a *failure model*. A failure model, used for reliability assessment and engineering design, is the specification of a probability distribution of the time to failure of an item.

To be of real value, a failure model must be validated against actual data and, if necessary, be updated (or modified) so that the model better describes the data. A coherent approach to updating is via the calculus of probability, namely Bayes' law.

Thus, the art of failure modeling is an interactive interplay between fracture mechanics, probability modeling, and data analysis. However, often, it has been data analysis that has played the dominant role with the result that predictions have fallen short of actual experiences.

16.4 The Weibull Distribution in Material Failure: Some History

The tradition of empirically fitting data to identical laboratory specimens can be traced to Weibull [7], who did introduce the notion of the *Weakest Link Theory*, but whose emphasis turned out to be data analysis. Thus, in fracture mechanics, a combination of the weakest link principle with some astute data analysis gave birth to the Weibull distribution – a distribution of smallest extremes.

However, it was not until 1948 (see [1]) that a formal justification of the Weibull distribution was provided by Epstein who adopted the Fisher–Tippett–Gnedenko Theory of Extreme Values to the Weakest Link Principle. Epstein's contribution is the first step in a major direction, namely combining fracture mechanics with known results in probability theory to produce a failure model that can be supported by actual data.

But certainly this is not the best we can do, for underlying many of the results of probability theory are assumptions (e.g., independence, finite moments), and these assumptions may not be valid in practice. Additionally, results that are shown to hold under some elementary principles of fracture mechanics may not hold when more advanced principles are invoked, or when additional principles are introduced. Thus, an automatic use of the Weibull for calculating *safety factors*, etc., needs re-examination.

16.5 Preliminaries and Notation

Let m be the number of flaws in a material specimen of unit volume, and suppose that the specimen is subjected to a tensile stress of magnitude σ_a , where σ_a is assumed

constant over time. Let S_i be the strength of the i th flaw, $i = 1, \dots, m$; we assume that the S_i 's are independent random variables. The i th flaw will cause failure of the specimen if $S_i < \sigma_a$. Let $P(S_i < x) = F_i(x)$, and assume that F_i is absolutely continuous. Let S be the strength of the specimen, assumed random, and let $P(S < x) = F(x)$, $x \geq 0$. Also, F is assumed to be absolutely continuous. The specimen fails when $S < \sigma_a$, and our goal is to assess $P(S < \sigma_a) = F(\sigma_a)$.

16.6 The Weakest Link Principle and an Application

The Weakest Link Principle (WLP) says that “The strength of the material is the strength of its weakest flaw,” i.e., $S = \min(S_i)$, $i = 1, \dots, m$.

According to the WLP, the reliability of the specimen, under an application of constant load σ_a , is given by

$$\begin{aligned} P(S \geq \sigma_a) &= P\left(\min_i S_i \geq \sigma_a\right) \\ &= P(S_1 \geq \sigma_a, \dots, S_m \geq \sigma_a) \\ &= \prod_{i=1}^m P(S_i \geq \sigma_a), \end{aligned}$$

since the S_i 's are assumed independent. But,

$$P(S_i \geq \sigma_a) = 1 - P(S_i < \sigma_a) = 1 - F_i(\sigma_a).$$

Therefore,

$$P(S \geq \sigma_a) = \prod_{i=1}^m (1 - F_i(\sigma_a)).$$

Thus, under the WLP, with constant loading σ_a , the *unreliability* of the specimen is

$$F(\sigma_a) = P(S < \sigma_a) = 1 - P(S \geq \sigma_a) = 1 - \prod_{i=1}^m (1 - F_i(\sigma_a)),$$

which is the probability that the specimen fails under stress σ_a . If we also assume that the S_i 's are identically distributed, i.e., $F_i(\sigma_a) = G(\sigma_a)$, $i = 1, \dots, m$, then

$$\overline{F}(\sigma_a) = (1 - G(\sigma_a))^m,$$

where $\overline{F} = 1 - F$.

16.7 Weibull's Approximation and Analysis

Weibull arrived at the famous distribution named after him by invoking the WLP and then using a questionable approximation to which he added some empirical results. His approximation was that

$$\begin{aligned}
 F(\sigma_a) &= P(S < \sigma_a) = 1 - \prod_{i=1}^m (1 - F_i(\sigma_a)) \\
 &\approx 1 - \prod_{i=1}^m \exp[-F_i(\sigma_a)] \\
 &= 1 - \exp\left[-\sum_{i=1}^m F_i(\sigma_a)\right].
 \end{aligned} \tag{16.2}$$

Weibull referred to $\sum_{i=1}^m F_i(\sigma_a)$ as the *risk of rupture* and, based on an empirical fit of failure data on laboratory specimens, simply set $\sum_{i=1}^m F_i(\sigma_a) = \left(\frac{\sigma_a}{\eta}\right)^\beta$. Thus he argued that for $\beta, \eta > 0$,

$$P(S < \sigma_a | \eta, \beta) \approx 1 - \exp\left[-\left(\frac{\sigma_a}{\eta}\right)^\beta\right],$$

for $\sigma_a \geq 0$. Note that this is the cumulative distribution function for the Weibull distribution.

Weibull then argued that if the material specimen is of volume V , then a repeated application of the WLP gives, for the constant application of a tensile stress $\sigma_a \geq 0$, the result that

$$F(\sigma_a) = P(S < \sigma_a | \eta, \beta) = 1 - \exp\left[-V \left(\frac{\sigma_a}{\eta}\right)^\beta\right]. \tag{16.3}$$

Weibull's approximation, namely

$$\prod_{i=1}^m (1 - F_i(\sigma_a)) \approx \exp\left[-\sum_{i=1}^m F_i(\sigma_a)\right],$$

for all $\sigma_a \geq 0$, will hold if two conditions are met, namely

- (i) $P(S_i < \sigma_a) = F_i(\sigma_a)$ is infinitesimally small, for all i and any $\sigma_a \geq 0$. This does not seem reasonable.
- (ii) m is small to moderate, a questionable hypothesis.

In other words, Weibull's choice of the Weibull distribution as a failure model is based on the above questionable assumptions.

16.8 Critique of Weibull's Analysis

Recall Weibull's result:

$$\begin{aligned}\bar{F}(\sigma_a) &\approx \exp \left[-\sum_{i=1}^m F_i(\sigma_a) \right] \\ &= \exp [-mG(\sigma_a)],\end{aligned}$$

when the S_i 's are identically distributed. But if h is the *failure rate* of F , then it is a standard result of reliability theory (cf. Singpurwalla [5]) that

$$\bar{F}(\sigma_a) = \exp \left[-\int_0^{\sigma_a} h(u) du \right].$$

Thus, we may look upon $G(\sigma_a)$ as the cumulative failure rate of F and its derivative as the failure rate of F .

In other words, Weibull's result implies that the probability density of the distribution of flaw strengths is the failure rate of the distribution of specimen strength. Consequently, if the flaw strength distribution is unimodal (like that of a Weibull), then the failure rate of the specimen strength cannot be monotonic. Thus, if the flaw strengths are assumed Weibull (with shape parameter $\beta \neq 1$), then the specimen strength cannot be a Weibull. This conclusion appears to be unrecognized in engineering practice. An exception is that if the flaw strengths are assumed exponential, then the specimen strength can be a Weibull with a decreasing failure rate.

Also recall that, based on empirical considerations, Weibull set

$$\sum_{i=1}^m F_i(\sigma_a) = \left(\frac{\sigma_a}{\eta} \right)^\beta,$$

for $\beta, \eta > 0$.

Is this choice meaningful? Let

$$F^*(\sigma_a) = \frac{1}{m} \sum_{i=1}^m F_i(\sigma_a).$$

Note that this is an average. Then

$$mF^*(\sigma_a) = \sigma_a^\beta \left(\frac{1}{\eta} \right)^\beta,$$

and we may identify $F^*(\sigma_a)$ with σ_a^β , but $F^*(\sigma_a) \in [0, 1]$ and $\sigma_a \geq 0$, which implies that β cannot be > 0 . Therefore, Weibull's empirical analysis cannot be correct; it results in an incoherence.

16.9 The Theory of Extreme Values

Recall that when $F_i(\sigma_a) = G(\sigma_a)$, for all i , then the WLP gives us the result that

$$\bar{F}(\sigma_a) = P(S \geq \sigma_a) = P\left(\min_i S_i \geq \sigma_a\right) = (1 - G(\sigma_a))^m.$$

The theory of extreme values (EVT) tells us what happens to $\bar{F}(\sigma_a)$ when $m \rightarrow \infty$. According to EVT, there exist constants C_m and d_m , which depend on m and also on G , such that as $m \rightarrow \infty$, and for $\sigma_a \notin (-\infty, 0)$,

$$P\left(\min_i S_i \geq C_m + d_m \sigma_a\right) = \begin{cases} \exp(-\sigma_a^\beta), & \sigma_a \geq 0, \beta > 0 \\ \exp(-e^{\sigma_a}), & -\infty \leq \sigma_a \leq \infty \end{cases}.$$

If the first of the above two expressions holds, then for $m \rightarrow \infty$,

$$F(\sigma_a) = 1 - \exp\left[-\left(\frac{\sigma_a - C_m}{d_m}\right)^\beta\right], \quad (16.4)$$

$\sigma_a \geq 0, \beta > 0$.

This expression parallels Weibull's result which was based on an approximation and an empirical analysis; see (16.3).

But there are some features of (16.4) which make it more attractive than Weibull's derivation (16.3):

- (i) The effect of m , even though it is infinite, is captured in the norming constants C_m and d_m .
- (ii) The power law relationship σ_a^β was motivated by Weibull via empirical considerations, but in (16.4), namely $\left(\frac{\sigma_a - C_m}{d_m}\right)^\beta$, it is a consequence of the EVT.

But are (16.3) and (16.4) supportive of other considerations of fracture mechanics or are they just a manifestation of our assumptions and approximations?

16.10 Comments on Using the Weibull Distribution Motivated via WLP and EVT

We have seen that Weibull's approximation implies a relationship between flaw strength, density, and specimen failure rate that is restrictive (if not unacceptable), and his empirical analysis implies a choice that is incoherent. But the Weibull distribution can also be motivated by the EVT, and thus one may argue that the consequences mentioned above have to do with Weibull's approximation, which de facto ignores terms of order $(F_i(\sigma_a))^2$ and above. But a derivation of the Weibull via the EVT is justified only when

$$G(\sigma_a) = F_i(\sigma_a),$$

for all i , is of a certain form. Technically speaking, G belongs to the “Domain of Attraction of . . .”

In Singpurwalla et al. [3], we argue, using the fact that $S_i \propto (a_i)^{-1}$, where a_i is the length of the i th crack (flaw), that the Weibull distribution for specimen strength can arise iff the a_i 's are themselves i.i.d. Weibull.

Whereas a Weibull distribution for crack lengths can be argued (using microstructure), the independence of crack lengths is hard to accept. There has been much discussion about the *Journal of the American Statistical Association (JASA)* missing the boat by not accepting Weibull's paper. We do not know the grounds on which the paper was rejected, but if it was because of the underlying analysis, then JASA did not miss the boat!

Acknowledgments. This chapter has been written on the occasion to commemorate Bill Meeker's 60th birthday. Bill's contributions to accelerated life testing and the analysis of degradation data are ground breaking, noteworthy, and legendary. Supported by The Army Research Office Grant W911NF-09-1-0039.

References

1. Epstein, B. (1948). Statistical aspects of fracture problems, *Journal of Applied Physics*, **19**, 140–147.
2. Lu, C.J. and Meeker, W.Q. (1993). Using degradation measures to estimate a time-to-failure distribution, *Technometrics* **35**, 161–174.
3. Singpurwalla, N.D., Wilson, S., and Fuller, E. (1996). Statistical Aspects of Failure Processes in Ceramics, in *Bayesian Statistics 5* J. Berger, J. Bernardo, P. David, and A. Smith (eds), pp. 369–386, Clarendon Press, Oxford.
4. Singpurwalla, N.D. (2006a). The hazard potential: introduction and overview, *Journal of the American Statistical Association*, **101**, 1705–1717.
5. Singpurwalla, N.D. (2006b). *Reliability and Risk: A Bayesian Perspective*, John Wiley & Sons, New York.
6. Singpurwalla, N.D. (2006c). On Competing risk and degradation processes, in *The Second Erich L. Lehman Symposium – Optimality, Institute of Mathematical Statistics – Monograph Series, Vol. 49*, J. Rojo (ed.), pp. 289–304. The Institute of Mathematical Statistics.
7. Weibull, W. (1939). A statistical theory of the strength of materials, *Ingeniors Vetenskaps Akademien, Handlingar*, **151**, 45–55.

Reliability Estimation and ALT

Reliability Estimation of Mechanical Components Using Accelerated Life Testing Models

Fabrice Guérin, M. Barreau, A. Charki, A. Todoskoff, S. Cloupet and D. Bigaud

Institut des Sciences et Techniques de l'Ingénieur d'Angers, Angers, France

Abstract: This chapter presents an overview of using accelerated life testing (ALT) models for reliability estimation on mechanical components. The reliability is estimated by considering two test plans: a classical one testing a sample system under accelerated conditions only and a second plan with previous accelerated damage. The principle of the test plan with previous accelerated damage is testing the sample under step-stress. In the beginning (until time N_1), the sample is tested under stress s_1 (accelerated testing: $s_1 > s_0$); when the tested units have used many of their “resources,” the stress s_1 is replaced by the operating conditions s_0 (until the time N_2). Therefore, failure times under the accelerated conditions can be used to estimate reliability function in operating conditions. The time transformation function is considered as log-linear and four types of estimation are studied: parametric, Extended Hazard Regression (GPH), semi-parametric, and nonparametric models. The chapter is illustrated by a simulation example of ball bearings testing. The results are used to analyze and compare these estimation methods. The simulations have been performed both with censored data and without censoring, in order to examine the asymptotic behavior of the different estimates.

Keywords and phrases: Reliability, parametric estimation, extended hazard regression model, semi-parametric estimation, nonparametric estimation, regression, Kaplan–Meier, ball bearings

17.1 Introduction

Since devices are more and more reliable, measuring their reliability becomes more and more difficult: a large number of devices need to be tested for an extensive amount of time under normal operating conditions in order to obtain accurate measures of their reliabilities. One commonly used approach for estimating such systems reliability in a much shorter time is accelerated life testing (ALT). ALT is achieved by applying on test units stresses which are more severe than the normal operating conditions [1–3]. The results obtained under accelerated conditions are then extrapolated in order

to estimate the reliability functions for normal operating conditions. Some inference methods have been developed assuming that (1) the reliability function in a constant stress environment belongs to a common distribution family and (2) the scale parameter of such a distribution is related to the stress environment via a parametric function known as TTF (time transformation function) [1, 2, 4, 5].

Later studies use the proportional hazard (PH) model [1, 2, 4, 6]. The PH model is based on a reliability function defined by the hazard function decomposed into the baseline hazard function and the function dependent on the covariates. PH model can be used to extrapolate in stress but not in time. Therefore, Bagdonavicius and Nikulin [17] and Etezadi-Amoli and Ciampi [7] have introduced the generalized proportional hazard (GPH) model, and Etezadi-Amoli and Ciampi propose to use the quadratic splines method to approximate the baseline hazard rate allowing thus extrapolating in stress and in time.

Other papers propose the use of semi-parametric ALT models, where the reliability function is unknown (e.g., defined by Kaplan–Meier estimator) and TTF is parameterized [4, 8, 9].

However, the regression test plans require a good knowledge of the model, which is rather difficult to obtain. Therefore, Bagdonavicius and Nikulin [15, 17] have proposed a new test plan to reduce these difficulties. This test plan, with previous accelerated damage, consists in testing a sample under step-stress. At first (until time T_1), the sample is tested under stress s_1 , with $s_1 > s_0$ (accelerated testing) and when the tested units have used many of their “resources,” the stress s_1 is replaced by the normal conditions s_0 (until the time T_2). Failure times under the accelerated conditions can be used to estimate reliability function for the operating conditions.

This chapter presents an overview of using ALT models for mechanical components, by defining a regression test plan and a test plan with previous accelerated damage. For the case of the regression test plan, the parametric, GPH, and semi-parametric models with a log-linear TTF model are presented. For the case of the test plan with previous accelerated damage, the parametric and nonparametric models are studied. First, Section 17.2 provides a brief background on ALT model. Then, Section 17.3 presents the regression test plan and the parametric, GPH, and semi-parametric estimates and illustrates them by a simulation example. Finally, in Section 17.4, the test plan with previous accelerated damage and the parametric and nonparametric estimates are studied, before some concluding remarks.

17.2 Accelerated Life Testing Model

ALT model is generally used when the exact relationship between the applied stresses and the failure time of the component (or product) is difficult to determine only on physics bases [1]. In this case, components are tested at different accelerated stress levels. The failure times at the different stress levels are then used to determine the most appropriate reliability function, along with its parameter(s).

The most commonly used failure time probability distributions are exponential, Weibull, and lognormal. The reliability function is expected to be the same at different stress levels, as well as for the normal operating conditions. For practical purposes,

the failure times at different stress levels are assumed to be linearly related to each other. The timescale transformation is constant, implying that we have a true linear acceleration. Based on this assumption (see Figure 17.1), the reliability functions at normal operating R_{z_0} conditions and accelerated conditions R_z are related as

$$R_z(t) = R_{z_0}(TTF \times t), \quad (17.1)$$

where TTF is the time transformation function.

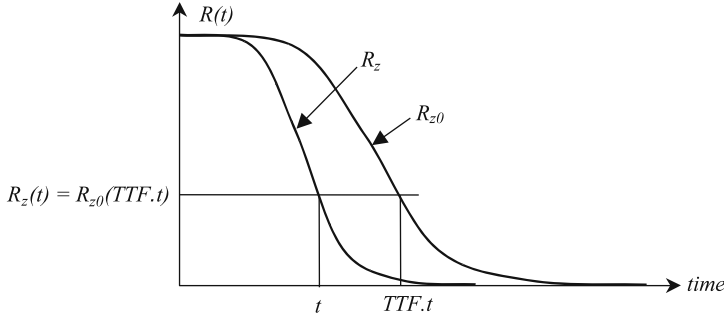


Figure 17.1. Definition of time transformation function

In ALT model, the shape parameters of the distribution are expected to be the same for all stress levels. To define the TTF , the log-linear models are the most used regression models [1–4]. TTF specifies that the effect of the covariates is a multiplicative factor for the hazard rate, a scale change, or location shift for the reliability function, respectively. For some constant stress z , the relationship (17.1) becomes

$$R_z(t) = R_{z_0} \left(e^{\beta^T z} \cdot t \right), \quad (17.2)$$

where $\beta = (\beta_0, \dots, \beta_m)^T$ is a vector of unknown parameters and $z = (z_0, \dots, z_m)$ is a vector of covariates (with $z_0 = 1$).

In the Appendix different log-linear models used to define the TTF are presented.

When the failure data conceal a complex lifetime distribution, proportional hazard (with the generalized proportional hazard (GPH) model) and semi parametric models appear to be attractive approaches [10]. The main advantage of these models is that they are essentially distribution-free.

However, the regression test plans require a good knowledge of the model, which is rather hard to obtain. Therefore, Bagdonavicius and Nikulin [15, 17] have proposed a new test plan with previous accelerated damage to reduce these difficulties. The authors have studied the parametric and nonparametric estimates.

In following sections, the regression test plan and the test plan with previous accelerated damage are applied to mechanical components, to ball bearings in particular, for studying the parametric, GPH, semi-parametric, and nonparametric estimates.

17.3 Regression Test Plan

17.3.1 Introduction

ALT models can be applied to various fields such as electronics, electromechanics, or mechanics [1–3, 11, 12] and are used for mechanical components and assemblies such as automobile parts, hydraulic components, tools, and gears. Acceleration stresses include mechanical load, vibration, temperature or other environmental factors, and combinations of such stresses. The applications of these models depend on the life distributions (for example, in mechanics time is expressed in number of cycles N), on the accelerated stress, and on the type of the regression model given by the relationship (17.2). Thus, in mechanics it is common to consider the Weibull, Birnbaum–Saunders, or Lognormal distributions as underlying life distributions [13, 14], and Weibull distribution is considered as correctly describing the lifetime of mechanical components.

The studied test plan (Figure 17.2) determines the parameters of the TTF ($e^{\beta^T z}$) by tests performed only under accelerated conditions. The TTF is used to estimate the reliability function in operating conditions.

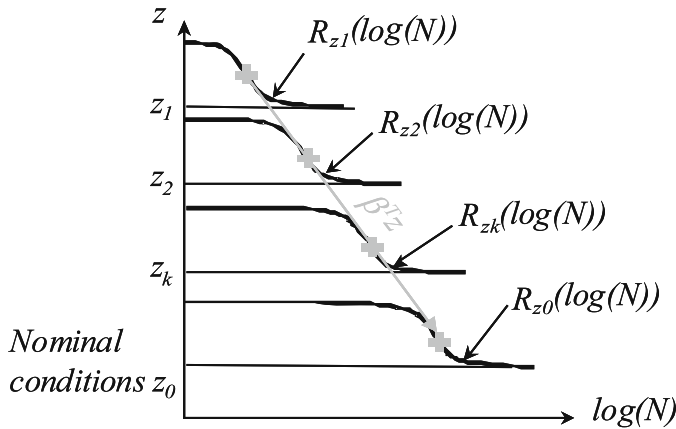


Figure 17.2. Definition of the studied test plan

The tests are performed under accelerated stresses z_1, z_2, \dots, z_n , while estimating vector β (Equation 17.2). Consequently, an estimate of the reliability function R_{z_0} can be computed.

Assume a test with fixed maximal time N_i (in number of cycles) of i th sample (of size n_i) under accelerated stresses z_i (with $i = 1, \dots, k$ and $z_i = z_{i0}, \dots, z_{im}$) and let m_i be the number of failures.

Three types of reliability estimates illustrate this test plan: the parametric model, the GPH model, and the semi-parametric model.

17.3.2 Parametric ALT model

The parametric estimate of the ALT model consistently chooses a common probability distribution in order to estimate the reliability function [1–3]. The Weibull distribution is widely used for life data including mechanical damage. The reliability function $R_{z0}(u)$ for the Weibull distribution of product life u is defined as

$$R_{z0}(u) = e^{-\left(\frac{u}{\eta}\right)^\nu}, \quad (17.3)$$

where ν and η are the distribution parameters and u is the number of cycles.

Then, for a constant stress z , the reliability function, defined by Equation (17.2), is given by

$$R_z(u) = e^{-\left(\frac{r(s) \cdot u}{\eta}\right)^\nu} = e^{-\left(\frac{e^{\beta T z} \cdot u}{\eta}\right)^\nu} = e^{-\left(\frac{u}{e^{-\beta T z}}\right)^\nu}, \quad (17.4)$$

with $\beta = (\beta_0, \dots, \beta_m)$, $\beta_0 = \log(\eta)$.

For example, for a test with fixed maximal time N_i (in number of cycles) of i th sample (of size n_i), under accelerated stresses z_i (with $i = 1, \dots, k$ and $z_i = z_{i0}, \dots, z_{im}$), with m_i failures, the sample likelihood is defined by

$$L(\beta, \nu) = \prod_{i=1}^k \left(\prod_{j=1}^{m_i} \frac{\nu}{e^{-\beta T z}} \left(\frac{N_{ij}}{e^{-\beta T z}} \right)^{\nu-1} e^{-\left(\frac{N_{ij}}{e^{-\beta T z}} \right)^\nu} \right) \times \left(e^{-\left(\frac{N_i}{e^{-\beta T z}} \right)^\nu} \right)^{n_i - m_i}, \quad (17.5)$$

with N_{ij} the number of cycles to failure under stress z_i , N_i the censored number of cycles to failure (Figure 17.3).

The maximum likelihood estimates (MLE) of vector β and shape parameter ν are obtained. The reliability function under operating conditions is estimated by

$$\hat{R}_{z0}(u) = e^{-\left(\frac{u}{e^{-\hat{\beta} T z_0}} \right)^{\hat{\nu}}}, \quad (17.6)$$

where $\hat{\beta}$ and $\hat{\nu}$ are the maximum likelihood estimators and z_0 is stress vector under operating conditions (Figure 17.2).

17.3.3 Generalized proportional hazard (GPH) model

The proportional hazard (PH) model (introduced by Cox in 1972) is well known in survival analysis [1, 4, 6, 14, 15]. The PH model does not assume a form of the life distribution. The PH model can be used to extrapolate in stress but not in time, because it is distribution-free. The PH model involves more than one accelerating stress and is considered to be a multidimensional model. The PH model for the failure (hazard) rate function of the distribution at stress $z(\cdot)$ is

$$\lambda(u, z) = \lambda_0(u) e^{\beta^T z}, \quad (17.7)$$

with $\lambda_0(u)$ the baseline failure rate and u the lifetime.

Then, the reliability is defined as

$$R_{s0}(u) = e^{-\int_0^u \lambda(x, z) e^{-\beta^T z} dx}. \quad (17.8)$$

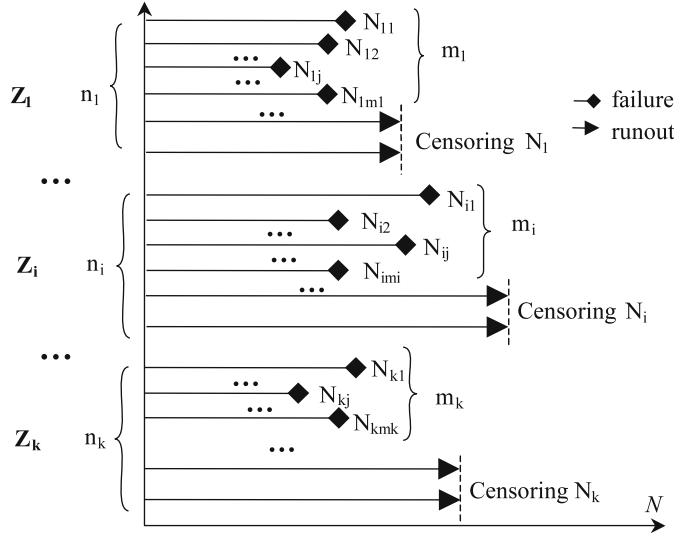


Figure 17.3. Record of failure times

Shyur, Elsayed, and Luxhoj [6] and Bagdonavicius and Nikulin [17] have extended the PH model and proposed the generalized proportional hazard (GPH) model. The general expression of the GPH model is written as follows:

$$\lambda(u, z) = g_1(\alpha^T z) \lambda_0(g_2(\beta^T z) \cdot u), \quad (17.9)$$

where u is the number of cycles

$Z = (z_1, z_2, \dots, z_m)$: vector of the applied stress (covariants)

$\lambda(u, Z)$: hazard failure rate for a given Z

$\lambda_w(\cdot)$: baseline hazard rate

$g_1(x)$ and $g_2(x)$: positive functions equal to 1 at $x = 0$

α and β : vectors of regression coefficients.

In this model, λ_0 , g_1 , and g_2 involve unknown parameters and $\lambda_0(\cdot)$ is an arbitrary and unspecified baseline hazard function. Etezadi, Amoli and Ciampi [9] use the quadratic splines method to approximate the baseline hazard rate $\lambda_0(\cdot)$. Etezadi-Amoli and Ciampi [9] introduce the GPH model and assume that $g_1(x) = g_2(x) = \exp(x)$, which reduces this model to the proportional hazards when $\beta = 0$ and to the accelerated failure time model when $\alpha = \beta$. The maximum likelihood method is used to evaluate the unknown parameters of GPH models. Supposing a test with fixed maximal time N_i (in number of cycles) of i th sample (of size n_i), under accelerated stresses s_i (with $i = 1, \dots, k$ and $z^{(i)} = z_{i0}, \dots, z_{im}$) and m_i the number of failures, the log-likelihood function can be written as

$$L(\beta, \nu) = \sum_{i=1}^k \left(\sum_{j=1}^{m_i} \left(\alpha^T z + \log \left(\lambda_0 \left(e^{\beta^T z} \cdot N_{ij}, z \right) \right) - \frac{e^{\alpha^T z}}{e^{\beta^T z}} \lambda_0 \left(e^{\beta^T z} \cdot N_{ij}, \gamma \right) \right) \right. \\ \left. - (n_i - m_i) \cdot \left(\frac{e^{\alpha^T z}}{e^{\beta^T z}} \lambda_0 \left(e^{\beta^T z} \cdot N_{ij}, \gamma \right) \right) \right), \quad (17.10)$$

with N_{ij} the number of cycle to failure under stress s_i , N_i the censored number of cycle to failure, $\Lambda_0(u, z) = \int_0^u \lambda_0(t, z) dt$ cumulative hazard function, γ denotes the vector of parameters defining λ_0 when a spline function is used.

The reliability function under operating conditions is defined by

$$\hat{R}_{z_0}(u) = e^{-e^{\hat{\alpha}^T z_0} \cdot \Lambda_0(e^{\hat{\beta}^T z_0} \cdot u, z)}, \quad (17.11)$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the maximum likelihood estimates and z_0 is the stress vector in operating conditions.

17.3.4 Semi-parametric ALT model

In the semi-parametric ALT model, the form of the reliability function is unknown and the TTF is parameterized by $TTF = e^{\beta^T z}$ [4, 7, 8, 16, 17]. The Kaplan–Meier estimator gives an estimate of the reliability function.

$K_i(l)$ is the number of failures in sample i during the interval $[0, l]$ and $S_i(l)$ the number of surviving samples immediately before l . N_{i1}, \dots, N_{imi} represent the number of cycles to failure and N_i the censored number of cycles (with $m_i = K_i(N_i)$) for sample i , under stress s_i .

The random variables, $e^{\beta^T z_i} \cdot N_{ij}$ ($i = 1, \dots, k; j = 1, \dots, m_i$) can be considered as observed pseudo-failures (time converted into equivalent time under operating time) in tests where $n = \sum_{i=1}^k n_i$ systems (with baseline reliability function R_{s0}) have been tested and $n_i - m_i$ censored systems at equivalent censoring time $e^{\beta^T z_i} \cdot N_i$ are observed.

The number of failures in interval $[0, l]$ is defined by

$$K(l, \beta) = \sum_{i=1}^k K_i(e^{-\beta^T z_i} l) \quad (17.12)$$

and the surviving number immediately before l :

$$S(l, \beta) = \sum_{i=1}^k S_i(e^{-\beta^T z_i} l). \quad (17.13)$$

The reliability function is estimated using Kaplan–Meier point estimate: for all $u \leq \max_i(e^{\beta^T z_i} l_i)$ where $l_i = N_i$

$$\hat{R}_{z_0}(u, \beta) = \prod_{l \leq u} \left(1 - \frac{\Delta K(l, \beta)}{S(l, \beta)} \right), \quad (17.14)$$

where $\Delta K(l, \beta) = K(l, \beta) - K(l-, \beta)$.

The sample likelihood function is defined as

$$L(\beta) = \prod_{i=1}^k \prod_{j=1}^{m_i} \left[\hat{R}_{s_0}(e^{\beta^T z_i}(N_{ij}-), \beta) - \hat{R}_{s_0}(e^{\beta^T z_i}(N_{ij}), \beta) \right] \times \left(\hat{R}_{s_0}(e^{\beta^T z_i} l_i, \beta) \right)^{n_i - m_i}. \quad (17.15)$$

The unknown density function $f_{s_0} = -R_{s_0}$ is approximated by

$$R_{z_0}(e^{\beta^T z_i}(N_{ij}-), \beta) - R_{z_0}(e^{\beta^T z_i}(N_{ij}), \beta). \quad (17.16)$$

The reliability function under operating conditions is estimated with MLE value $\hat{\beta}$.

17.3.5 Application to ball bearings

In this section, the ALT models are applied to ball bearings, and a simulation model is proposed to observe reliability testing results.

Ball bearing life

Rolling element bearings are used in an extremely wide variety of tools, like machinery and equipment. They may be used in fans, gear boxes, transmissions, axles, compressors, electric motors, engines, final drives, jet engine main shafts, blenders, saws, mixers, and so on. Most rotating shafts use a rolling element bearing. Geisler et al. recently patented bearing with fluid flow bypass [18]. Kuwabara has recently patented rolling device employing lubricating grease composition and electric power steering apparatus employing the rolling device [19].

Ball bearing life refers to the amount of time a bearing will perform a specified operation before failure. Bearing life is commonly defined in terms of L_{10} life, which is sometimes referred to as B_{10} . This is the lifetime which 90% of identical bearings subjected to identical usage applications and environments will attain (or surpass) before bearing material fails from fatigue. In practice, life (in million of revolutions) given as a function of load is represented with Palmgren's relationship [1] for the L_{10} of the life distribution:

$$L_{10} = \left(\frac{C}{P} \right)^p, \quad (17.17)$$

where C is the basic dynamic load rating

P is the equivalent radial load

$p = 3$ for ball bearings and $p = 10/3$ for rolling bearings.

Bearing life is usually defined by the Weibull distribution, which is combined with the Palmgren's relationship:

$$R(u, P) = e^{-0.105 \left(\frac{u}{L_{10}} \right)^{1.5}}. \quad (17.18)$$

The shape parameter ν 1.5 is used and the scale parameter η is given by

$$\eta(P) = L_{10} \left(\frac{1}{0.105} \right)^{\frac{1}{1.5}} \quad (17.19)$$

or scale parameter η_H in hours:

$$\eta_H(P) = L_{10} \frac{60.N}{10^6} \left(\frac{1}{0.105} \right)^{\frac{1}{1.5}}, \quad (17.20)$$

with N the speed in revolution per minute.

Simulation model definition of testing result

For this chapter, the times to failure are drawings from a Weibull distribution. The ball bearing is characterized by the dimensional parameters given in Table 17.1.

To illustrate the ALT model, the load P is used to accelerate the test. The simulation parameters are defined in Table 17.2.

From these simulation parameters, the values N_{ij} are simulated from following Weibull distributions:

Table 17.1. Ball bearing parameters


	$D = 32$ mm : outside diameter
	$d = 15$ mm: inside diameter
	$B = 8$ mm: width bearing
	$C = 400$ daN: basic dynamic load rating
	$N = 1500$ tr/min: operating rotation shaft speed

Table 17.2. Simulation parameters

	Load P in daN	Scale parameter η (eq 23.) in hours	Shape parameter ν
Operating conditions	125	1635.86	1.5
Test 1	150	946.68	1.5
Test 2	175	596.16	1.5

Table 17.3. Simulation results (N in hours)

	1	2	3	4	5	6	7	8	9	10
Test 1	156.5	280.4	378.3	421.5	433.2	558.4	601.8	699.1	716.4	726.9
Test 2	499.5	534.9	794	808.4	825.4	834.3	868.8	1126.2	1143.5	1492.2
	11	12	13	14	15	16	17	18	19	20
Test 1	749.6	778.8	800.9	884.9	1061.5	1229.9	1269.3	1468.9	1804.2	1910.1
Test 2	499.5	534.9	794	808.4	825.4	834.3	868.8	1126.2	1143.5	1492.2

$$\text{Test 1: } N_{1j} \propto W(946.68, 1.5) \quad \text{Test 2: } N_{2j} \propto W(596.16, 1.5)$$

The simulation results are given in Table 17.3.

Testing result analysis

The parametric, GPH, and semi-parametric ALT models are used to analyze the accelerated test data. Table 17.4 gives the model definitions and the MLE values of unknown parameters.

Finally, the reliability function under operating conditions ($P_0 = 125$ daN) can be estimated by parametric, GPH, and semi-parametric models (see Figure 17.4).

One can notice a good correlation of the parametric, GPH, and semi-parametric models with the reference model.

Asymptotic behavior study

In order to analyze the asymptotic behavior of the different points estimate, the Monte Carlo simulation is repeated several times (20 simulations with samples sizes $n_1 = n_2 = 100$ units). Parametric, GPH, and semi-parametric points estimates of the unknown parameters are characterized by their mean and their standard deviation (Table 17.5).

In each simulation, the reliability function $R_{s0}(N)$ under operating conditions is plotted using the parametric (Figure 17.5), GPH (Figure 17.6), and semi-parametric (Figure 17.7) estimations.

Table 17.4. MLE estimations of model parameters

	Model definition	MLE estimates
Parametric ALT model	$R_P(u) = e^{-\left(\frac{u}{e^{\beta_0 + \beta_1 \log(P)}}\right)^\nu}$	$\hat{\beta}_0 = 18.2422$ $\hat{\beta}_1 = -2.2758$ $\hat{\nu} = 1.7086$
GPH model	$R_P(u) = e^{-(\gamma_0 u + \gamma_1 u^2)} e^{\beta \log(P)}$	$\hat{\gamma}_0 = 4.1018\text{e-}006$ $\hat{\gamma}_1 = 4.1314\text{e-}008$ $\hat{\beta} = 0.0252$
Semi-parametric model	$\hat{R}_{P_0}(u, \beta) = \prod_{l \leq u} \left(1 - \frac{\Delta K(l, \beta)}{S(l, \beta)}\right)$ where $\Delta K(l, \beta) = K(l, \beta) - K(l-, \beta)$ with $K(l, \beta) = \sum_{i=1}^k K_i(e^{-\beta_0 - \beta_1 \log(P)} l)$ $S(l, \beta) = \sum_{i=1}^k S_i(e^{-\beta_0 - \beta_1 \log(P)} l)$	$\hat{\beta}_0 = -1.8746\text{e-}006$ $\hat{\beta}_1 = 2.8816$

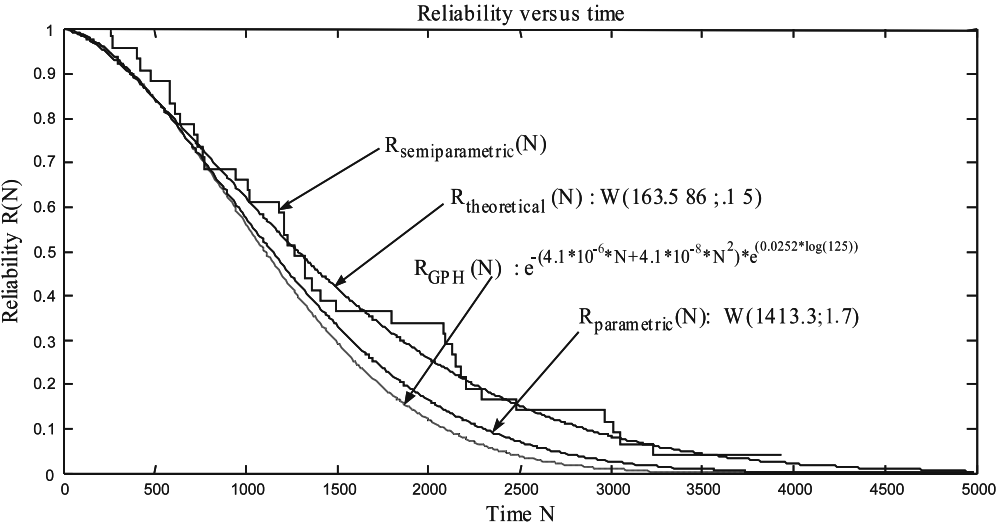


Figure 17.4. Reliability functions by the different ALT models (N in hours)

Table 17.5. Mean and standard deviation of MLE estimations after 20 repetitions

		Mean	Standard deviation
Parametric model	$\hat{\beta}_0$	20.9797	2.9695
	$\hat{\beta}_1$	-2.8224	0.5877
	$\hat{\nu}$	1.4861	0.0989
	$\hat{\eta}$	1.5743e+003	216.6656
GPH model	$\hat{\gamma}_0$	7.0053e-006	7.2418e-006
	$\hat{\gamma}_1$	2.1232e-008	1.9557e-008
	$\hat{\beta}$	0.0301	0.0060
Semi-parametric model	$\hat{\beta}_0$	-3.3203e-004	5.3523e-004
	$\hat{\beta}_1$	-651.8708	1.0265e+003

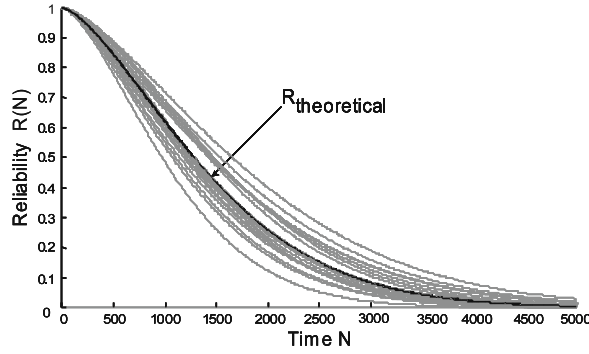


Figure 17.5. Reliability function by parametric model (N in hours)

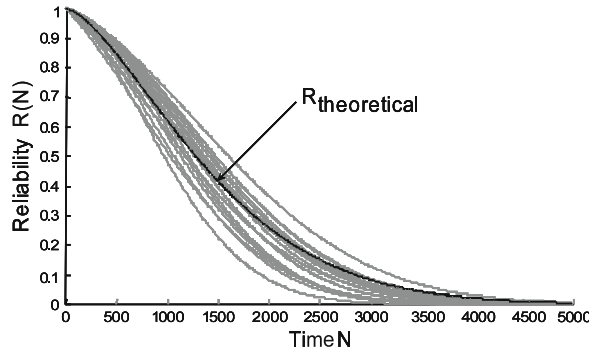


Figure 17.6. Reliability function by GPH model (N in hours)

In previous applications, all units were considered as having failures. In the following simulation (using the same simulation parameters given Table 17.2 and sample sizes $n_1 = n_2 = 100$), the parametric, GPH, and semi-parametric estimations are studied with censored data. The simulation is repeated to study the asymptotic behavior of the estimators. Two cases are studied corresponding to two censoring times (1,000 hours and 300 hours, respectively). The unknown parameters of parametric, GPH, and semi-parametric models are characterized by their mean and their standard deviation (Table 17.6).

At each repeating of simulation, the reliability function $R_{s_0}(N)$ under operating conditions is plotted using the parametric (Figure 17.8), GPH (Figure 17.9), and semi-parametric (Figure 17.10) estimations.

In the case of small sample sizes, it is rather difficult to compare the variances of the different estimators from a theoretical point of view. Twenty samples of the Weibull

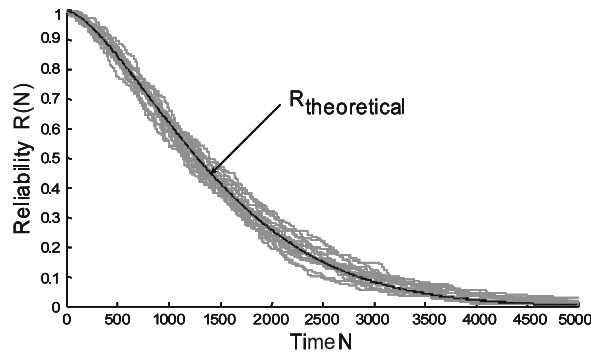


Figure 17.7. Reliability function by semi-parametric model (N in hours)

Table 17.6. Mean and standard deviation of MLE estimations after 20 repeatings for different censoring times

		Censoring at 1,000 hours		Censoring at 300 hours	
		Mean	Standard deviation	Mean	Standard deviation
Parametric model	$\hat{\beta}_0$	22.6286	2.7090	22.8761	6.1846
	$\hat{\beta}_1$	-3.1473	0.5342	-3.1901	1.1973
	$\hat{\nu}$	1.5329	0.0774	1.5128	0.2568
	$\hat{\eta}$	1.7046e+003	235.1529	1.9386e+003	928.2192
GPH model	$\hat{\gamma}_0$	8.8955e-006	1.0058e-005	4.3653e-005	1.7451e-004
	$\hat{\gamma}_1$	3.7524e-008	4.6808e-008	4.1147e-007	1.3179e-006
	$\hat{\beta}$	0.0292	0.0078	0.0313	0.0158
Semi-parametric model	$\hat{\beta}_0$	-3.5321e-004	5.6398e-004	-5.9933e-005	2.5866e-004
	$\hat{\beta}_1$	-654.6965	1.0307e+003	-107.6598	494.8533

distribution were generated in order to estimate the variance, and the corresponding parametric, GPH, and semi-parametric points estimates have been deduced.

Analysis

One can notice a good correlation of the parametric, GPH, and semi-parametric models with the reference model (Figures 17.8, 17.9 and 17.10), considering both complete and censored data.

However, this type of test plan by regression requires knowledge of the function TTF , which is not always obvious, in particular when many variables are used to accelerate the tests. Therefore, tests should be performed with various combinations of

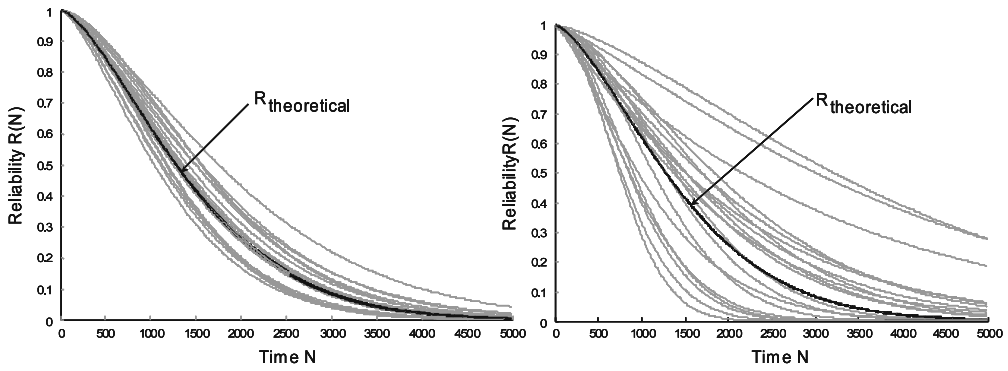


Figure 17.8. Reliability functions by parametric model with censoring at 1,000 hours (a) and 300 hours (b)

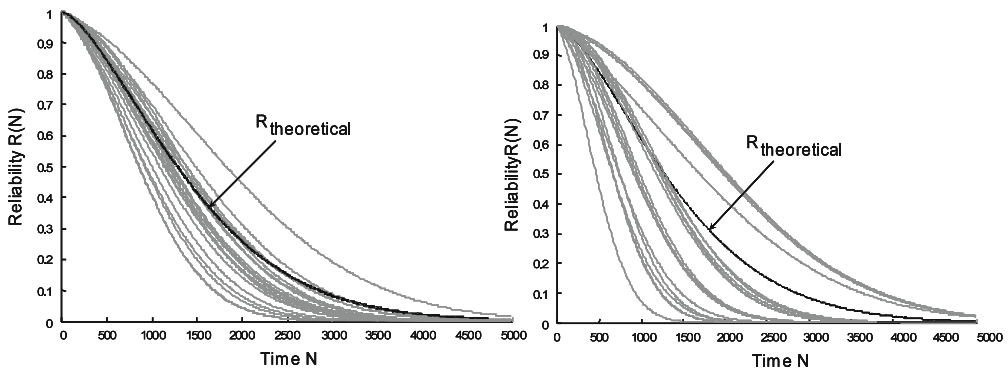


Figure 17.9. Reliability functions by GPH model with censoring at 1,000 hours (a) and 300 hours (b)

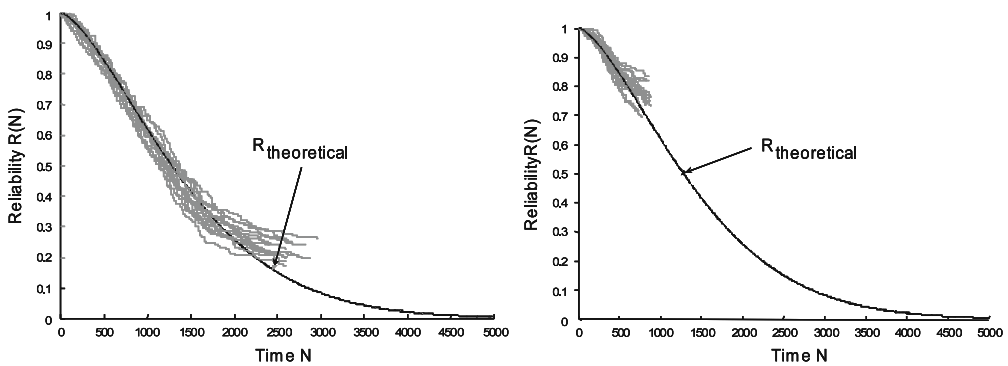


Figure 17.10. Reliability functions by semi-parametric model with censoring at 1,000 hours (a) and 300 hours (b)

levels of stress in order to estimate the *TTF* parameters. Additionally, the levels of stress selected to perform the tests are relatively high when compared to operating conditions. The accuracy of the calculations of the various estimators is largely uncertain, which increases the uncertainty in the estimation of the reliability function for operating conditions (Figure 17.11).

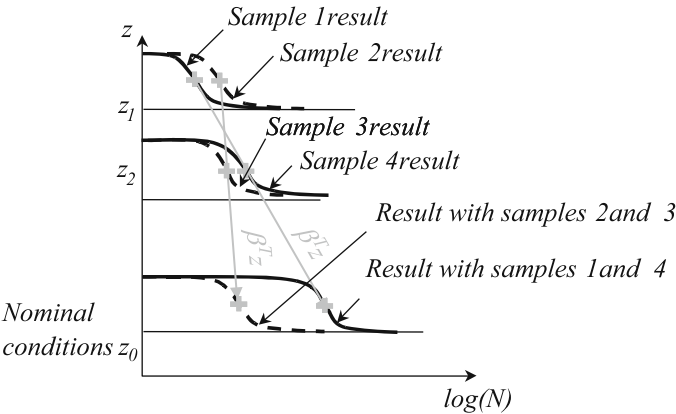


Figure 17.11. Example of error propagation

The parametric models can be used to analyze censored and complete data when the lifetime distribution is assumed to be known. These models allow extrapolating the reliability function in stress and in time (see Figure 17.8a, b with censored data). When the lifetime distribution is unknown, the GPH and semi-parametric models can be used. The GPH model does not consider a form for the distribution, and the reliability function is estimated using parameterized failure rate. This model can be used to extrapolate in stress and in time (see the Figure 17.9a, b with censoring data). However, the number of unknown parameters can be more important than in parametric estimation case. For parametric and GPH models, the reliability curves variability increases with the censoring level (Figures 17.8a, b, 17.9a, b). The semi-parametric model can be used with complete and censored data and the reliability function is assessed directly without any assumption on the lifetime distribution form. It can be used to extrapolate in stress but not in time. In highly censored cases, the semi-parametric model does not allow to completely define the reliability function (see Figures 17.10a, b).

17.4 Reliability Test With Previous Accelerated Damage

17.4.1 Principle

The regression test plan requires a good knowledge of the model, which is rather hard to obtain. Therefore, Bagdonavicius and Nikulin [15, 17] have proposed a new test plan to reduce these difficulties.

Suppose that the lifetime under normal stress s_0 is long and most of the failures appear after the time N_2 . Two samples are then tested:

1. The first sample under accelerated stress s_1 ($s_1 > s_0$),
2. The second one under step-stress $s_2:s_1$ until time N_1 ($N_1 < N_2$) and then nominal stress s_0 until N_2 . Thus, the tested units use many of their “resources” until the time N_1 under accelerated stress. Under nominal stress s_0 the failures are observed in the $[N_1, N_2]$ interval.

Figure 17.12 illustrates this test plan. The stress s_2 corresponds to the step-stress (s_1 until N_1 , s_0 between N_1 and N_2) and $F_{s0}(N)$, $F_{s1}(N)$, and $F_{s2}(N)$ represent the cumulative distribution function under the stresses s_0 , s_1 , and s_2 . For this test plan, the parameterization of acceleration function r is not required.

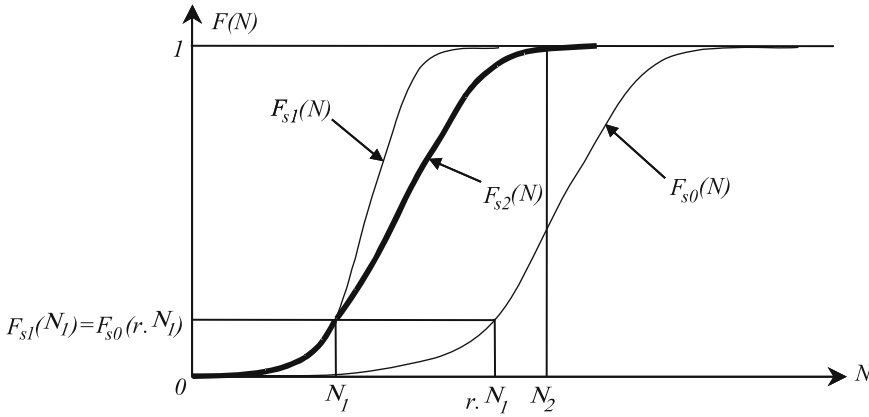


Figure 17.12. Principle of test plan with previous accelerated damage

Bagdonavicius and Nikulin [15, 17] have proposed to estimate the reliability function by two methods: parametric and nonparametric estimations.

17.4.2 Test plan definition

Assume that the function $r(s)$ and $R_{s_0}(n)$ are completely unknown and suppose that the variation coefficient under operating conditions is not large. Most of failures occur in some rather far area and cannot be reached during an experiment under the operating conditions s_0 . To quickly reach this area, an accelerated stress $s_1 > s_0$ can be used at the beginning (until the cycles number $N_1 < N_2$) and when this area is reached, the stress s_1 can be replaced by the normal conditions (until the cycles number N_2). Thus, failures under the normal conditions can be obtained and used to estimate reliability (see Figure 17.13).

In this test plan, the parameterization of function $r(s)$ is not required. Suppose that two groups of items are observed, the first group of items is subjected to accelerated condition s_1 and the second one is subjected to step-stress $s_2(u)$:

$$s_2(u) = \begin{cases} s_1 & \text{if } 0 \leq u \leq N_1 \\ s_0 & \text{if } N_1 \leq u \leq N_2 \end{cases},$$

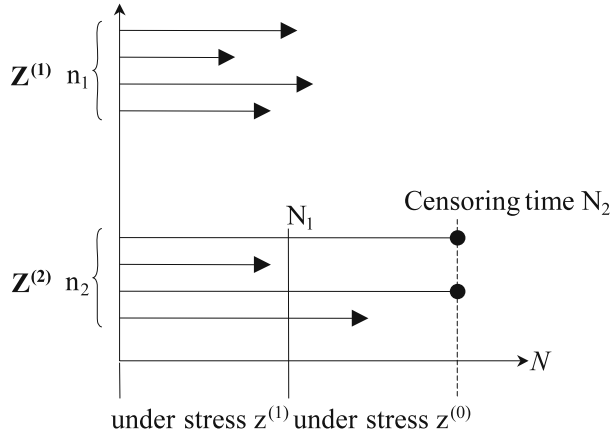


Figure 17.13. Definition of test plan with previous accelerated damage

where u is the number of cycles.

Then the model (17.2) implies

$$R_{s_1}(u) = R_{s_0}(ru) \quad (17.21)$$

and

$$R_{s_2}(u) = \begin{cases} R_{s_0}(ru) & 0 \leq u \leq N_1 \\ R_{s_0}(r.N_1 + u - N_1) & N_1 \leq u \leq N_2. \end{cases} \quad (17.22)$$

The following Figure 17.14 illustrates the method (with $f_{s_1} = -R'_{s_1}$, $f_{s_2} = -R'_{s_2}$ and $f_{s_0} = -R'_{s_0}$).

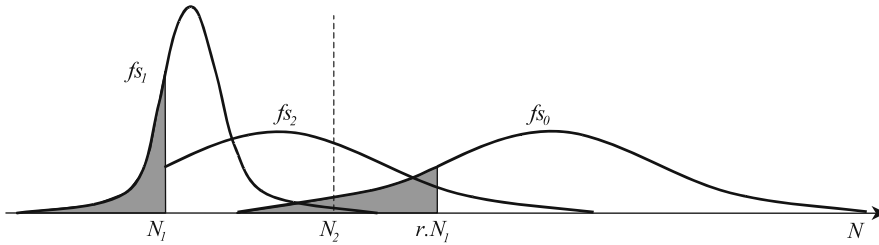


Figure 17.14. Definition of pdf under different stresses

In the following sections the assessment of reliability function $R_{s_0}(u)$ by the parametric and nonparametric estimations are studied.

17.4.3 Parametric model

The parametric estimation consists in choosing an usual probability distribution in order to estimate the reliability function [15, 17]. The Weibull distribution is widely used for mechanical component. Consider two samples of units, the first group of n_1 items is tested under the constant stress s_1 and the complete sample N_{11}, \dots, N_{1n_1} is obtained. The second group of n_2 items is tested under following step-stress s_2 :

$$s_2(u) = \begin{cases} s_1 & \text{if } 0 \leq u \leq N_1 \\ s_0 & \text{if } N_1 < u \leq N_2. \end{cases}$$

We obtain the censored sample of the numbers of cycles N_{21}, \dots, N_{2m_2} , ($m_2 \leq n_2$) with k_2 failures under stress s_1 , $m_2 - k_2$ failures under stress s_0 , and $n_2 - m_2$ surviving systems at the end of N_2 cycles. The reliability function $R_{s_0}(u)$ for the Weibull distribution of lifetime u is defined as

$$R_{s_0}(u) = e^{-\left(\frac{u}{\eta}\right)^v}, \quad (17.23)$$

where η is the scale parameter, v the shape parameter, u the lifetime.

The model (17.21) can be written as

$$R_s(u) = R_{s_0}(r.u) = e^{-\left(\frac{r.u}{\eta}\right)^v}. \quad (17.24)$$

Then

$$R_{s_1}(u) = e^{-\left(\frac{r.u}{\eta}\right)^v} \quad (17.25)$$

and

$$R_{s_2}(u) = \begin{cases} R_{s_0}(ru) = e^{-\left(\frac{r.u}{\eta}\right)^v} & u \leq N_1 \\ R_{s_0}(r.N_1 + u - N_1) = e^{-\left(\frac{r.N_1 + u - N_1}{\eta}\right)^v} & N_1 < u. \end{cases} \quad (17.26)$$

The likelihood function is defined by

$$L(r, \mu, \sigma) = \prod_{i=1}^{n_1} \frac{v}{\eta} \left(\frac{r.N_{1i}}{\eta}\right)^v e^{-\left(\frac{r.N_{1i}}{\eta}\right)^v} \prod_{j=1}^{k_2} \frac{v}{\eta} \left(\frac{r.N_{2j}}{\eta}\right)^v e^{-\left(\frac{r.N_{2j}}{\eta}\right)^v} \prod_{l=k_2+1}^{m_2} \frac{v}{\eta} \left(\frac{r.N_1 + N_{2l} - N_1}{\eta}\right)^v e^{-\left(\frac{r.N_1 + N_{2l} - N_1}{\eta}\right)^v} \times \left(e^{-\left(\frac{r.N_1 + N_{2l} - N_1}{\eta}\right)^v}\right)^{n_2 - m_2}. \quad (17.27)$$

The reliability function under operating conditions is estimated with the maximum likelihood estimators \hat{r} , $\hat{\eta}$, and \hat{v} :

$$\hat{R}_{s_0}(u) = e^{-\left(\frac{u}{\hat{\eta}}\right)^{\hat{v}}}. \quad (17.28)$$

17.4.4 Nonparametric model

The nonparametric ALT model does not need an assumption on probability distribution. The reliability function is estimated using the Kaplan–Meier estimator. The estimation is performed using the same test plan presented in the previous section.

Denote

- $K_1(u)$ the observed failure number from first sample in interval $[0, u]$,
- $K_2(u)$ the observed failure number from second sample in interval $[0, u]$,
- $S_1(u)$ the surviving number of first sample before the time u .
- $S_2(u)$ the surviving number of second sample before the time u (with $n_2 - m_2$ items censored).

Let $K(u)$ be the total number of “observed failures” in interval $[0, u]$; it is given by (u always representing the moment in logarithmic scale)

$$K(u) = \begin{cases} K_1\left(\frac{u}{r}\right) + K_2\left(\frac{u}{r}\right) & u \leq N_1 \\ K_1\left(\frac{u}{r}\right) + K_2(N_1 + u - r.N_1) & u > N_1. \end{cases} \quad (17.29)$$

And $S(u)$ is the total number of surviving components before the time u :

$$S(u) = \begin{cases} S_1\left(\frac{u}{r}\right) + S_2\left(\frac{u}{r}\right) & u \leq N_1 \\ S_1\left(\frac{u}{r}\right) + S_2(N_1 + u - r.N_1) & u > N_1. \end{cases} \quad (17.30)$$

The reliability function is defined using the Kaplan–Meier estimator:

$$\hat{R}_{s_0}(x) = \prod_{u \leq x} \left(1 - \frac{\Delta K(u)}{S(u)}\right), \quad (17.31)$$

where $\Delta K(u) = K(u) - K(u-)$.

The likelihood function is defined by

$$L(r) = \prod_{i=1}^{n_1} \left[\hat{R}_{s_0}(r.u_{1i}-) - \hat{R}_{s_0}(r.u_{1i}) \right] \prod_{i=1}^{k_2} \left[\hat{R}_{s_0}(r.u_{2i}-) - \hat{R}_{s_0}(r.u_{2i}) \right] \prod_{l=k_2+1}^{m_2} \left[\hat{R}_{s_0}([r.N_1 + u_{2i} - N_1]-) - \hat{R}_{s_0}(r.N_1 + u_{2i} - N_1) \right] \left[\hat{R}_{s_0}(r.N_1 + N_2 - N_1) \right]^{n_2 - m_2}. \quad (17.32)$$

The unknown probability density function $f_{s_0} = -R'_{s_0}$ is approximated by $\hat{R}_{s_0}(u-) - \hat{R}_{s_0}(u)$

The reliability function in operating conditions is evaluated with the maximum likelihood estimator \hat{r} .

17.4.5 Simulation example

The example on ball bearing, presented in Section 17.3.5, is used to illustrate the test plan with previous accelerated damage. The simulation parameters are defined in Table 17.7.

Table 17.7. Simulation data

	Load P in daN	Scale parameter η (eq 23.) in hours	Shape parameter ν	Censoring time
Operating conditions	125	1635.86	1.5	No censoring time
Accelerated conditions	175	946.68	1.5	$N_1 = 300$ hours

The studied test plan is simulated in drawing the times to failure:

- Sample 1 : 20 values N_{1i} are drawn from the Weibull distribution $W(946.68 ; 1.5)$,
- Sample 2 : 20 values N_{2i} are drawn from the Weibull distribution $W(946.68 ; 1.5)$ until $N_1 = 300$ hours for the first values and from Weibull distribution $W(1635.86 ; 1.5)$ for the following values.

In considering the parametric and nonparametric estimations, the simulation results are analyzed:

1. Parametric estimation: the reliability function is defined by a normal distribution and the estimators of η , v , and r are computed by the maximum likelihood method (Equation (17.29)).
2. Nonparametric estimation: the reliability function is estimated by Kaplan–Meier estimator and r is estimated by the maximum likelihood method (Equation (17.32)).

The simulation is repeated (30 drawings) to examine the asymptotic behavior of these different estimators. The mean and the standard deviation estimators are presented in Table 17.8 for each estimator.

Table 17.8. Mean and standard deviation of MLE estimations after 20 repetitions

	Parameters	Mean	Standard deviation
Parametric estimation	$\hat{\eta}$	1636.8	1.4273
	\hat{v}	1.4915	0,061
	\hat{r}	1.3931	0.0142
Nonparametric estimation	\hat{r}	1.3840	0.0506

The estimators of acceleration factor \hat{r} is equal to 1.39 for the parametric model (respectively $\hat{r}=1.38$ for the nonparametric model); we can say that, on average, the lifetime of the part under the accelerated conditions accounts for $71.43\% = 100 \cdot (1/1.40)$ ($72.76\% = 100 \cdot (1/1.38)$, respectively) of the lifetime in normal operating conditions. We have a time saving of about 28.57% (27.53%, respectively) in logarithmic scale.

In each simulation, the reliability function $\hat{R}_{s_0}(\cdot)$ under normal conditions is plotted using the parametric and nonparametric estimations (Figure 17.15).

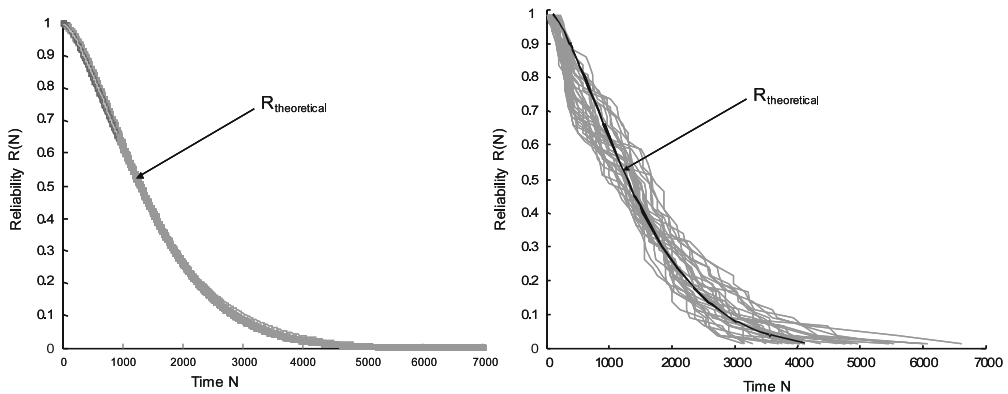


Figure 17.15. Estimation of the reliability function by parametric and nonparametric models

We observe that the estimations by the two models are correct. Indeed, we notice a good adequacy of the parametric and nonparametric models with the theoretical model. This plan is very interesting because it does not require a good knowledge of the

acceleration model, which minimizes the prediction errors. In the case of small samples, it is rather difficult to compare the variances of the estimators from a theoretical point of view. To estimate this variance, we generate 30 something samples of the normal distribution and build the parametric and nonparametric corresponding estimators. The variability seems to be small (Figure 17.15). These examples show the interest of using the plan with a preliminary damage to obtain good results.

17.5 Conclusions

In this chapter, two test plans used to evaluate the lifetime of a mechanical component have been studied. The first test plan by regression consists in testing a sample system under constant acceleration stress. The most frequently used regression models in ALT are log-linear models. The test plan was studied along with three reliability estimates: the parametric model, the GPH model, and the semi-parametric model. Thus, this chapter illustrates a general study on ALT model estimates. These estimates are developed and verified with simulation data. On the basis of the simulation results, the following conclusions can be drawn:

There is good correlation of the parametric, GPH, and semi-parametric models with reference model obtained under normal conditions ($P_0 = 125$ daN). This example shows that these estimates can be used for conducting ALT on mechanical components with various stresses such as loading, strain, or temperature.

These models can be used to analyze both censored and complete data.

The parametric and GPH models allow extrapolation of the reliability function in stress and in time, while the semi-parametric models can be used to extrapolate in stress but not in time.

The difficulty of using the GPH model is to choose the model form and the number of parameters to estimate.

It is difficult to choose a good time transfer function when applying these ALT models on mechanical components. However, the inverse power model (with single stress) is adapted to describe life-stress relationships for mechanical damage such as cracking propagation and fatigue.

A second test plan with previous accelerated damage has been studied. This test plan consists in evaluating the lifetime of a mechanical component under both normal (s_0) and accelerated (s_1) conditions. This test plan consists of test two samples:

1. The first one under accelerated stress s_1 ($s_1 > s_0$),
2. The second one under step-stress s_2 : under s_1 until time T_1 ($T_1 < T_2$) and then under nominal stress s_0 until T_2 . Thus, the tested units use many of their “resources” until the time T_1 under accelerated stress.

This test plan does not require a complex parameterization of the function $r(\cdot)$. Two estimation types have been developed:

- Parametric: this estimation consists in choosing an usual probability distribution in order to estimate the reliability function.

- Nonparametric: this estimation does not make an assumption on the probability distribution. Therefore, the reliability function is estimated using the Kaplan–Meier estimator.

The example shows that the test plan is adequate to assess the reliability function for mechanical components. We note a good correlation of the parametric and nonparametric models with the theoretical model. A Monte Carlo simulation has been performed to examine the asymptotic behavior of the different estimators.

Appendix

Model	Classical definition	Exponential definition
Arrhenius	$\tau = A e^{\frac{B}{T}}$ With A and B : model constants T temperature	$\tau = e^{\beta_0 + \beta_1 z_1}$ With $\beta_0 = \log(A)$ $\beta_1 = B$ $z_1 = \frac{1}{T}$
Inverse power	$\tau = \frac{A}{V^\gamma}$ With A and γ : model constants V: stress	$\tau = e^{\beta_0 + \beta_1 z_1}$ With $\beta_0 = \log(A)$ $\beta_1 = \gamma$ $z_1 = \log(V)$
Aering	$\tau = \frac{A}{T} e^{\frac{B}{T}} e^{(V(C + \frac{D}{T}))}$ With A, B, C and D: model constants T temperature V: stress	$\tau = e^{\beta_0 + z_1 + \beta_1 z_2 + \beta_2 z_3 + \beta_3 z_2 z_3}$ With $\beta_0 = \log(A)$, $\beta_1 = B$, $\beta_2 = C$ and $\beta_3 = D$ $z_1 = -\log(T)$, $z_2 = \frac{1}{T}$ and $z_3 = V$

References

1. Nelson, W. *Accelerated Testing: Statistical models, test plans, and data analysis*. New York: Wiley Series in Probability and Mathematical Statistics 1990.
2. Meeker, W. and Escobar, L. *Statistical Methods for Reliability* New York: Wiley Series in Probability and Mathematical Statistics 1998.
3. Pantelis, V. and Mettas, A. *Understanding accelerated life-testing analysis*. Tutorial Notes: Annual Reliability and Maintainability Symposium 2001, 1–22, ISSN 0897-500.
4. Bagdonavicius, V. and Nikulin, M.S. *Transfer functional and semi-parametric regression*. Biometrika 1997; 84 (2):365–378.
5. Singpurwalla, N.D. *Inference from accelerated life tests when observations are obtained from censored samples*. Technometrics 1971;13: 295–306.
6. Shyur, H-J., Elsayed, E.A., and Luxhoj, J.T. *A general model for accelerated life testing with time-dependent covariates*. Naval Research Logistics 1999; (46):303–321.

7. Ciampi, A. and Etezadi-Amoli, J., *A general model for testing the proportional hazards and the accelerated failure time hypotheses in the analysis of censored survival data with covariates*. Stat Theor Meth 1985; 14: 651–667.
8. Basu, A.P. and Ebrahimi N. *Nonparametric accelerated life testing*. IEEE Trans Reliab 1982; 31(5):432–435.
9. Devarajan, K. and Ebrahimi, N. *Nonparametric approach to accelerated life testing under multiple stress*. Naval Research Logistics 1998; (45):629–644.
10. Bagdonavicius, V. and Nikulin, M.S. *Accelerated life models: Modeling and Statistical Analysis*. London, New York, Washington DC: Chapman & Hall/CRC, 2001.
11. Kececioglu, D., Jiang, M., and Sun, F-B. *A unified approach to random fatigue reliability quantification under random loading*. Proceedings of the Annual Reliability and Maintainability Symposium 1998, 308–313, ISSN 0149-144X.
12. Zhang, C. *Mechanical component lifetime estimation based on accelerated life testing with singularity extrapolation*. Mechanical Systems and Signal Processing 2002; 16:705–718.
13. Owen, W.J and Padgett, W.J. *Accelerated test models for system strength based on Birnbaum-Saunders distributions*. Lifetime Data Anal 1999; 5:133–147.
14. Pham, H. *Handbook of reliability engineering*. Verlag London Limited: Springer, 2003, 149–149.
15. Bagdonavicius, V. and Nikulin, M.S. *On accelerated testing of systems*. European Journal of Diagnosis and Safety in Automation, 1995; 5: 307–316.
16. Velichko, S.A., Dorrough, M.J., Blunn, R.G.: *US20087337088* (2008).
17. Bagdonavicius V, and Nikulin M.S. *Semi-parametrics models in accelerated life testing*. Queen’s papers in pure and applied mathematics. Queen’s University, Kingston, Ontario, Canada, 1995.
18. Geisler, W.L., Bristol, B.J., Grout, J.D., Garrod, T.C., and Tornquist, D.B.: *EP1895177* (2008).
19. Kuwabara, H., Sugimori, S., and Asakura, T.: *EP1889897* (2008).

Reliability Estimation from Failure-Degradation Data with Covariates

V. Bagdonavičius¹, I. Masiulaitytė¹, M.S. Nikulin²

¹ Vilnius University, Vilnius, Lithuania

² IMB, Victor Segalen University, Bordeaux, France

Abstract: Models and estimation methods for failure process parameters and various reliability characteristics using censored multi-mode failure time-degradation data with covariates. Degradation data is supposed to be measured at discrete times. Modified likelihood function which uses predictors of degradation process is defined. Examples of predictors when degradation process is modelled by various stochastic processes are given.

Keywords and phrases: Covariates, degradation, failures, gamma process, path models, shock process, traumatic failure, wear

18.1 Introduction

An important part of modern reliability theory is modelling and statistical analysis of degradation processes of technical units or systems.

The most applied stochastic processes used as degradation models are Meeker's general path models [23–25, 6, 7] and time-scaled stochastic processes with stationary and independent increments such as the gamma process ([2–4, 18, 27, 29]), compound Poisson process ([12, 15, 16, 30, 31, 17]) and Wiener process with drift ([9–11, 32–34, 19, 20, 26, 27]). Harlamov [14] discusses inverse gamma process as a wear model and Zacks [37] discusses general compound renewal damage processes.

If joint failure time and degradation data are available models and estimation methods for analysis of such data are needed. Excellent introductions to failure time-degradation models is given in [28], [36]. More recent developments can be found in [5, 13, 20, 35]. Methods of estimation from failure time-degradation data may be found in [4, 5, 19, 21, 22, 6, 7, 18, 8].

In this chapter we introduce the modified maximum likelihood method for estimation of failure process characteristics using degradation and multi-mode failure time data with covariates.

18.2 Modelling Simultaneous Traumatic Events and Degradation Data Under Covariates

Suppose that the following data are available for reliability characteristics estimation: failure times (possibly censored), explanatory variables (covariates, stresses) and the values of some observable quantity characterizing the degradation of units. The failure rate of units may depend on covariates, degradation level and time. For example, many covariates influence the wear rate of tires: state and type of road covering, weight of the load, weather conditions (temperature, humidity), pressure inside tires, type of a vehicle, steep turns. The covariates may influence the intensity of traumatic failures via the wear and also directly.

We call a failure non-traumatic when the degradation attains a critical level z_0 . Other failures are called traumatic. Traumatic failures may be of different modes, related with production defects, caused by mechanical damages or by fatigue of components.

Suppose that under fixed constant covariate the degradation is stochastic process $Z(t)$, $t \geq 0$.

Suppose that the degradation process $Z(t)$ is non-decreasing with cadlag trajectories.

Denote by $T^{(k)}$ the moment of the traumatic failure of the k th mode, $k = 1, \dots, s$.

We suppose that the random variables $T^{(1)}, \dots, T^{(s)}$ are conditionally independent given the degradation Z .

Denote by $\tilde{\lambda}^{(k)}(t|Z) = \tilde{\lambda}^{(k)}(t|Z(s), 0 \leq s \leq t)$ the conditional failure rate of the traumatic failure of the k th mode given the degradation.

Suppose that this conditional failure rate has two additive components: one related to observed degradation values and the, other to non-observable degradation (aging) and to possible shocks causing sudden traumatic failures. For example, observable degradation of tires is the wear of the protector. The failure rate of tire explosion depends on thickness of the protector, on non-measured degradation level of other tire components and on intensity of possible shocks (hitting a kerb, nail, etc.). So

$$\tilde{\lambda}^{(k)}(t|Z) = \lambda^{(k)}(Z(t)) + \mu^{(k)}(t). \quad (18.1)$$

The function $\lambda^{(k)}(z)$ characterizes the dependence of the rate of traumatic failures of the k th mode on degradation.

Suppose that external covariates [20] influence degradation rate and traumatic event intensity.

Let $x(t) = (x_1(t), \dots, x_s(t))^T$ be a vector of possibly time-dependent covariates. We assume in what follows that x_i are deterministic or realizations of bounded right continuous with finite left-hand limits stochastic processes.

Denote by $Z(t|x)$ the degradation level at the moment t for units functioning under the covariate x .

We suppose that the covariates influence locally the scale of the traumatic failure time distribution component related to aging (non-observable degradation) and to possible shocks, i.e., the accelerated failure time (AFT) [1] model is true for this component. Let us explain it in detail. Denote by

$$S_1^{(k)}(t|Z) = \exp \left\{ - \int_0^t \lambda^{(k)}[Z(u)] du \right\}, \quad S_2^{(k)}(t) = \exp \left\{ - \int_0^t \mu^{(k)}(u) du \right\}$$

the survival functions corresponding to the failure rates $\lambda^{(k)}(Z(u))$ and $\mu^{(k)}(u)$. The first survival function is conditional given the degradation.

The AFT model defines the following relation of the second survival function and the covariates:

$$S_2^{(k)}(t|x) = S_2^{(k)} \left(\int_0^t e^{\beta_k^T x(s)} ds \right),$$

the parameters β_k have the same dimension as x . The covariate x may be replaced by some specified function $\varphi(x)$.

Set

$$f(t, x, \beta) = \int_0^t e^{\beta^T x(u)} du, \quad (18.2)$$

and denote by $g(t, x, \beta)$ the inverse of $f(t, x, \beta)$ with respect to the first argument. If $x = \text{const}$, then

$$f(t, x, \beta) = e^{\beta^T x} t, \quad g(t, x, \beta) = e^{-\beta^T x} t.$$

The function $f(t, x, \beta)$ is time transformation in dependence on x . For units functioning under different covariates $x^{(1)}$ and $x^{(2)}$ two moments t_1 and t_2 , respectively, are equivalent in the sense of degradation if they verify the equality $f(t_1, x^{(1)}, \beta) = f(t_2, x^{(2)}, \beta)$, i.e., we consider the following model for degradation process under covariates:

$$Z(t|x) = Z(f(t, x, \beta)).$$

The covariates have double influence on the distribution of the first traumatic failure component via degradation and directly. So we combine the AFT and the proportional hazards models:

$$S_1^{(k)}(t|x, Z) = \exp \left\{ - \int_0^t e^{\tilde{\beta}_k^T x(u)} \lambda^{(k)}(Z(u|x)) du \right\}.$$

Denote by

$$S^{(k)}(t|x, Z) = \mathbf{P}(T^{(k)} > t|x(u), Z(u|x), 0 \leq u \leq t), \quad \tilde{\lambda}^{(k)}(t|x, Z) = -\frac{d}{dt} \ln S^{(k)}(t|x, Z)$$

the conditional distribution function and the failure rate of the traumatic failure of the k th mode given the covariates and the degradation. So we consider the following model:

$$\mathbf{P}(T^{(1)} > t, \dots, T^{(s)} > t|x(u), Z(u|x), 0 \leq u \leq t) = \prod_{k=1}^s S^{(k)}(t|x, Z), \quad (18.3)$$

$$S^{(k)}(t|x, Z) = \exp \left\{ - \int_0^t \tilde{\lambda}^{(k)}(t|x, Z) du \right\} = \exp \left\{ - \int_0^t e^{\tilde{\beta}_k^T x(u)} \lambda^{(k)}(Z(u|x)) du - H^{(k)}(f(t, x, \beta_k)) \right\}, \quad (18.4)$$

where

$$\tilde{\lambda}^{(k)}(t|x, Z) = e^{\tilde{\beta}_k^T x(t)} \lambda^{(k)}(Z(t|x)) + e^{\beta_k^T x(t)} \mu^{(k)}(f(t, x, \beta_k)), \quad (18.5)$$

$$H^{(k)}(t) = \int_0^t \mu^{(k)}(u) du. \quad (18.6)$$

Denote by

$$T^{(0)} = \inf\{t : Z(t|x) \geq z_0\} \quad (18.7)$$

and

$$S^{(0)}(t|x) = \mathbf{P}\{T^{(0)} > t \mid x(u), 0 \leq u \leq t\} = \mathbf{P}\{Z(t|x) < z_0 \mid x(u), 0 \leq u \leq t\} \quad (18.8)$$

the time to non-traumatic failure and its survival function under the covariate x , respectively.

The time of the unit failure

$$T = \min(T^{(0)}, T^{(1)}, \dots, T^{(s)}) \quad (18.9)$$

may be traumatic or non-traumatic.

Denote by

$$V = k \quad \text{if} \quad T = T^{(k)}, \quad k = 0, \dots, s, \quad (18.10)$$

the indicator of the failure mode. The failure mode 0 is non-traumatic. Others are traumatic.

Let us consider reliability characteristics which are interesting for applications. These are the following:

(1) The survival function of the failure time under the covariate x :

$$S(t|x) = \mathbf{P}(T > t|x) = \mathbf{E}S(t|x, Z), \quad S(t|x, Z) = \mathbf{1}_{\{Z(t|x) < z_0\}} \prod_{k=1}^s S^{(k)}(t \mid x, Z). \quad (18.11)$$

(2) Mean failure time under the covariate x :

$$e(x) = \mathbf{E}(T|x) = \mathbf{E}(\mathbf{E}(T|x, Z)), \quad \mathbf{E}(T|x, Z) = \int_0^{T^{(0)}} \prod_{k=1}^s S^{(k)}(t \mid x, Z) dt. \quad (18.12)$$

(3) The probability that under the covariate x the non-traumatic failure is observed in the interval $[0, t]$:

$$P^{(0)}(t|x) = \mathbf{E}P^{(0)}(t|x, Z), \quad P^{(0)}(t|x, Z) = \mathbf{1}_{\{Z(t|x) \geq z_0\}} \prod_{k=1}^s S^{(k)}(T^{(0)} \mid x, Z). \quad (18.13)$$

In particular, the probability of observed non-traumatic failure under the covariate x in the interval $[0, \infty)$ is obtained.

(4) The probability that under the covariate x a traumatic failure is observed in the interval $[0, t]$:

$$P^{(tr)}(t|x) = \mathbf{E}P^{(tr)}(t|x, Z), \quad P^{(tr)}(t|x, Z) = 1 - \prod_{k=1}^s S^{(k)}(t \wedge T^{(0)} | x, Z). \quad (18.14)$$

(5) The probability that under the covariate x the traumatic failure of the k th mode, $k = 1, \dots, s$, is observed in the interval $[0, t]$:

$$P^{(k)}(t|x) = \mathbf{E}P^{(k)}(t|x, Z), \quad P^{(k)}(t|x, Z) = \int_0^{t \wedge T^{(0)}} \prod_{l=1}^s S^{(l)}(s | x, Z) \lambda^{(k)}(s | x, Z) ds. \quad (18.15)$$

Suppose that the cause of some traumatic failure modes is eliminated. Note that elimination of a failure mode may increase the number of failures of other modes. Indeed, a failure of the l th mode is not observed if it is preceded by a failure of the k th mode but this failure might be observed if the k th failure mode would be eliminated.

If i_1 th, ..., i_q th ($1 \leq i_1 < \dots < i_q \leq s$) traumatic failure modes are eliminated then the survival function $S(t|x)$, the mean $e(x)$ and the probabilities $P^{(0)}(t|x)$, $P^{(tr)}(t|x)$, and $P^{(k)}(t|x)$ ($k = 0, 1, \dots, s$) are modified taking $\prod_{l \neq i_1, \dots, i_q}$ instead of $\prod_{l=1}^s$ in the formulas (11)–(15). So an experiment using units with eliminated failure modes is not needed. The estimators of survival characteristics of units with eliminated failure modes are useful for planning possible ways of reliability improvement.

Suppose that a unit did not fail to the moment τ and we have some information about its covariable and degradation processes $(x(s), Z(s|x) | s \leq \tau)$.

Let \mathcal{G} denote the σ -algebra generated by the possessed information about the degradation process and

$$\bar{\mathcal{G}}_\tau = \sigma(\mathcal{G} \cup \{T > \tau\}).$$

The conditional probabilities of the events considered in the previous section given the σ -algebra $\bar{\mathcal{G}}_\tau$ are for $t > \tau$

$$S(t | x, \tau, \mathcal{G}) = \frac{\mathbf{E}_{\mathcal{G}}\{S(t | x, Z)\}}{\mathbf{E}_{\mathcal{G}}\{S(\tau | x, Z)\}}, \quad (18.16)$$

$$P^{(k)}(t | x, \tau, \mathcal{G}) = \frac{\mathbf{E}_{\mathcal{G}}\{P^{(k)}(t | x, Z)\} - \mathbf{E}_{\mathcal{G}}\{P^{(k)}(\tau | x, Z)\}}{\mathbf{E}_{\mathcal{G}}\{S(\tau | x, Z)\}}, \quad (18.17)$$

$$P^{(tr)}(t | \tau, \mathcal{G}) = \frac{\mathbf{E}_{\mathcal{G}}\{P^{(tr)}(t | x, Z)\} - \mathbf{E}_{\mathcal{G}}\{P^{(tr)}(\tau | x, Z)\}}{\mathbf{E}_{\mathcal{G}}\{S(\tau | x, Z)\}}. \quad (18.18)$$

Moreover, the mean residual life of the unit is

$$e(x, \tau, \mathcal{G}) = \frac{\mathbf{E}_{\mathcal{G}}\{(T - \tau)\mathbf{1}_{\{T > \tau\}} | x\}}{\mathbf{E}_{\mathcal{G}}\{S(\tau | x, Z)\}}. \quad (18.19)$$

If $\mathcal{G} = \sigma(Z)$, then

$$\mathbf{E}_{\mathcal{G}}\{S(t | x, Z)\} = \mathbf{E}\{S(t | x, Z)\}, \quad (18.20)$$

$$\mathbf{E}_{\mathcal{G}}\{(T - \tau)\mathbf{1}_{\{T > \tau\}} | x, Z\} = \int_{\tau}^{T^{(0)}} \prod_{l=1}^s S^{(l)}(t | x, Z) dt - \tau \mathbf{1}_{\{Z(\tau|x) < z_0\}} \prod_{l=1}^s S^{(l)}(\tau | x, Z). \quad (18.21)$$

18.3 Estimation of Model Parameters

18.3.1 The data

Suppose that n units are observed. The i th unit is tested under the vector of explanatory variables $x^{(i)}$, and at the moments

$$0 < t_{i1} < t_{i2} < \dots < t_{im_i},$$

the values $Z_{ij} = Z_i(t_{ij} | x^{(i)})$ of the degradation level are supposed to be measured. The moments t_{ij} correspond to the scale of real functioning. For example, in the case of tire wear, t_{ij} mean kilometers done by the i th tire until the j th measurement. The values of covariates are supposed to be observed during the experiment. Most often they should be constant in time or step functions.

Denote by $T_i = \min(T_i^{(0)}, \dots, T_i^{(s)})$ the failure time and V_i – the failure mode indicator. The data may be right censored. Denote by C_i the censoring time of the i th unit and set

$$\tilde{C}_i = C_i \wedge t_{im_i}, \quad X_i = T_i \wedge \tilde{C}_i, \quad \delta_i = \mathbf{1}_{\{T_i \leq \tilde{C}_i\}}, \quad \tilde{\delta}_i = \mathbf{1}_{\{T_i \leq \tilde{C}_i, V_i \neq 0\}}. \quad (18.22)$$

Denote by

$$\mu_i = \begin{cases} j, & \text{if } X_i \in (t_{ij}, t_{i,j+1}], j = 0, \dots, m_i - 1 \\ m_i, & \text{if } X_i = t_{im_i} \end{cases} \quad (18.23)$$

the observed number of measurements of the i th unit.

The data are the random vectors

$$(X_i, \delta_i, V_i, \mu_i, Z_{i1}, \dots, Z_{i\mu_i}, x^{(i)}), \quad i = 1, \dots, n. \quad (18.24)$$

If $\mu_i = 0$, then the degradation values are not observed.

18.3.2 Likelihood function construction

Suppose that the functions $\lambda^{(k)}(z)$ and $\mu^{(k)}(t)$ are from a class of functions

$$\lambda^{(k)}(z) = \lambda^{(k)}(z, \eta_k), \quad \mu^{(k)}(t) = \mu^{(k)}(t, \gamma_k), \quad (18.25)$$

where η_k, γ_k are possibly multidimensional parameters. For example, analysis of tire failure time and wear data shows that the intensities $\lambda^{(k)}(z)$ and $\mu^{(k)}(t)$ typically have the form $(z/\eta_{1k})^{\eta_{2k}}$ and $(t/\gamma_{1k})^{\gamma_{2k}}$.

Suppose at first that degradation processes $Z_i(t) = Z_i(t|x^{(i)})$ of all units are continuously observable. In this case conditional likelihood L and loglikelihood l functions given degradation for the parameters characterizing traumatic failures can be written as follows:

$$\begin{aligned}
 L &= \prod_{i=1}^n \left\{ \sum_{k=1}^s \mathbf{1}_{\{V_i=k\}} \left[e^{\tilde{\beta}_k^T x^{(i)}(X_i)} \lambda^{(k)}(Z_i(X_i); \eta_k) \right. \right. \\
 &\quad \left. \left. + e^{\beta_k^T x^{(i)}(X_i)} \mu^{(k)}(f(X_i, x^{(i)}, \beta_k); \gamma_k) \right] \right\}^{\tilde{\delta}_i} \\
 &\times \exp \left\{ - \sum_{k=1}^s \left(\int_0^{X_i} e^{\tilde{\beta}_k^T x^{(i)}(u)} \lambda^{(k)}(Z_i(u); \eta_k) du - H^{(k)}(f(X_i, x^{(i)}, \beta_k); \gamma_k) \right) \right\}, \\
 l &= \sum_{i=1}^n \sum_{k=1}^s \mathbf{1}_{\{V_i=k\}} \ln \left[e^{\tilde{\beta}_k^T x^{(i)}(X_i)} \lambda^{(k)}(Z_i(X_i); \eta_k) \right. \\
 &\quad \left. + e^{\beta_k^T x^{(i)}(X_i)} \mu^{(k)}(f(X_i, x^{(i)}, \beta_k); \gamma_k) \right] \\
 &\quad - \sum_{k=1}^s \left(\int_0^{X_i} e^{\tilde{\beta}_k^T x^{(i)}(u)} \lambda^{(k)}(Z_i(u); \eta_k) du - H^{(k)}(f(X_i, x^{(i)}, \beta_k); \gamma_k) \right). \quad (18.26)
 \end{aligned}$$

If covariates are absent then

$$\begin{aligned}
 l &= \sum_{i=1}^n \sum_{k=1}^s \mathbf{1}_{\{V_i=k\}} \ln \left[\lambda^{(k)}(Z_i(X_i); \eta_k) + \mu^{(k)}(X_i; \gamma_k) \right] \\
 &\quad - \sum_{k=1}^s \left(\int_0^{X_i} \lambda^{(k)}(Z_i(u); \eta_k) du - H^{(k)}(X_i; \gamma_k) \right). \quad (18.27)
 \end{aligned}$$

If the values of degradation processes are measured only at discrete times t_{ij} then the conditional likelihood function is modified replacing $Z_i(u)$ by their predictors $\hat{Z}_i(u)$ obtained from degradation data. The form of these predictors depends on the form of the degradation processes.

It was mentioned in the introduction that the most applied stochastic processes describing degradation are general path models and time-scaled stochastic processes with stationary and independent increments such as the gamma process, compound Poisson process and Wiener process with drift, the last not monotone.

Let us find the predictors for some specified degradation processes.

18.3.3 Example 1: Time-scaled gamma process

Let m be a real time function. The degradation process Z is *time-scaled* (by the scale function $m(t)$) *gamma process* if

- (1) it has independent increments, i.e., for any $0 < t_1 < \dots < t_k$ the random variables $Z(t_1)$, $Z(t_2) - Z(t_1)$, \dots , $Z(t_k) - Z(t_{k-1})$ are independent;
- (2) for any $t > 0$ the random variable $Z(t)$ has the gamma distribution;
- (3) for any $t \geq 0$

$$\mathbf{E}(Z(t)) = m(t), \quad \mathbf{Var}(Z(t)) = \sigma^2 m(t).$$

The definition implies that for any $x > 0$ the density of the r.v. $Z(t_j) - Z(t_{j-1})$, $j = 1, \dots, k$, $t_0 = 0$, is

$$p_{Z(t_j)-Z(t_{j-1})}(x) = \frac{1}{\sigma^2 \Gamma\left(\frac{\Delta m_j}{\sigma^2}\right)} \left(\frac{x}{\sigma^2}\right)^{\frac{\Delta m_j}{\sigma^2}-1} e^{-\frac{x}{\sigma^2}}, \quad (18.28)$$

$$\Delta m_j = m(t_j) - m(t_{j-1}), \quad \Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx.$$

The density of $Z(t_j)$ is of the same form: Δm_j must be replaced by $m(t_j)$ in (28).

The degradation and its characteristics under covariate x are

$$Z(t|x) = Z(f(t, \beta, x)), \quad m(t|x) = \mathbf{E}(Z(t|x)) = m(f(t, \beta, x)), \quad (18.29)$$

$$\sigma^2(t|x) = \mathbf{Var}(Z(t|x)) = \sigma^2 m(t|x).$$

$m(t|x)$ is the mean degradation under the covariate x .

(a) Parametric form of the mean degradation

The form of the mean degradation $m(t)$ may be suggested by the form of observed degradation curves. In such a case $m(t)$ is chosen from some parametric class of functions (power or other time function depending on a finite-dimensional unknown parameter): $m(t) = m(t; \nu)$, $\nu = (\nu_1, \dots, \nu_q)^T$. The data

$$(Z_{ij}, \mu_i), \quad i = \overline{1, n}, \quad j = \overline{1, \mu_i}, \quad (18.30)$$

are used for estimation of the parameters $\theta = (\beta, \nu, \sigma^2)^T$.

For any $z > 0$ the density of the increment $\Delta Z_{ij} = Z_{ij} - Z_{i,j-1}$ has the form

$$p_{\Delta Z_{ij}}(z; \theta) = \frac{1}{\sigma^2 \Gamma\left(\frac{\Delta \mu_{ij}(\beta, \nu)}{\sigma^2}\right)} \left(\frac{z}{\sigma^2}\right)^{\frac{\Delta \mu_{ij}(\beta, \nu)}{\sigma^2}-1} e^{-\frac{z}{\sigma^2}},$$

where

$$\Delta \mu_{ij}(\beta, \nu) = m\left(f(t_{i,j}, \beta, x^{(i)}); \nu\right) - m\left(f(t_{i,j-1}, \beta, x^{(i)}); \nu\right). \quad (18.31)$$

The likelihood function of the degradation data (18.30) is

$$L_d(\theta) = \prod_{i=1}^n \prod_{j=1}^{\mu_i} p_{\Delta Z_{ij}}(\Delta Z_{ij}; \theta), \quad (18.32)$$

where we set $p_{\Delta Z_{ij}}(\Delta Z_{ij}; \theta) = 1$, if $X_i < t_{i1}$, i.e., when a traumatic event occurs earlier than the first measurement of degradation.

Denote by $\hat{\theta}$ the maximum likelihood estimator. Then for any x the estimator of the mean degradation $m(t|x)$ under the covariate x is

$$\hat{m}(t|x) = m(f(t, \hat{\beta}, x); \hat{\nu}). \quad (18.33)$$

In the case of the data (18.24) degradation values are not measured continuously and the loglikelihood function (18.26) cannot be used for estimation of the parameters

$\beta_k, \tilde{\beta}_k, \eta_k$ and γ_k . For modification of the loglikelihood (18.26) we need predictors of $Z_i(t)$.

Set

$$\tilde{Z}_i(t) = \mathbf{E}(Z_i(t) | Z_i(t_{i1}), \dots, Z_i(t_{i\mu_i})). \quad (18.34)$$

For $j = 1, \dots, \mu_i$ we have $\tilde{Z}_i(t_{ij}) = Z_{ij}$. For $t \in (t_{i,j-1}, t_{ij})$, $j = 1, \dots, \mu_i$, the conditional means (18.34) are

$$\tilde{Z}_i(t; \theta) = Z_{i,j-1} + \frac{\Delta m_{ij}(t; \beta, \nu)}{\Delta m_{ij}(\beta, \nu)} \Delta Z_{ij},$$

$$\Delta m_{ij}(t; \beta, \nu) = m(f(t, \beta, x^{(i)}); \nu) - m(f(t_{i,j-1}, \beta, x^{(i)}); \nu). \quad (18.35)$$

For any $t > t_{i,\mu_i}$

$$\tilde{Z}_i(t; \theta) = Z_{i\mu_i} F_{\chi_{\kappa_i}^2}(2(z_0 - Z_{i\mu_i})) + \Delta m_{i,\mu_i+1}(t; \beta, \nu) F_{\chi_{\kappa_i+2}^2}(2(z_0 - Z_{i\mu_i})), \quad (18.36)$$

where $F_{\chi_n^2}(x)$ is the c.d.f. of the chi square distribution, $\kappa_i = 2\Delta m_{i,\mu_i+1}(t; \beta, \nu)$.

Note that for any $t > t_{i,\mu_i}$

$$\tilde{Z}_i(t; \theta) \rightarrow Z_{i\mu_i} + \Delta m_{i,\mu_i+1}(t; \beta, \nu), \quad \text{as } z_0 \rightarrow \infty. \quad (18.37)$$

The predictors \hat{Z}_i of Z_i are defined as

$$\hat{Z}_i(t) = \tilde{Z}_i(t; \hat{\theta}). \quad (18.38)$$

(b) Unknown form of mean degradation

If the function m is completely unknown then non-parametric estimator of this function is used seeking predictors of the stochastic processes Z_i .

A piecewise linear approximation of the process $Z_i(t) = Z(t|x^{(i)})$ on $[0, X_i]$ is

$$Z_i^*(t) = \sum_{j=1}^{\mu_i+1} \left[Z_i(t_{i,j-1}) + \frac{t - t_{i,j-1}}{t_{ij} - t_{i,j-1}} (Z_i(t_{ij}) - Z_i(t_{i,j-1})) \right] \mathbf{1}_{[t_{i,j-1}, t_{ij}]}(t), \quad (18.39)$$

$t_{i0} = 0$, $t_{i,\mu_i+1} = X_i$. We denoted by $g(t, x, \beta)$ the inverse of $f(t, x, \beta)$ with respect to the first argument. The distribution of the stochastic process

$$Z(t) = Z(g(t, x, \beta)|x) \quad (18.40)$$

does not depend on x , so the processes $Z_i^*(g(t, x^{(i)}, \beta))$ are approximations of the process Z and can be used constructing an estimator of the mean

$$m(t) = \mathbf{E}Z(t) = \mathbf{E}Z(g(t, x, \beta)|x). \quad (18.41)$$

These approximating processes are censored at the points $t_i^*(\beta) = g(X_i, x^{(i)}, \beta)$. Consider the ordered sequence of distinct moments

$$t_{(1)}^*(\beta) < \dots < t_{(d)}^*(\beta), \quad d \leq n.$$

Take the following pseudo-estimator (depending on β) of $m(t)$:

$$\begin{aligned}\tilde{m}(t, \beta) &= \frac{1}{n} \sum_{i=1}^n Z_k^*(g(t, \beta, x^{(i)})), \quad t \in [0, t_{(1)}^*(\beta)], \\ \tilde{m}(t, \beta) &= \tilde{m}(t_{(j-1)}^*(\beta), \beta) + \frac{\sum_{i: t_i^*(\beta) > t_{(j-1)}^*(\beta)} \left(Z_i^*(g(t, \beta, x^{(i)})) - Z_i^*(g(t_{(j-1)}^*(\beta), \beta, x^{(i)})) \right)}{\sum_{i: t_i^*(\beta) > t_{(j-1)}^*(\beta)} 1},\end{aligned}\tag{18.42}$$

$t \in (t_{(j-1)}^*(\beta), t_{(j)}^*(\beta)]$. The likelihood function from degradation data (18.30) is written in the form (18.32) putting

$$\theta = (\beta^T, \sigma^2)^T, \quad \Delta m_{ij}(\beta) = \tilde{m}\left(f(t_{i,j}, \beta, x^{(i)}), \beta\right) - \tilde{m}\left(f(t_{i,j-1}, \beta, x^{(i)}), \beta\right). \tag{18.43}$$

Denote by $\hat{\beta}$, $\hat{\sigma}^2$ the maximum likelihood estimators. The function $m(t)$ is estimated by the statistic $\hat{m}(t) = \tilde{m}(t, \hat{\beta})$.

Define $\tilde{Z}_i(t; \theta)$ by (18.35) and (18.36) replacing $\Delta m_{ij}(\beta, \nu)$ by $\Delta m_{ij}(\beta)$ given in (18.43).

The predictors of $Z_i(t)$ are $\hat{Z}_i(t) = \tilde{Z}_i(t; \hat{\theta})$.

18.3.4 Example 2: Shock processes

Assume that degradation results from shocks, each of them leading to an increment of degradation. Let T_n , ($n \geq 1$) be the time of the n th shock and X_n the n th increment of the degradation level. Denote by $N(t)$ the number of shocks in the interval $[0, t]$. Set $X_0 = 0$. The degradation process is given by

$$Z(t) = \sum_{n=1}^{\infty} \mathbf{1}\{T_n \leq t\} X_n = \sum_{n=0}^{N(t)} X_n.$$

Kahle and Wendt (2006) model T_n as the moments of transition of the doubly stochastic Poisson process, i.e., they suppose that the distribution of the number of shocks up to time t is given by

$$\mathbf{P}\{N(t) = k\} = \mathbf{E} \left\{ \frac{(Y\eta(t))^k}{k!} \exp\{-Y\eta(t)\} \right\},$$

where $\eta(t)$ is a deterministic function and Y is a non-negative random variable with finite expectation. If Y is non-random, N is a non-homogenous Poisson process, in particular, when $\eta(t) = \lambda t$, N is a homogenous Poisson process. Other models for η may be used, for example, $\eta(t) = t^\alpha$, $\alpha > 0$.

Assume that X_1, X_2, \dots are conditionally independent given $\{T_n\}$ and assume that the probability density functions of X_n given $\{T_n\}$ is g .

Let us consider the case when the number of shocks is modelled by non-homogenous Poisson process:

$$\mathbf{P}\{N(t_1) = i_1, N(t_2) - N(t_1) = i_2, \dots, N(t_m) - N(t_{m-1}) = i_m\} =$$

$$\frac{\eta(t_1)^{i_1}}{i_1!} e^{-\eta(t_1)} \frac{[\eta(t_2) - \eta(t_1)]^{i_2}}{i_2!} e^{-[\eta(t_2) - \eta(t_1)]} \dots \frac{[\eta(t_m) - \eta(t_{m-1})]^{i_m}}{i_m!} e^{-[\eta(t_m) - \eta(t_{m-1})]}. \quad (18.44)$$

The degradation process $Z(t)$ is a stochastic process with independent increments and for any $z \geq 0$ the density of the r.v. $Z(t) - Z(s)$, $0 \leq s, t$, is

$$p_{Z(t)-Z(s)}(z) = \sum_{k=1}^{\infty} g_k(z) \frac{[\eta(t) - \eta(s)]^k}{k!} e^{-[\eta(t) - \eta(s)]}, \quad (18.45)$$

where g_k is the convolution of k densities g . For example, if the sizes of the shocks X_i have exponential distribution $\mathcal{E}(\xi)$: $g(u) = \xi e^{-\xi u}$, $u \geq 0$, then

$$g_k(u; \xi) = \frac{\xi^k u^{k-1}}{(k-1)!} e^{-\xi u}, \quad u \geq 0, \quad p_{Z(t)-Z(s)}(z) = \xi b e^{-\xi z - b} \sum_{k=0}^{\infty} \frac{(\xi b z)^k}{k!(k+1)!},$$

where $b = \eta(t) - \eta(s)$.

Denote by $a_1 = \mathbf{E}X_1$ and $a_2 = \mathbf{E}X_1^2$ the first two moments of the random variable X_1 . The moments of $Z(t)$ are

$$\mathbf{E}(Z(t)) = a_1 \eta(t), \quad \mathbf{Var}(Z(t)) = a_2 \eta(t).$$

The degradation and its characteristics under covariate x are

$$Z(t|x) = Z(f(t, \beta, x)), \quad m(t|x) = \mathbf{E}(Z(t|x)) = \mu_1 \eta(f(t, \beta, x)), \quad (18.46)$$

$$\sigma^2(t|x) = \mathbf{Var}(Z(t|x)) = a_2 \eta(f(t, \beta, x)).$$

(a) Parametric form of the mean degradation

Suppose that g and η belong to some parametric classes $g(t) = g(t, \xi)$, $\xi = (\xi_1, \dots, \xi_p)^T$, and $\eta(t) = \eta(t; \nu)$, $\nu = (\nu_1, \dots, \nu_q)^T$. Set $\theta = (\beta^T, \nu^T, \xi^T)^T$. The likelihood function of the degradation data (30) is of the form (32), where for any $z > 0$, $0 \leq s < t$, the density of the increment $Z(t) - Z(s)$ is

$$p_{Z(t)-Z(s)}(z; \theta) = \sum_{k=1}^{\infty} g_k(z; \xi) \frac{[\Delta \eta(s, t; \beta, \nu)]^k}{k!} e^{-\Delta \eta(s, t; \beta, \nu)}, \quad (18.47)$$

here

$$\Delta \eta(s, t; \beta, \nu) = \eta(f(t, \beta, x^{(i)}; \nu)) - \eta(f(s, \beta, x^{(i)}; \nu)). \quad (18.48)$$

Denote by $\hat{\theta}$ the maximum likelihood estimator. Then for any x the estimator of the mean degradation $m(t|x)$ under the covariate x has the form (18.33).

For $t \in (t_{i,j-1}, t_{i,j})$, $j = 1, \dots, \mu_i$, the conditional means (18.34) are

$$\tilde{Z}_i(t; \theta) = \frac{1}{p_{\Delta Z_{i,j}}(\Delta Z_{i,j}; \theta)} \int_{Z_{i,j-1}}^{Z_{i,j}} z p_{Z(t)-Z_{i,j-1}}(z - Z_{i,j-1}; \theta) p_{Z_{i,j}-Z(t)}(Z_{i,j} - z; \theta) dz. \quad (18.49)$$

For any $t > t_{i,\mu_i}$

$$\tilde{Z}_i(t; \theta) = \int_{Z_{i, \mu_i}}^{z_0} z p_{Z(t)-Z_{i, \mu_i}}(z - Z_{i, \mu_i}; \theta) dz. \quad (18.50)$$

The predictors \hat{Z}_i of Z_i are defined by the formula (18.38).

Note that as in the case of the gamma process (we set $\sigma^2 = a_2/a_1$)

$$\frac{\mathbf{Var}(Z(t))}{\mathbf{E}(Z(t))} = \sigma^2, \quad \text{Cov}(Z(s), Z(t)) = \mathbf{Var}(Z(s \wedge t)),$$

so in terms of the first two moments the considered shock process and the gamma process are of identical structure.

(b) Unknown form of mean degradation

The predictors of $Z_i(t)$ are defined as $\hat{Z}_i(t) = \tilde{Z}_i(t; \hat{\theta})$ and $\tilde{Z}_i(t; \theta)$ are defined by (18.49) and (18.50) replacing $\Delta\eta(s, t; \beta, \nu)$ (given in (18.48)) by

$$\Delta\eta(s, t; \beta, \xi) = \left[\tilde{m} \left(f(t_{i,j}, \beta, x^{(i)}), \beta \right) - \tilde{m} \left(f(t_{i,j-1}, \beta, x^{(i)}), \beta \right) \right] / a_1(\xi),$$

where $\tilde{m}(t; \beta)$ is the pseudoestimator of the mean $m(t) = \mathbf{E}Z(t)$ given by (18.42).

18.3.5 Example 3: Path models

Suppose that the degradation process $Z(t)$ is of the following form:

$$Z(t) = \varphi(t, A, \nu), \quad (18.51)$$

where φ is a deterministic function and $A = (A_1, \dots, A_p)$ is a finite-dimensional random vector and ν is a finite-dimensional non-random parameter.

The form of the function φ may be suggested by the form of individual degradation curves. The degradation under the covariate x is modelled by

$$Z(t|x) = \varphi(f(t, x, \beta), A), \quad m(t|x) = E\varphi(f(t, x, \beta), A).$$

Let us consider the following typical example:

$$Z(t) = (t/A)^\nu; \quad (18.52)$$

here A is a positive random variable with unknown cumulative distribution function F and ν is a positive parameter. In particular case $\nu = 1$ this model fits well as the tire wear model (see [24]).

The degradation process under covariate x is

$$Z(t|x) = Z(f(t, x, \beta)) = (f(t, x, \beta)/A)^\nu. \quad (18.53)$$

Even in the case $\nu = 1$ it is not necessarily linear.

Suppose that n units are on test. The i th unit is tested under explanatory variable $x^{(i)}$. Denote by T_i , V_i the failure times and the failure modes, respectively. Suppose that the degradation values $Z^{(i)}$ at the moments T_i are observed. So the data have the form

$$(T_i, V_i, Z^{(i)}, x^{(i)}), \quad i = 1, \dots, n. \quad (18.54)$$

The covariates $x^{(i)}$ are observed until the moments X_i .

Taking into account that the random variables

$$\ln A_i = \nu f(T_i, x^{(i)}, \beta) - \ln Z^{(i)}$$

are independent identically distributed with the mean, say m , which does not depend on β and ν , these parameters are estimated by the method of least squares, minimizing the sum

$$\sum_{i=1}^n (\nu \ln f(T_i, x^{(i)}, \beta) - \ln Z^{(i)} - m)^2,$$

which gives the system of equations

$$\begin{aligned} n \sum_{i=1}^n \frac{\int_0^{T_i} x^{(i)} e^{\beta^T x^{(i)}(u)} du [\nu \ln f(T_i, x^{(i)}, \beta) - \ln Z^{(i)}]}{f(T_i, x^{(i)}, \beta)} \\ - \sum_{i=1}^n \frac{\int_0^{T_i} x^{(i)} e^{\beta^T x^{(i)}(u)} du}{f(T_i, x^{(i)}, \beta)} \sum_{j=1}^n [\nu \ln f(T_j, x^{(j)}, \beta) - \ln Z^{(j)}] = 0, \\ n \sum_{i=1}^n \ln f(T_i, x^{(i)}, \beta) [\nu \ln f(T_i, x^{(i)}, \beta) - \ln Z^{(i)}] \\ - \sum_{i=1}^n \ln f(T_i, x^{(i)}, \beta) \sum_{j=1}^n [\nu \ln f(T_j, x^{(j)}, \beta) - \ln Z^{(j)}] = 0. \end{aligned}$$

If $x^{(i)}$ are constant then this system is

$$\begin{aligned} n \sum_{i=1}^n \beta^T x^{(i)} [\nu \beta^T x^{(i)} + \nu \ln T_i - \ln Z^{(i)}] - \sum_{i=1}^n \beta^T x^{(i)} \sum_{j=1}^n [\nu \beta^T x^{(j)} + \nu \ln T_j - \ln Z^{(j)}] = 0, \\ n \sum_{i=1}^n x^{(i)} [\nu \beta^T x^{(i)} + \nu \ln T_i - \ln Z^{(i)}] - \sum_{i=1}^n x^{(i)} \sum_{j=1}^n [\nu \beta^T x^{(j)} + \nu \ln T_j - \ln Z^{(j)}] = 0. \end{aligned}$$

Denote by $\hat{\beta}$ and $\hat{\nu}$ the obtained estimator.

The predictors of $Z_i(t)$ are defined as

$$\hat{Z}_i(t) = \left(\frac{f(t, x^{(i)}, \hat{\beta})}{f(T_i, x^{(i)}, \hat{\beta})} \right)^{\hat{\nu}} Z^{(i)}. \quad (18.55)$$

Note that $\hat{Z}_i(T_i) = Z^{(i)}$. If $x^{(i)}$ are constant over time then

$$\hat{Z}_i(t) = \left(\frac{t}{T_i} \right)^{\hat{\nu}} Z^{(i)}. \quad (18.56)$$

18.3.6 Modified loglikelihood

The modified loglikelihood function for the parameters $\beta_k, \tilde{\beta}_k, \eta_k$ and γ_k from the data (18.24) and (18.54) is obtained modifying the loglikelihood function (18.26): the stochastic processes Z_i are replaced by their predictors \hat{Z}_i in (18.26) (in the case of the data (18.54) take $X_i = T_i$):

$$\begin{aligned} \tilde{l} = & \sum_{i=1}^n \sum_{k=1}^s \mathbf{1}_{\{V_i=k\}} \ln \left[e^{\tilde{\beta}_k^T x^{(i)}(X_i)} \lambda^{(k)}(\hat{Z}_i(X_i); \eta_k) + e^{\beta_k^T x^{(i)}(X_i)} \mu^{(k)}(f(X_i, x^{(i)}, \beta_k); \gamma_k) \right] \\ & - \sum_{k=1}^s \left(\int_0^{X_i} e^{\tilde{\beta}_k^T x^{(i)}(u)} \lambda^{(k)}(\hat{Z}_i(u); \eta_k) du - H^{(k)}(f(X_i, x^{(i)}, \beta_k); \gamma_k) \right). \end{aligned} \quad (18.57)$$

If covariants are absent then

$$\begin{aligned} \tilde{l} = & \sum_{i=1}^n \sum_{k=1}^s \mathbf{1}_{\{V_i=k\}} \ln \left[\lambda^{(k)}(\hat{Z}_i(X_i); \eta_k) + \mu^{(k)}(X_i; \gamma_k) \right] \\ & - \sum_{k=1}^s \left(\int_0^{X_i} \lambda^{(k)}(\hat{Z}_i(u); \eta_k) du - H^{(k)}(X_i; \gamma_k) \right), \end{aligned} \quad (18.58)$$

where the predictors \hat{Z}_i are defined replacing $f(u, \beta, x^{(i)})$ by u in all formulas.

The loglikelihood (18.24) function can be modified to the case when the two functions $\lambda^{(k)}$ (or the functions $H^{(k)}$, but not both) are completely unknown. In the case of linear path models such modifications and properties of estimators are given in [7].

Investigating the case of other degradation models is a subject for separate work.

18.4 Estimation of Reliability Characteristics

Let us consider estimation of reliability characteristics (18.11)–(18.20) when the mean degradation $m(t)$ is of parametric form. Set

$$\hat{Z}_i(t|x) = \hat{Z}_i(g(f(t, \hat{\beta}, x), x^{(i)}, \hat{\beta})), \quad (18.59)$$

where $\hat{Z}_i(t)$ are the predictors of the discretely observed processes $Z_i(t|x^{(i)})$, $i = 1, \dots, n$. We considered construction of the predictors $\hat{Z}_i(t)$ in the previous section.

In the particular case when $x, x^{(i)}$ are constant over time

$$\hat{Z}_i(t|x) = \hat{Z}_i(e^{\hat{\beta}(x-x^{(i)})}t). \quad (18.60)$$

The predictor of the non-traumatic failure of the i th unit under the covariate x is

$$\hat{T}_i^{(0)}(x) = \inf\{t : \hat{Z}_i(t|x) \geq z_0\}. \quad (18.61)$$

The formulas (18.11)–(18.15) imply the following estimators:

(1) The estimator of the survival function of the failure time under the covariate x :

$$\hat{S}(t|x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{\hat{Z}_i(t|x) < z_0\}} \prod_{k=1}^s \hat{S}^{(k)}(t | x, \hat{Z}_i), \quad (18.62)$$

where

$$\hat{S}^{(k)}(t | x, \hat{Z}_i) = \exp \left\{ - \int_0^t e^{\hat{\beta}_k^T x(u)} \lambda^{(k)}(\hat{Z}_i(u|x), \hat{\eta}_k) du - H^{(k)}(f(t, x, \hat{\beta}_k), \hat{\gamma}_k) \right\}. \quad (18.63)$$

(2) Mean failure time under the covariate x :

$$\hat{e}(x) = \frac{1}{n} \sum_{i=1}^n \int_0^{\hat{T}_i^{(0)}(x)} \prod_{k=1}^s \hat{S}^{(k)}(t | x, \hat{Z}_i) dt. \quad (18.64)$$

(3) The estimator of the probability that under the covariate x the non-traumatic failure is observed in the interval $[0, t]$:

$$\hat{P}^{(0)}(t|x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{\hat{Z}_i(t|x) \geq z_0\}} \prod_{k=1}^s \hat{S}^{(k)}(\hat{T}_i^{(0)}(x) | x, \hat{Z}_i). \quad (18.65)$$

(4) The estimator of the probability that under the covariate x a traumatic failure is observed in the interval $[0, t]$:

$$\hat{P}^{(tr)}(t|x) = 1 - \frac{1}{n} \sum_{i=1}^n \prod_{k=1}^s \hat{S}^{(k)}(t \wedge \hat{T}_i^{(0)}(x) | x, \hat{Z}_i). \quad (18.66)$$

(5) The estimator of the probability that under the covariate x the traumatic failure of the k th mode, $k = 1, \dots, s$, is observed in the interval $[0, t]$:

$$\hat{P}^{(k)}(t|x) = \frac{1}{n} \sum_{i=1}^n \int_0^{t \wedge \hat{T}_i^{(0)}(x)} \prod_{l=1}^s \hat{S}^{(l)}(s | x, \hat{Z}_i) \lambda^{(k)}(s | x, \hat{Z}_i) ds. \quad (18.67)$$

The estimators of survival characteristics of units with eliminated failure modes are obtained taking $\prod_{l \neq i_1, \dots, i_q}$ instead of $\prod_{l=1}^s$ in the formulas (62)–(67).

Acknowledgements. The authors are grateful to the University Victor Segalen and to “StatXpert” for the financial support.

References

1. Bagdonavičius, V. (1978). Testing the hypothesis of the additive accumulation of damages. *Probability Theory and Its Applications*, **23**, 403–408.

2. Bagdonavičius, V. and Nikulin, M.S. (1999). Modelling dependence of tire wear on explanatory variables. Technical Report 9902, Université de Bretagne-Sud, Vannes.
3. Bagdonavičius, V. and Nikulin, M.S. (2000). Modèle statistique de dégradation models avec des covariables dépendant du temps, *Comptes Rendus de l'Académie des Sciences de Paris*, Série I, **330**, 131–134.
4. Bagdonavičius, V. and Nikulin, M.S. (2001). Estimation in degradation models with explanatory variables, *Lifetime Data Analysis*, **7**, 85–103.
5. Bagdonavičius, V. and Nikulin, M.S. (2002). *Accelerated life models: Modeling and Statistical Analysis*. Chapman & Hall/CRC: Boca Raton.
6. Bagdonavičius, V., Bikelis, A., and Kazakevičius, V. (2004). Statistical analysis of linear degradation and failure time data with multiple failure modes, *Lifetime Data Analysis*, **10**, 65–81.
7. Bagdonavičius V., Bikelis, A., Kazakevičius, V., and Nikulin, M.S. (2007). Non-parametric estimation from simultaneous renewal-failure-degradation data with competing risks. *Journal of Statistical Planning and Inference*, **137**, 2191–2207.
8. Couallier V. (2006). Some recent results on joint degradation and failure time modelling. In *Probability, Statistics and Modelling in Public Health* M.S. Nikulin, D. Commenges, and C. Huber (eds.), pp. 73–89. Springer Science and Business Media.
9. Doksum K.A. and Hoyland A. (1992). Models for variable-stress accelerated life testing experiment based on a Wiener process and the inverse Gaussian distribution, *Technometrics*, **34**, 74–82.
10. Doksum, K.A. and Normand, S.L.T. (1995). Gaussian models for degradation processes – part I: Methods for the analysis of biomarker data, *Lifetime Data Analysis*, **1**, 131–144.
11. Doksum, K.A. and Normand, S.L.T. (1996). Models for degradation processes and event times based on Gaussian processes. In: *Lifetime Data: Models in Reliability and Survival Analysis*, N.P. Jewell et al. (eds), pp. 85–91, Kluwer Academic Publishers.
12. Esary, J.D., Marshall, A.W. and Proshan, F. (1973). Shock models and wear processes, *The Annals of Probability*, **1**, 627–649.
13. Finkelstein, M.S. (2004). On the exponential formula of reliability, *IEEE Transactions on Reliability*, **53**, 265–269.
14. Harlamov, B. (2004). Inverse Gamma-Process as a Model of Wear. In: *Longevity, Aging and Degradation Models in Reliability, Health, Medicine and Biology*, v.2, V. Antonov, C. Huber, M.S. Nikulin, and V. Politschok (eds.), pp. 180–190. St. Petersburg State Polytechnical University, Saint Petersburg.
15. Kahle, W. (1994). Simultaneous confidence regions for the parameters of damage processes. *Statistical Papers*, **35**, 27–41.
16. Kahle, W. and Lehmann, A. (1998). Parameter estimation in damage processes: Dependent observations of damage increments and first passage time. In: *Advances in Stochastic Models for Reliability, Quality and Safety* W. Kahle, E. von Collani, F. Franz, and U. Jensen (eds.), pp. 139–152. Birkhauser, Boston.
17. Kahle, W. and Wendt, H. (2006). Statistical Analysis of Some Parametric Degradation Models. In: *Probability, Statistics and Modelling in Public Health*, M.S. Nikulin, D. Commenges, and C. Huber (eds.), pp. 266–279. Springer Science and Business Media.
18. Lawless, J. and Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure, *Lifetime Data Analysis*, **10**, 213–227.
19. Lehmann, A. (2001). A Wiener process based model for failure and degradation data in dynamic environments, *Dresdner Schriften zur Mathemat. Stochastik*, **4**, 35–40.
20. Lehmann, A. (2004). On a degradation-failure model for repairable items. In: *Semiparametric Models and its Applications for Reliability, Survival Analysis and Quality of Life*, M.S. Nikulin, N. Balakrishnan, N. Limnios, and M. Mesbah (eds.), pp. 65–79. Birkhauser, Boston.
21. Lehmann, A. (2005). Joint models for degradation and failure time data, *Proceedings of the International Workshop Statistical Modelling and Inference in Life Sciences*, September 1–4, Potsdam, 90–94.

22. Lehmann, A. (2006). Degradation-Threshold-Shock Models. In : *Probability, Statistics and Modelling in Public Health*, M.S. Nikulin, D. Commenges, and C. Huber (eds.), pp. 286–298. Springer, New York.
23. Lu, C.J. and Meeker, W.Q. (1993). Using degradation measures to estimate a time-to-failure distribution, *Technometrics*, **35**, 161–174.
24. Meeker, W. Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*, Wiley, New York.
25. Meeker, W.Q., Escobar, L.A., and Lu, C.J. (1998). Accelerated degradation test: Modeling and analysis, *Technometrics*, **40**, 89–99.
26. Padgett, W.J. and Tomlinson, M.A. (2004). Inference from accelerated degradation and failure data based on Gaussian process models, *Lifetime Data Analysis*, **10**, 191–206.
27. Padgett, W.J. and Tomlinson, M.A. (2005). Accelerated degradation models for failure based on geometric Brownian motion and gamma processes, *Lifetime Data Analysis*, **11**, 511–527.
28. Singpurwalla, N.D. (1995). Survival in dynamic environments, *Statistical Science*, **10**, 86–103.
29. Van Noortwijk, J.M. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, **94**, 2–21.
30. Wendt, H. and Kahle, W. (2004a). On a cumulative damage process and resulting first passage times, *Applied Stochastic Models in Business and Industry*, **20**, 17–26.
31. Wendt, H., Kahle, W. (2004b). On Parametric Estimation for a Position-Dependent Marking of a Doubly Stochastic Poisson Process. In: *Parametric and Semiparametric Models with Applications to Reliability, Survival Analysis, and Quality of Life*, M.S. Nikulin, N. Balakrishnan, M. Mesbah, and N. Limnios), pp. 473–486. Birkhauser, Boston.
32. Whitmore G.A. (1995). Estimation degradation by a Wiener diffusion process subject to measurement error, *Lifetime Data Analysis*, **1**, 307–319.
33. Whitmore, G.A. and Schenkelberg, F. (1997). Modelling accelerated degradation data using Wiener diffusion with a time scale transformation, *Lifetime Data Analysis*, **3**, 27–45.
34. Whitmore G.A., Crowder M.I., and Lawless J. (1998). Failure inference from a marker process based on a bivariate Wiener model, *Lifetime Data Analysis*, **4**, 229–251.
35. Yashin, A.I. (2004). Semiparametric Models in the Studies of Aging and Longevity. In: *Parametric and Semiparametric Models with Applications for Reliability, Survival Analysis, and Quality of Life*, M.S. Nikulin, N. Balakrishnan, N. Limnios, and M. Mesbah (eds.), pp. 149–166. Birkhauser, Boston.
36. Yashin, A.I. and Manton, G.M. (1997). Effects of unobserved and partially observed covariate processes on system failure: A review a models and estimation strategies, *Statistical Science*, **12**, 20–34.
37. Zacks, Sh. (2004). Failure Distributions Associated With General Compound Renewal Damage Processes. In: *Longevity, Aging and Degradation Models in Reliability, Public Health, Medicine and Biology*, v.2, V. Antonov, C. Huber, M.S. Nikulin, and V. Polischook (eds.), pp. 336–344. St.Petersburg State Polytechnical University, St. Petersburg, Russia.

Asymptotic Properties of Redundant Systems Reliability Estimators

V. Bagdonavičius¹, I. Masiulaitytė¹, M.S. Nikulin²

¹ Vilnius University, Vilnius, Lithuania

² IMB, Victor Segalen University, Bordeaux, France

Abstract: Nonparametric and parametric methods of estimation of redundant systems with “warm” standby units reliability are given. Asymptotic properties of the estimators and asymptotic confidence intervals are obtained. Power of goodness-of-fit tests from finite samples is investigated by simulation.

Keywords and phrases: AFT model, confident interval, failure time, maximum likelihood, nonparametric estimation, parameter estimation, redundant system, reliability, Sedyakin’s model, warm standby unit, Weibull distribution

19.1 Introduction

Let us consider redundant systems with one main unit and $m - 1$ standby units operating in “warm” conditions, i.e., under lower stress than the main one. We shall use notation $S(1, m - 1)$ for such systems.

The problem is to obtain confidence intervals for the cumulative distribution functions of redundant systems using failure data of two groups of units, the first group functioning in “hot” and second in “warm” conditions.

We suppose that switching from “warm” to “hot” conditions does not do any damage to units. Bagdonavičius et al. [4] give mathematical formulation of “fluent switch on” and propose tests for verification of this hypothesis. The formulation is based on the “principle of Sedyakin”, see [5].

Denote by T_1 , F_1 , and f_1 the failure time, the c.d.f. and the probability density function of the main unit, respectively. The failure times of the standby units are denoted by T_2, \dots, T_m . In “hot” conditions their distribution functions are also F_1 . In “warm” conditions the c.d.f. of T_i is F_2 and the p.d.f. is f_2 , $i = 2, \dots, m$. If a standby unit is switched to “hot” conditions, its c.d.f. is different from F_1 and F_2 . For $i = 1, 2$ denote by $S_i = 1 - F_i$, $\lambda_i = f_i/S_i$, and $A_i = -\ln S_i$ the survival function, hazard rate, and cumulative hazard, respectively.

The failure time of the system $S(1, m - 1)$ is $T^{(m)} = T_1 \vee T_2 \vee \dots \vee T_m$. Denote by K_j and k_j the c.d.f. and the p.d.f. of $T^{(j)}$, respectively ($j = 2, \dots, m$), $K_1 = F_1$, $k_1 = f_1$. The c.d.f. K_j can be written in terms of the c.d.f. K_{j-1} and F_1 :

$$K_j(t) = \mathbf{P}(T^{(j)} \leq t) = \int_0^t \mathbf{P}(T_j \leq t | T^{(j-1)} = y) dK_{j-1}(y). \quad (19.1)$$

The “fluent switch on” hypothesis H_0 formulated in [4] states that

$$f_{T_j | T^{(j-1)}=y}(t) = \begin{cases} f_2(t) & \text{if } t \leq y, \\ f_1(t + g(y) - y) & \text{if } t > y; \end{cases}, \quad g(y) = F_1^{-1}(F_2(y)). \quad (19.2)$$

This model implies that

$$K_j(t) = \int_0^t F_1(t + g(y) - y) dK_{j-1}(y). \quad (19.3)$$

So the distribution function K_m of the system with $m - 1$ standby units is defined recurrently using the formula (19.3).

In particular, if we suppose that the distribution of units functioning in “warm” and “hot” conditions differ only in scale, i.e., $F_2(t) = F_1(rt)$ for all $t \geq 0$ and some $r > 0$, then $g(y) = ry$. Combining this assumption and the model (2) we have more strict hypothesis H_0^* . This hypothesis can also be considered as generalization of the accelerated failure time (AFT) model (Bagdonavičius [2]) to the case of stress with random switch on.

19.2 Point Estimators of the c.d.f. of Redundant Systems

Suppose that the following data are available:

- (a) complete ordered sample T_{11}, \dots, T_{1n_1} of the failure times of units tested in “hot” conditions;
- (b) the time to obtain complete data in “warm” conditions may be long, so we suppose that n_2 units are tested time t_1 in “warm” conditions and the ordered first failure times T_{21}, \dots, T_{2m_2} are obtained.

19.2.1 Nonparametric estimation

Denote by

$$\hat{F}_j(t) = \frac{1}{n_j} \sum_{i=1}^{n_j} \mathbf{1}_{\{T_{ji} \leq t\}}, \quad \hat{F}_j^{-1}(y) = \inf\{s : \hat{F}_j(s) \geq y\},$$

the empirical distribution function and its inverse, respectively, for the j th sample.

The estimator of the function $g(t)$ is

$$\hat{g}(t) = \hat{F}_1^{-1}(\hat{F}_2(t)), \quad t \leq t_1.$$

Under H_0 for any $t \leq t_1$ the value $K_j(t)$ of the c.d.f is estimated recurrently:

$$\hat{K}_j(t) = \int_0^t \hat{F}_1(t + \hat{g}(y) - y) d\hat{K}_{j-1}(y), \quad \hat{K}_1(t) = \hat{F}_1(t). \quad (19.4)$$

If we suppose that the distribution of units functioning in “warm” and “hot” conditions differ only in scale, i.e., $g(y) = ry$ then the c.d.f. $K_j(t)$ can be estimated at any point $t \geq 0$ replacing $\hat{g}(y)$ by $\hat{r}y$ in (4), where \hat{r} is a convenient estimator of r . Set

$$N_1(t) = \sum_{i=1}^{n_1} \mathbf{1}_{\{T_{1i} \leq t\}}, \quad N_2(t) = \sum_{i=1}^{n_2} \mathbf{1}_{\{T_{2i} \leq t, t \leq t_1\}},$$

$$Y_1(t) = \sum_{i=1}^{n_1} \mathbf{1}_{\{T_{1i} \geq t\}}, \quad Y_2(t) = \sum_{i=1}^{n_2} \mathbf{1}_{\{T_{2i} \geq t, t \leq t_1\}}.$$

Bagdonavičius et al. (2008) give the following estimator of the parameter r :

$$\hat{r} = \tilde{U}^{-1}(0) = \sup\{r : \tilde{U}(r) > 0\};$$

here

$$\tilde{U}(r) = - \int_0^{rt_1} \frac{Y_2(v/r) dN_1(v)}{Y_1(v) + Y_2(v/r)} + \int_0^{t_1} \frac{Y_1(ru) dN_2(u)}{Y_1(ru) + Y_2(u)},$$

is càdlàg stochastic process with trajectories which are nonincreasing step functions satisfying the inequalities $\tilde{U}(0+) > 0$, $\tilde{U}(+\infty) < 0$.

In Figures 19.1 and 19.2 the graphs of the trajectories of the nonparametric estimators \hat{F}_1, \hat{K}_i are given using data simulation when the main unit has Weibull or loglogistic distribution, respectively, and the hypothesis H_0^* is verified. It can be seen from these pictures how the reliability of redundant systems increases including supplementary standby units.

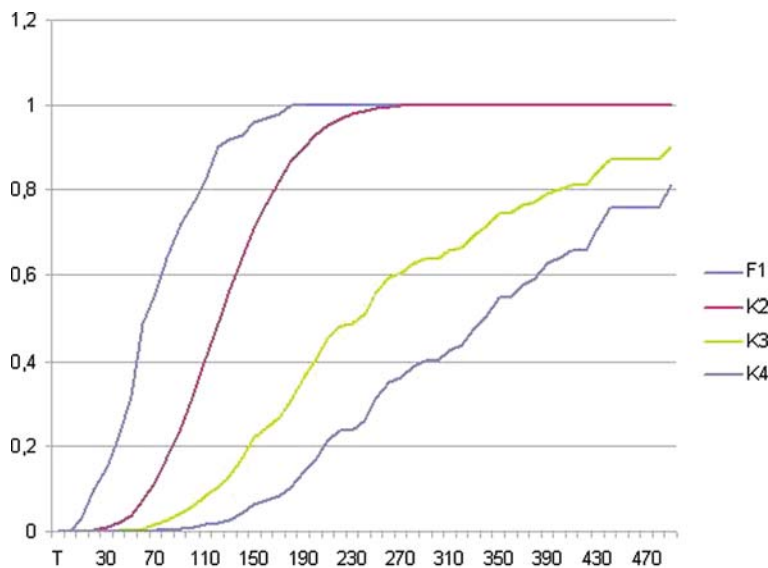


Figure 19.1. Graphs of the trajectories of the nonparametric estimators \hat{F}_1, \hat{K}_i (Weibull distribution)

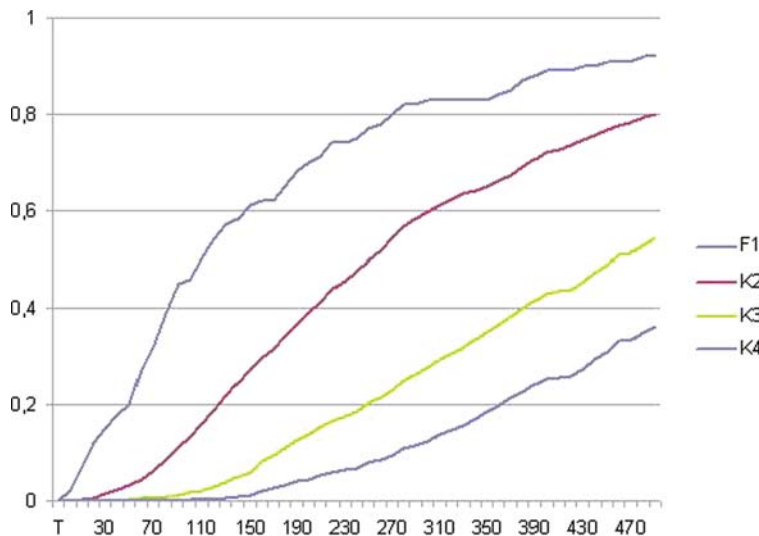


Figure 19.2. Graphs of the trajectories of the nonparametric estimators \hat{F}_1, \hat{K}_i (loglogistic distribution)

19.2.2 Parametric estimation

Suppose that in hot conditions the c.d.f $F_1(t; \theta)$ is absolutely continuous and depends on finite dimensional parameter $\theta \in \Theta \subset \mathbf{R}^k$. Set $\gamma = (r, \theta^T)^T$.

The maximum likelihood estimator $\gamma^* = (r^*, (\theta^*)^T)^T$ of the parameter γ maximizes the loglikelihood function

$$\ell(\gamma) = \sum_{i=1}^{n_1} \ln f_1(T_{1i}; \theta) + m_2 \ln r + \sum_{i=1}^{m_2} \ln f_1(rT_{2i}; \theta) + (n_2 - m_2) \ln S_1(rt_1; \theta).$$

Under H_0^* for any $t \geq 0$ and $j \geq 2$ the c.d.f. $K_j(t)$ is estimated recurrently:

$$\hat{K}_j(t) = \int_0^t F_1(t + r^*y - y; \theta^*) d\hat{K}_{j-1}(y), \quad \hat{K}_1(t) = F_1(t; \theta^*). \quad (19.5)$$

19.3 Asymptotic Distribution of \hat{K}_j and Confidence Intervals for $K_j(t)$

Suppose that

$$\frac{n_i}{n} = l_i + O\left(\frac{1}{n}\right), \quad l_i \in (0, 1), \quad \text{as } n = n_1 + n_2 \rightarrow \infty.$$

19.3.1 Nonparametric case

The limit distribution of the empirical distribution functions is well known:

$$\sqrt{n}(\hat{F}_i - F_i) \xrightarrow{\mathcal{D}} U_i \quad (19.6)$$

on $D(A_i)$, where $\xrightarrow{\mathcal{D}}$ means weak convergence, $A_1 = [0, \infty)$, $A_2 = [0, t_1]$, U_1, U_2 are independent Gaussian martingales with $U_i(0) = 0$ and the covariances

$$\mathbf{cov}(U_i(u), U_i(v)) = \frac{1}{l_i} F_i(u \wedge v) S_i(u \vee v). \quad (19.7)$$

Let us find the asymptotic distribution of the estimator \hat{r} . Denote by $r_0 \in (0, 1)$ the true value of r . Under the model H_0^* it is the ratio of the mean failure times μ_1 and μ_2 of units functioning in “hot” and “warm” conditions, respectively.

Lemma 1. *Suppose that the c.d.f. F_1 is absolutely continuous with positive p.d.f. f_1 on $(0, \infty)$ and the hypothesis $F_2(t) = F_1(r_0 t)$ is true. If*

$$A = -\frac{1}{r_0} \int_0^{r_0 t_1} u f_1(u) dA_1(u) - t_1 f_1(r_0 t_1) \neq 0, \quad (19.8)$$

then

$$\sqrt{n}(\hat{r} - r_0) \xrightarrow{d} Y = -\frac{W}{A}, \quad (19.9)$$

where

$$W = -\int_0^{t_1} [U_1(r_0 u) - U_2(u) dA_2(u) - U_1(r_0 t_1) + U_2(t_1)], \quad (19.10)$$

Proof. Set $\hat{U}(r) = \frac{n}{n_1 n_2} \tilde{U}(r)$, $\hat{S}_i = 1 - \hat{F}_i$. For any $r > 0$

$$\begin{aligned} \hat{U}(r) &= \int_0^{t_1} \frac{\hat{S}_2(u-) d\hat{S}_1(ru)}{\frac{n_1}{n} \hat{S}_1(ru-) + \frac{n_2}{n} \hat{S}_2(u-)} - \int_0^{t_1} \frac{\hat{S}_1(ru-) d\hat{S}_2(u)}{\frac{n_1}{n} \hat{S}_1(ru-) + \frac{n_2}{n} \hat{S}_2(u-)} \\ &=: \int_0^{t_1} \hat{Z}_2(u, r) d\hat{S}_1(ru) - \int_0^{t_1} \hat{Z}_1(u, r) d\hat{S}_2(u), \quad \hat{Z}_2(u, r) = \frac{n}{n_2} \left(1 - \frac{n_1}{n} \hat{Z}_1(u, r)\right). \end{aligned}$$

The convergence $\sup_{u \in A_i} |\hat{S}_i(u) - S_i(u)| \xrightarrow{P} 0$, $\sup_{u \in A_i} |\hat{S}_i(u-) - S_i(u)| \xrightarrow{P} 0$ implies that $\hat{U}(r) \xrightarrow{P} U(r)$, where

$$\begin{aligned} U(r) &= \int_0^{t_1} \frac{S_2(u) dS_1(ru)}{l_1 S_1(ru) + l_2 S_2(u)} - \int_0^{t_1} \frac{S_1(ru) dS_2(u)}{l_1 S_1(ru) + l_2 S_2(u)} \\ &=: \int_0^{t_1} Z_2(u, r) dS_1(ru) - \int_0^{t_1} Z_1(u, r) dS_2(u), \quad Z_2(u, r) = \frac{1}{l_2} (1 - l_1 Z_1(u, r)). \end{aligned}$$

Note that $U(r_0) = 0$.

By functional delta method

$$\sqrt{n}(\hat{Z}_1 - Z_1)(u, r) \xrightarrow{d} l_2 \left(\frac{-U_1(ru)S_2(u) + U_2(u)S_1(ru)}{(l_1S_1(ru) + l_2S_2(u))^2} \right) =: U_1^*(u, r)$$

and

$$\sqrt{n}(\hat{Z}_2 - Z_2)(u, r) \xrightarrow{d} -\frac{l_1}{l_2}U_1^*(u, r) =: U_2^*(u, r)$$

on $D([0, t_1] \times [0, 1])$. Note that

$$U_1^*(u, r_0) = l_2 \frac{U_2(u) - U_1(r_0u)}{S_2(u)}, \quad Z_1(u, r_0) \equiv 1.$$

By the functional delta method for stochastic integrals

$$\begin{aligned} \sqrt{n}(\hat{U}(r) - U(r)) \xrightarrow{d} W(r) &:= \int_0^{t_1} (U_2^*(u, r)dS_1(ru) - Z_2(u, r)dU_1(ru) \\ &\quad - \int_0^{t_1} (U_1^*(u, r)dS_2(u) - Z_1(u, r)dU_2(u)) \end{aligned} \quad (19.11)$$

on $[0, 1]$. By the functional delta method

$$\sqrt{n}(\hat{r} - r_0) \rightarrow Y = -\frac{W(r_0)}{U'(r_0)}.$$

Note that $W = W(r_0)$ and $A = U'(r_0)$ are given by formulas (19.9) and (19.10).

The proof is complete.

Remark 1. If samples are complete then

$$W = -\int_0^\infty [U_1(r_0u) - U_2(u)]d\Lambda_2(u), \quad A = -\frac{1}{r_0} \int_0^\infty u f_1(u)d\Lambda_1(u) \quad (19.12)$$

and

$$\sqrt{n}(\hat{r} - r_0) \xrightarrow{\mathcal{D}} Y = -\frac{W}{A} \sim N\left(0, \frac{1}{l_1 l_2 A^2}\right). \quad (19.13)$$

Theorem 1. *If F_1 is continuously differentiable on $[0, \infty)$ then under H_0^* for any $t > 0$ and any natural $j \geq 2$*

$$\begin{aligned} \sqrt{n}(\hat{K}_j(s) - K_j(s)) &\xrightarrow{\mathcal{D}} W_j(s) \\ &= \int_0^s U_1(s + r_0y - y)dK_{j-1}(y) + \mu^{(j-1)}(s)Y \\ &\quad + \int_0^s F_1(s + r_0y - y)dW_{j-1}(y) \end{aligned} \quad (19.14)$$

on $D[0, t]$, where $W_1(s) = U_1(s)$, $\mu^{(j-1)}(s) = \int_0^s y f_1(s + r_0y - y)dK_{j-1}(y)$.

Proof. Let us prove that under conditions of Lemma 1 for any $t \geq 0$

$$\sqrt{n}(\hat{F}_1(t + \hat{r}y - y) - F_1(t + r_0y - y)) \xrightarrow{d} U_1(t + r_0y - y) + yf_1(t + r_0y - y)Y \quad (19.15)$$

on $D[0, t]$. It is sufficient to verify the conditions of the following theorem (see [1]):

Suppose that

- (1) x is a continuously differentiable function on $[0, \tau]$;
- (2) $\varphi = \varphi(t, \theta) : A \times B_\varepsilon(\theta_0) \rightarrow \mathbf{R}$, $B_\varepsilon(\theta_0) \subset \mathbf{R}^s$, $A = [0, \tau_0]$ or $(0, \tau_0)$ is a continuous nonincreasing in t function such that $0 < \varphi(t, \theta_0) < \tau$ for $t \in A$;
- (3) $\{X^{(n)} \in D[0, \tau]\}$ is a sequence of stochastic processes such that

$$\sqrt{n}(X^n - x) \xrightarrow{\mathcal{D}} Z$$

on $D[0, \tau]$, where Z is a continuous on $[0, \tau]$ stochastic process;

- (4) $\{\hat{\theta}^{(n)}\}$ is a sequence of random variables such that

$$\sqrt{n}(\hat{\theta}^{(n)} - \theta_0) \xrightarrow{\mathcal{D}} Y.$$

Then

$$\sqrt{n}(X^n(\varphi(\cdot, \hat{\theta}_0^{(n)})) - x(\varphi(\cdot, \theta_0))) \xrightarrow{\mathcal{D}} Z(\varphi(\cdot, \theta_0)) + x'(\varphi(\cdot, \theta_0))\varphi'_\theta(\cdot, \theta_0)Y$$

on $D(A)$.

Take $x = F_1$, $X^n = \hat{F}_1$, $\theta = r$, $\varphi(y, r) = t + ry - y$, $\tau = t + \Delta$, Δ – any positive real number, $\tau_0 = t$, $Z = U_1$. All four conditions are verified because $\varphi(0, r_0) = t < \tau$, $\varphi(t, r_0) = r_0t > 0$. So (19.15) is true.

We prove the theorem by induction. If $j = 2$ then by the functional delta method for integrals and using (19.4) (taking $g(y) = y$), (19.6), and (19.15), we obtain

$$\begin{aligned} & \sqrt{n}(\hat{K}_2(s) - K_2(s)) \\ & \xrightarrow{\mathcal{D}} \int_0^s U_1(s + r_0y - y)dF_1(y) + \mu^{(1)}(s)Y \\ & \quad + \int_0^s F_1(s + r_0y - y)dU_1(y) = W_2(s) \end{aligned} \quad (19.16)$$

on $D[0, t]$. So (14) is true for $j = 2$. Suppose that it is true for some $j = l \geq 2$. Then by the functional delta method for integrals and using (19.6), (19.15) we obtain the result (19.14) for $j = l + 1$.

The proof is complete.

The asymptotic variance of $\sqrt{n}(\hat{K}_j(t) - K_j(t))$, $j \geq 2$, might be estimated recurrently, using Equation (19.14): the covariances $\mathbf{Cov}(W_j(s), W_j(t)) = \mathbf{E}(W_j(s)W_j(t))$ can be written in terms of the covariances

$$\begin{aligned} & \mathbf{E}(W_{j-1}(u)W_{j-1}(v)), \quad \mathbf{E}(W_{j-1}(u)U_1(v)), \quad \mathbf{E}(W_{j-1}(u)U_2(v)), \\ & \mathbf{E}(U_1(u)U_1(v)), \quad \mathbf{E}(U_2(u)U_2(v)). \end{aligned}$$

Note that for $j = 2$ these covariances are

$$\mathbf{E}(W_1(u)W_1(v)) = \mathbf{E}(W_1(u)U_1(v)) = \mathbf{E}(U_1(u)U_1(v)) = \frac{1}{l_1} F_1(u \wedge v)S_1(u \vee v),$$

$$\mathbf{E}(W_1(u)V_2(v)) = \mathbf{E}(U_1(u)U_2(v)) = 0, \quad \mathbf{E}(U_2(u)U_2(v)) = \frac{1}{l_2} F_2(u \wedge v)S_2(u \vee v).$$

Let us find the asymptotic variance of $\sqrt{n}(\hat{K}_2(t) - K_2(t))$. Suppose first that samples are complete. In the following we skip the index in r_0 .

The formulas (19.12), (19.13), and (19.16) imply that $\sqrt{n}(\hat{K}_2(s) - K_2(s)) \xrightarrow{\mathcal{D}} W_2(s)$, where

$$\begin{aligned} W_2(t) &= F_2(t)U_1(t) + \int_0^t U_1(t+ry-y)dF_1(y) - \int_0^t U_1(y)dF_1(t+ry-y) \\ &\quad + \frac{\mu(t)}{A} \left(\int_0^\infty U_1(ry)d\Lambda_2(y) - \int_0^\infty U_2(y)d\Lambda_2(y) \right) \\ &= (V_1 + V_2 + V_3 + V_4)(t). \end{aligned} \quad (19.17)$$

$$\mu(t) = \mu^{(1)}(t) = \int_0^t y f_1(t+ry-y)dF_1(y), \quad A = -\frac{1}{r} \int_0^\infty u f_1(u)d\Lambda_1(u). \quad (19.18)$$

The random variable $W_2(t)$ has zero mean. Set $\nu(t) = \int_0^t F_1(t+ry-y)dF_1(y)$. For any $t \geq 0$ the variances and the covariances of the random variables $V_i(t)$ multiplied by l_1 are

$$\begin{aligned} l_1 \mathbf{E}V_1^2(t) &= F_2^2(t)F_1(t)S_1(t), \\ l_1 \mathbf{E}V_2^2(t) &= 2 \int_0^t F_1(z)F_1(t+rz-z)dF_1(z) - \nu^2(t), \\ l_1 \mathbf{E}V_3^2(t) &= 2F_2(t)(F_1(t)F_2(t) - \nu(t)) - 2 \int_0^t F_1(z)F_1(t+rz-z)dF_1(t+rz-z) \\ &\quad - (F_1(t)F_2(t) - \nu(t))^2, \\ l_1 \mathbf{E}V_4^2(t) &= \frac{\mu^2(t)}{l_2 A^2}, \quad l_1 \mathbf{E}V_1(t)V_2(t) = S_1(t)F_2(t)\nu(t), \\ l_1 \mathbf{E}V_1(t)V_3(t) &= -S_1(t)F_2(t)(F_1(t)F_2(t) - \nu(t)), \\ l_1 \mathbf{E}V_1(t)V_4(t) &= -F_2(t)\frac{\mu(t)}{A}S_1(t)\ln S_1(t), \\ -l_1 \mathbf{E}V_2(t)V_3(t) &= F_1(t) \int_0^{rt} F_1(y)dF_1(t+ry-y) \\ &\quad + \int_{rt}^t F_1((t-y)/(1-r))F_1(y)dF_1(t+ry-y) + F_2(t)\nu(t) \\ &\quad + \int_{rt}^t F_1(t+rz-z)F_1(z)dF_1((t-z)/(1-r)) \\ &\quad - \nu(t)(F_1(t)F_2(t) - \nu(t)), \\ l_1 \mathbf{E}V_2(t)V_4(t) &= -\frac{\mu(t)}{A} \int_0^t S_1(t+ry-y)\ln S_1(t+ry-y)dF_1(y), \\ l_1 \mathbf{E}V_3(t)V_4(t) &= \frac{\mu(t)}{A} \int_0^t S_1(y)\ln S_1(y)dF_1(t+ry-y). \end{aligned}$$

So the variance $\mathbf{Var}(W_2(t))$ is defined by the following formula:

$$\begin{aligned} l_1 \mathbf{Var}(W_2(t)) &= -4\nu^2(t) + \int_0^t F_1(t+ry-y)[F_1(t+ry-y) + 2F_1(y)] dF_1(y) \\ &\quad + 2F_1(t)\nu(rt) + 2 \int_{rt}^t F_1(t+ry-y)F_1((t-y)/(1-r)) dF_1(y) + \frac{\mu^2(t)}{l_2 A^2} \\ &\quad + \frac{2\mu(t)}{A} \left[\nu(t) + \int_0^t [F_1(t+ry-y) \ln S_1(y) - S_1(t+ry-y) \ln S_1(t+ry-y)] dF_1(y) \right]. \end{aligned}$$

Set

$$\begin{aligned} Z_{1i} &= \hat{F}_1(t + (\hat{r} - 1)T_{1i}-), \quad \hat{F}_1(t-) = \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{1}_{\{T_{1i} < t\}}, \quad Z_{2i} = \hat{F}_1\left(\frac{t - T_{1i}}{1 - \hat{r}}-\right), \\ Z_{3i} &= \hat{F}_1(T_{1i}-), \quad Z_{4i} = \hat{f}_1(t + (\hat{r} - 1)T_{1i}-), \quad \hat{\mu}(t) = \frac{1}{n_1} \sum_{T_{1i} \leq t} T_{1i} Z_{4i}, \quad Z_{5i} = \hat{f}_1(T_{1i}-). \end{aligned}$$

The variance $\mathbf{Var}(W_2(t))$ is estimated using the statistic

$$\frac{n_1}{n} \hat{\mathbf{Var}}(W_2(t)) = -4\hat{\phi}_1^2(t) + \hat{\phi}_2(t) + \frac{n\hat{\mu}^2(t)}{n_2 \hat{A}^2} + \frac{2\hat{\mu}(t)}{\hat{A}} \hat{\phi}_3(t);$$

here

$$\begin{aligned} \hat{\phi}_1(t) &= \frac{1}{n_1} \sum_{T_{1i} \leq t} Z_{1i}, \quad \hat{A} = -\frac{1}{\hat{r} n_1} \sum_{i=1}^{n_1} \frac{T_{1i} Z_{5i}}{1 - Z_{3i}}, \\ \hat{\phi}_2(t) &= \frac{1}{n_1} \sum_{T_{1i} \leq t} Z_{1i} [Z_{1i} + 2Z_{3i} + 2\hat{F}_1(t) \mathbf{1}_{\{T_{1i} \leq \hat{r}t\}} + 2Z_{2i} \mathbf{1}_{\{T_{1i} > \hat{r}t\}}], \\ \hat{\phi}_3(t) &= \frac{1}{n_1} \sum_{T_{1i} \leq t} [Z_{1i}(1 + \ln(1 - Z_{3i})) - (1 - Z_{1i}) \ln(1 - Z_{1i})]. \end{aligned}$$

So the variance $\sigma_{\hat{K}_2}^2$ of the estimator $\hat{K}_2(t)$ is estimated by

$$\hat{\sigma}_{\hat{K}_2}^2 = \frac{1}{n} \hat{\mathbf{Var}}(W_2(t)).$$

The asymptotic $1 - \alpha$ confidence interval for $K_2(t)$ is $(\underline{K}_2(t), \overline{K}_2(t))$, where

$$\begin{aligned} \underline{K}_2(t) &= \left(1 + \frac{1 - \hat{K}_2(t)}{\hat{K}_2(t)} \exp \left\{ \frac{\hat{\sigma}_{\hat{K}_2} z_{1-\alpha/2}}{\sqrt{\hat{K}_2(t)(1 - \hat{K}_2(t))}} \right\} \right)^{-1}, \\ \overline{K}_2(t) &= \left(1 + \frac{1 - \hat{K}_2(t)}{\hat{K}_2(t)} \exp \left\{ -\frac{\hat{\sigma}_{\hat{K}_2} z_{1-\alpha/2}}{\sqrt{\hat{K}_2(t)(1 - \hat{K}_2(t))}} \right\} \right)^{-1}. \end{aligned} \quad (19.19)$$

Remark 2. In the case of censoring the expression in parenthesis of the term V_4 in (19.17) is replaced by $\int_0^{t_1} [U_1(ru) - U_2(u)d\Lambda_2(u) + U_1(rt_1) + U_2(t_1)]$, so only minor modifications are needed.

We investigated finite sample confidence level of the proposed asymptotic confidence intervals. The failure times T_{1j} and T_{2j} were simulated from exponential distribution:

$$T_{1j} \sim \mathcal{E}(\lambda_1), \quad T_{2j} \sim \mathcal{E}(\lambda_2), \quad \lambda_1 = \frac{1}{100}; \quad \lambda_2 = \frac{1}{300}.$$

The number of replications was 2000. For various values of t the proportions of confidence interval realizations covering the true value of the distributional function $K_2(t)$ are given in Table 19.1.

Table 19.1. Confidence level for finite samples ($n_1 = n_2 = 100$)

Time, t	50	100	150	200	300	400	500
$K_2(t)$	0.114	0.319	0.513	0.667	0.856	0.941	0.977
Confidence level (%)	91.1	90.8	90.5	90.3	89.8	89.4	89.1

19.3.2 Parametric case

Denote by $I_n(\gamma) = -\mathbf{E}\ddot{\ell}(\gamma)$ the Fisher information matrix and suppose that $\frac{1}{n}I_n(\gamma) \rightarrow i(\gamma)$. Under classical assumptions on the family of distributions $f_1(t, \theta)$ the maximum likelihood estimator γ^* is asymptotically normal:

$$\sqrt{n}(\gamma^* - \gamma) \xrightarrow{d} Y = (Y_1, Y_2^T)^T \sim N_{k+1}(0, i^{-1}(\gamma)).$$

Y_1 is one dimensional Y_2 k dimensional.

Using delta method we obtain:

$$\sqrt{n}(\hat{K}_2(t) - K_2(t)) \xrightarrow{\mathcal{D}} W_2(t) = C_2^T(t; \gamma)Y,$$

where

$$C_2(t; \gamma) = (C_{21}(t; \gamma), C_{22}^T(t; \gamma))^T, \quad C_{21}(t; \gamma) = \int_0^t \frac{\partial}{\partial r} F_1(t + ry - y; \theta) dF_1(y; \theta),$$

$$C_{22}(t; \gamma) = \int_0^t \frac{\partial}{\partial \theta} F_1(t + ry - y; \theta) dF_1(y; \theta) + F_1(t + ry - y; \theta) d\left(\frac{\partial}{\partial \theta} F_1(y; \theta)\right).$$

The random variable $W_2(t)$ is a linear function of Y .

If $j \geq 2$ then

$$\sqrt{n}(\hat{K}_j(t) - K_j(t)) \xrightarrow{\mathcal{D}} W_j(t).$$

The random variable $W_j(t)$, $j \geq 2$, is also a linear function of Y :

$$W_j(t) = Y^T C_j(t; \gamma), \quad C_j(t; \gamma) \in (C[0, t])^{k+1}.$$

It follows by induction:

$$W_j(t) = Y^T \left(\int_0^t \frac{\partial}{\partial \gamma} F_1(t + ry - y; \theta) dK_{j-1}(y; \gamma) + F_1(t + ry - y; \theta) dC_{j-1}(t; \gamma) \right).$$

So the variance

$$\mathbf{Var}(W_j(t)) = \mathbf{Var}(C_j(t; \gamma)^T Y) = C_j^T(t; \gamma) i^{-1}(\gamma) C_j(t; \gamma)$$

is estimated by $n C_2^T(t; \hat{\gamma}) I^{-1}(\hat{\gamma}) C_j(t; \hat{\gamma})$, and the variance $\sigma_{\hat{K}_j(t)}^2$ of the estimator $\hat{K}_j(t)$ is estimated by

$$\hat{\sigma}_{\hat{K}_j(t)}^2 = C_j^T(t; \hat{\gamma}) I^{-1}(\hat{\gamma}) C_j(t; \hat{\gamma}).$$

The matrix $I(\hat{\gamma})$ may be replaced by $-\ddot{\ell}(\hat{\gamma})$.

The asymptotic $1 - \alpha$ confidence interval for $K_j(t)$ is $(\underline{K}_j(t), \overline{K}_j(t))$, where

$$\begin{aligned} \underline{K}_j(t) &= \left(1 + \frac{1 - \hat{K}_j(t)}{\hat{K}_j(t)} \exp \left\{ \frac{\hat{\sigma}_{\hat{K}_j} z_{1-\alpha/2}}{\sqrt{\hat{K}_j(t)(1 - \hat{K}_j(t))}} \right\} \right)^{-1}, \\ \overline{K}_j(t) &= \left(1 + \frac{1 - \hat{K}_j(t)}{\hat{K}_j(t)} \exp \left\{ -\frac{\hat{\sigma}_{\hat{K}_j} z_{1-\alpha/2}}{\sqrt{\hat{K}_j(t)(1 - \hat{K}_j(t))}} \right\} \right)^{-1}. \end{aligned} \quad (19.20)$$

Example 1. Exponential distribution: $S_1(t) = e^{-\lambda t}$.

Suppose that samples are complete. The logarithm of the parametric likelihood function is

$$\ell(r, \mu, \nu) = n \ln \lambda - \lambda \left(\sum_{i=1}^{n_1} T_{1i} + r \sum_{i=1}^{n_2} T_{2i} \right) + n_2 \ln r,$$

the MLE estimators are $r^* = \hat{\mu}_1 / \hat{\mu}_2$, $\lambda^* = 1 / \hat{\mu}_1$, where $\hat{\mu}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} T_{ji}$, $j = 1, 2$. The Fisher information matrix is

$$I(r, \lambda) = \begin{pmatrix} \frac{n_2}{r^2} & \frac{n_2}{r\lambda} \\ \frac{n_2}{r\lambda} & \frac{n}{\lambda^2} \end{pmatrix}, \quad \frac{1}{n} I(r, \lambda) \rightarrow i(r, \lambda) = \begin{pmatrix} \frac{l_2}{r^2} & \frac{l_2}{r\lambda} \\ \frac{l_2}{r\lambda} & \frac{1}{\lambda^2} \end{pmatrix}, \quad i^{-1}(r, \lambda) = \begin{pmatrix} \frac{r^2}{l_1 l_2} & -\frac{r\lambda}{l_1} \\ -\frac{r\lambda}{l_1} & \frac{\lambda^2}{l_1} \end{pmatrix}.$$

The partial derivatives are

$$\frac{\partial}{\partial r} F_1(t + ry - y; \lambda) = \lambda y e^{-\lambda(t+ry-y)}, \quad \frac{\partial}{\partial \lambda} F_1(t + ry - y; \lambda) = (t + ry - y) e^{-\lambda(t+ry-y)}.$$

Note that $K_1(t) = F_1(t)$.

So $C_2 = (C_{21}, C_{22})^T$, where

$$C_{21}(t; r, \lambda) = \frac{1}{r^2} S_1(t) [F_2(t) - \lambda r t S_2(t)], \quad C_{22}(t; r, \lambda) = S_1(t) F_2(t) \frac{1+r}{r} t.$$

The equality

$$\mathbf{Var}(W_2(t)) = \mathbf{Var}(C_2(t; r, \lambda)^T Y) = \frac{S_1^2(t)}{l_1 l_2 r^2} \{ l_1 [F_2(t) - \lambda r t S_2(t)]^2 + l_2 [(1 - \lambda t) F_2(t) - \lambda r t]^2 \}$$

implies that

$$\hat{\sigma}_{\hat{K}_2(t)}^2 = \frac{\hat{S}_1^2(t)}{n_1 n_2 \hat{r}^2} \{n_1 [\hat{F}_2(t) - \lambda^* r^* t \hat{S}_2(t)]^2 + n_2 [(1 - \lambda^* t) \hat{F}_2(t) - \lambda^* r^* t]^2\},$$

where

$$\hat{S}_1(t) = e^{-\lambda^* t}, \quad \hat{S}_2(t) = e^{-\lambda^* r^* t}, \quad \hat{F}_i(t) = 1 - \hat{S}_i(t), \quad i = 1, 2.$$

The asymptotic $(1 - \alpha)$ confidence interval for K_2 has the form (19) with

$$\hat{K}_2(t) = \hat{F}_1(t) - \frac{\hat{S}_1(t) \hat{F}_2(t)}{r^*},$$

Taking into consideration the equality

$$K_2(t) = F_1(t) - \frac{S_1(t) F_2(t)}{r},$$

the weights $C_3 = (C_{31}, C_{32})^T$ and the estimator

$$\hat{\sigma}_{\hat{K}_3(t)}^2 = C_3^T(t; \hat{\gamma}) I^{-1}(\hat{\gamma}) C_3(t; \hat{\gamma})$$

can be computed and the asymptotic confidence interval for $K_3(t)$ can be found.

We investigated finite sample confidence level of the proposed asymptotic confidence intervals. The same simulated data was used as in Section 19.3.1. For various values of t the proportions of confidence interval realizations covering the true value of the distributional function $K_2(t)$ are given in Table 19.2.

Table 19.2. Confidence level (parametric estimation) for finite samples ($n_1 = n_2 = 100$)

Time, t	50	100	150	200	300	400	500
$K_2(t)$	0.114	0.319	0.513	0.667	0.856	0.941	0.977
Confidence level (%)	91.75	91.0	91.7	91.6	91.1	89.5	88.9

Example 2. Weibull distribution: $S_1(t) = e^{-(t/\mu)^\nu}$.

The logarithm of the parametric likelihood function is

$$\begin{aligned} \ell(r, \mu, \nu) = & n(\ln \nu - \nu \ln \mu) + n_2 \nu \ln r + (\nu - 1) \left(\sum_{i=1}^{n_1} \ln T_{1i} + \sum_{i=1}^{n_2} \ln T_{2i} \right) \\ & - \mu^{-\nu} \left(\sum_{i=1}^{n_1} T_{1i}^\nu + r^\nu \sum_{i=1}^{n_2} T_{2i}^\nu \right), \end{aligned}$$

and the Fisher information matrix

$$I(r, \mu, \nu) = \begin{pmatrix} \frac{n_2 \nu^2}{r^2} & -\frac{n_2 \nu^2}{r \nu} & \frac{n_2 \Gamma'(2)}{r} \\ -\frac{n_2 \nu^2}{r \nu} & \frac{n \nu^2}{\mu^2} & -\frac{n \Gamma'(2)}{\mu} \\ \frac{n_2 \Gamma'(2)}{r} & -\frac{n \Gamma'(2)}{\mu} & \frac{n(1 + \Gamma''(2))}{\mu^2} \end{pmatrix}.$$

The partial derivatives are

$$\begin{aligned}\frac{\partial}{\partial r}F_1(t+ry-y;\mu,\nu) &= \frac{\nu y}{\mu} \left(\frac{t+ry-y}{\mu} \right)^{\nu-1} S_1(t+ry-y), \\ \frac{\partial}{\partial \mu}F_1(t+ry-y;\mu,\nu) &= -\frac{\nu}{\mu} \left(\frac{t+ry-y}{\mu} \right)^{\nu} S_1(t+ry-y), \\ \frac{\partial}{\partial \nu}F_1(t+ry-y;\mu,\nu) &= \left(\frac{t+ry-y}{\mu} \right)^{\nu} \ln \left(\frac{t+ry-y}{\mu} \right) S_1(t+ry-y).\end{aligned}$$

So $C_2 = (C_{21}, C_{22}, C_{23})^T$, where

$$\begin{aligned}C_{21}(t; r, \mu, \nu) &= \int_0^t \frac{\partial}{\partial r} F_1(t+ry-y; \mu, \nu) dF_1(y), \\ C_{22}(t; r, \mu, \nu) &= \int_0^t \frac{\partial}{\partial \mu} F_1(t+ry-y; \mu, \nu) dF_1(y) + F_1(t+ry-y; \mu, \nu) d \left(\frac{\partial}{\partial \mu} F_1(y; \mu, \nu) \right), \\ C_{23}(t; r, \mu, \nu) &= \int_0^t \frac{\partial}{\partial \nu} F_1(t+ry-y; \mu, \nu) dF_1(y) + F_1(t+ry-y; \mu, \nu) d \left(\frac{\partial}{\partial \nu} F_1(y; \mu, \nu) \right).\end{aligned}$$

The last three integrals can be computed numerically.

Let us investigate the efficiency $AE(\hat{r}, r^*)$ of the nonparametric estimator \hat{r} with respect to the efficient parametric maximum likelihood estimator r^* for some parametric models used in reliability (complete samples).

Example 1. (continuation): exponential distribution.

In this case $A = -1/r_0$, so by (19.12)

$$\sqrt{n}(\hat{r} - r_0) \xrightarrow{\mathcal{D}} \hat{Y} \sim N \left(0, \frac{r_0^2}{l_1 l_2} \right).$$

The parametric MLE

$$r^* = \frac{\hat{\mu}_1}{\hat{\mu}_2}, \quad \hat{\mu}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} T_{ji}.$$

Here $\hat{\mu}_j$ is the estimator of the mean $\mu_j = \mathbf{E}(T_{ji})$, $i = 1, \dots, n_i$, $j = 1, 2$. Set $\sigma_j^2 = \mathbf{Var}(T_{ji})$, $j = 1, 2$. The convergence

$$\sqrt{n}(\hat{\mu}_j - \mu_j) \xrightarrow{\mathcal{D}} Y_j^* = - \int_0^\infty U_j(y) dy \sim N(0, \sigma_j^2/l_j)$$

implies

$$\sqrt{n}(r^* - r_0) \xrightarrow{\mathcal{D}} \tilde{Y} = \frac{1}{\mu_2}(rY_2^* - Y_1^*) \sim N \left(0, \frac{\sigma_1^2}{l_1 l_2 \mu_2^2} \right) = N \left(0, \frac{r_0^2}{l_1 l_2} \right).$$

So $AE(\hat{r}, r^*) = 1$.

Example 2. (continuation): Weibull distribution.

The first diagonal element of the inverse matrix I^{-1} is

$$I^{11} = \frac{nr^2}{n_1 n_2 \nu^2},$$

So

$$\sqrt{n}(r^* - r_0) \xrightarrow{\mathcal{D}} Y \sim N\left(0, \frac{r_0^2}{l_1 l_2 \nu^2}\right),$$

which means that in the Weibull case the nonparametric estimator \hat{r} has the same asymptotic efficiency as the parametric MLE r^* .

Example 3. Loglogistic distribution: $S_1(t) = \frac{1}{1+(t/\mu)^\nu}$.

The logarithm of the parametric likelihood function is

$$\begin{aligned} \ell(r, \mu, \nu) = & n(\ln \nu - \nu \ln \mu) + n_2 \nu \ln r + (\nu - 1) \left(\sum_{i=1}^{n_1} \ln T_{1i} + \sum_{i=1}^{n_2} \ln T_{2i} \right) \\ & - 2 \left(\sum_{i=1}^{n_1} \ln \left(1 + \left(\frac{T_{1i}}{\mu} \right)^\nu \right) + \sum_{i=1}^{n_2} \ln \left(1 + \left(\frac{r T_{2i}}{\mu} \right)^\nu \right) \right), \end{aligned}$$

and the Fisher information matrix

$$I(r, \mu, \nu) = \begin{pmatrix} \frac{n_2 \nu^2}{3r^2} & -\frac{n_2 \nu^2}{3r\mu} & 0 \\ -\frac{n_2 \nu^2}{3r\mu} & \frac{n\nu^2}{3\mu^2} & 0 \\ 0 & 0 & \frac{n(3+2\Gamma''(2)-2(\Gamma'(2))^2)}{3\mu^2} \end{pmatrix}.$$

The first diagonal element of the inverse matrix I^{-1} is

$$I^{11} = \frac{3nr^2}{n_1 n_2 \nu^2}.$$

So

$$\sqrt{n}(r^* - r_0) \xrightarrow{\mathcal{D}} Y \sim N\left(0, \frac{3r_0^2}{l_1 l_2 \nu^2}\right),$$

which means that differently from the Weibull case relative asymptotic efficiency of the nonparametric estimator \hat{r} with respect to the parametric MLE r^* is small: $ASE(\hat{r}, r^*) = 1/\sqrt{3} = 0,58$.

19.4 Power of Goodness-of-Fit Tests

Bagdonavičius et al. [4, 3] give goodness-of-fit tests for the “fluent switch on” hypotheses H_0 and H_0^* . Let us consider the power of the tests.

The test statistics are constructed as follows. Suppose that the following data are available:

- (a) the failure times T_{11}, \dots, T_{1n_1} of n_1 units tested in “hot” conditions;
- (b) the failure times T_{21}, \dots, T_{2n_2} of n_2 units tested in “warm” conditions;
- (c) the failure times T_1, \dots, T_n of n redundant systems (with “warm” standby units).

The tests are based on the difference of two estimators of the c.d.f. F . The first estimator is the empirical distribution function

$$\hat{F}^{(1)}(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{T_i \leq t\}}.$$

The second is based on the formula (3), i.e.,

$$\hat{F}^{(2)}(t) = \int_0^t \hat{F}_1(t + \hat{g}(y) - y) d\hat{F}_1(y),$$

where (hypothesis H_0)

$$\hat{g}(y) = \hat{F}_1^{-1}(\hat{F}_2(y)), \quad \hat{F}_j(t) = \frac{1}{n_j} \sum_{i=1}^{n_j} \mathbf{1}_{\{T_{ji} \leq t\}}, \quad \hat{F}_1^{-1}(y) = \inf\{s : \hat{F}_1(s) \geq y\},$$

or (hypothesis H_0^*)

$$\hat{g}(y) = \hat{r}y, \quad \hat{r} = \frac{\hat{\mu}_1}{\hat{\mu}_2}, \quad \hat{\mu}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} T_{ji}.$$

The test is based on the statistic

$$X = \sqrt{n} \int_0^\infty [\hat{F}^{(1)}(t) - \hat{F}^{(2)}(t)] dt.$$

Bagdonavičius et al. [3] proved that in the case of both hypothesis H_0 and H_0^* the limit distribution (as $n_i/n \rightarrow l_i \in (0, 1)$, $n \rightarrow \infty$) of the statistic X is normal with zero mean and finite variance σ^2 .

The test statistic is

$$T = \frac{X}{\hat{\sigma}},$$

where $\hat{\sigma}$ is a consistent estimator of σ . The distribution of the statistic T is approximated by the standard normal distribution and the hypothesis H_0 (or H_0^*) is rejected with approximative significance value α if $|T| > z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ -quantile of the standard normal distribution.

Bagdonavičius et al. [3] proved that consistent estimators of the variance σ^2 are

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (T_i - \hat{\mu})^2 + \frac{n}{n_1^2} \sum_{i=1}^{n_1} [\hat{H}(T_{1i}) - \hat{H}]^2 + \frac{n}{n_2^2} \sum_{i=1}^{n_2} [\hat{Q}(T_{2i}) - \hat{Q}]^2,$$

where

$$\hat{H}(x) = \hat{Q}(x) - x\hat{F}_1(\hat{g}^{-1}(x)) + \hat{g}(x)[1 - \hat{F}_2(x)] + \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{1}_{\{\hat{g}(T_{1i}) \leq x\}} \hat{g}(T_{1i})$$

$$+ \frac{1}{n_2} \sum_{i=1}^{n_2} \mathbf{1}_{\{T_{2i} \leq x\}} \hat{g}(T_{2i}) - x,$$

$$\hat{Q}(x) = \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{1}_{\{T_{1i} \leq x\}} [1 - \hat{F}_2(T_{1i})] / \hat{f}_1(\hat{g}(T_{1i})),$$

$$\hat{g}^{-1}(x) = \hat{F}_2^{-1}(\hat{F}_1(x)), \quad \hat{H} = \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{H}(T_{1i}), \quad \hat{Q} = \frac{1}{n_2} \sum_{i=1}^{n_2} \hat{Q}(T_{2i}),$$

the density f_1 is estimated by the kernel estimator

$$\hat{f}_1(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x - X_{1i}}{h}\right)$$

(hypothesis H_0) or

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (T_i - \hat{\mu})^2 + \frac{n}{n_1^2} \sum_{i=1}^{n_1} [\hat{H}(T_{1i}) - \hat{\hat{H}}]^2 + \frac{\hat{c}^2 \hat{r}^2 n}{n_2^2} \sum_{i=1}^{n_2} (T_{2i} - \hat{\mu}_2)^2,$$

where

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n T_i, \quad \hat{c} = \frac{1}{\hat{\mu}_2} \int_0^\infty y[1 - \hat{F}_2(y)] d\hat{F}_1(y) = \frac{1}{\hat{\mu}_2 n_1} \sum_{i=1}^{n_1} T_{1i} [1 - \hat{F}_2(T_{1i})],$$

$$\hat{H}(x) = x[\hat{c} + \hat{r} - 1 - \hat{F}_1(x/\hat{r}) - \hat{r}\hat{F}_2(x)] + \frac{\hat{r}}{n_1} \sum_{i=1}^{n_1} \mathbf{1}_{\{T_{1i} \leq x/\hat{r}\}} T_{1i} + \frac{\hat{r}}{n_2} \sum_{i=1}^{n_2} \mathbf{1}_{\{T_{2i} \leq x\}} T_{2i},$$

$$\hat{\hat{H}} = \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{H}(T_{1i}).$$

(hypothesis H_0^*).

Example 4. The test for the hypothesis H_0^* : Weibull null distribution. Times are simulated from Weibull distribution:

$$T_{1j} \sim W(\alpha_1, \beta_1), \quad T_{2j} \sim W(\alpha_2, \beta_2), \quad \alpha_1 = \alpha_2 = 2, \quad \beta_2 = 3\beta_1.$$

Under the hypothesis H_0^* the distribution function of a redundant system is

$$F(t) = 1 - e^{-\frac{t^2}{\beta_1^2}} - \int_0^t \frac{2y}{\beta_1^2} e^{-1/\beta_1^2(t^2 - \frac{4}{3}ty + \frac{13}{9}y^2)} dy.$$

The hypothesis H_0^* is tested using 5% asymptotic significance level under the different sample size n . Number of replications is 3,000. In Table 19.3 the values of the significance level for finite samples are given.

Table 19.3. Significance level

Sample size	Significance level (%)
50	4.8
100	5.2
200	5.0

Under the alternative hypothesis A_0^* (the c.d.f. of the standby unit has a jump of size $p(1 - F_2(y))$ at the switch-on moment y ,

Table 19.4. Power of the test

Sample size \ p	0.1	0.25	0.5	0.65	0.75
50	9	21	30	59	71
100	14	35	45	67	76
200	25	58	74	90	100

Table 19.5. Significance level

Sample size	Significance level (%)
50	5.6
100	5.2
200	5.1

Table 19.6. Power of the test

Sample size \ p	0.1	0.25	0.5	0.65	0.75
50	7	11	45	54	79
100	15	33	56	70	92
200	24	45	77	89	100

$$F_2^{(y)}(t) = 1 - e^{-\frac{t^2}{\beta_1^2}} - \int_0^t \frac{2y}{\beta_1^2} e^{-\frac{y^2}{\beta_1^2}} e^{\left(-\frac{1}{\beta_1^2} \left(t + \sqrt{-\beta_1^2 \log(1-p) + \frac{y^2}{9}} - y\right)^2\right)} dy.$$

In Table 19.4 power of the test under the alternative is given.

Example 5. The test for the hypothesis H_0 : Weibull null distribution (Tables 19.5, 19.6). Times are simulated from Weibull distribution:

$$T_{1j} \sim W(\alpha_1, \beta_1), \quad T_{2j} \sim W(\alpha_2, \beta_2), \quad \alpha_1 = 1, \quad \alpha_2 = 2, \quad \beta_2 = 3\beta_1.$$

Example 6. The test for the hypothesis H_0 : loglogistic null distribution (Table 19.7, 19.8). Failure times are simulated from loglogistic distribution:

$$T_{1j} \sim L(\alpha_1, \beta_1), \quad T_{2j} \sim L(\alpha_2, \beta_2), \quad \alpha_1 = \alpha_2 = 2, \quad \beta_2 = 3\beta_1.$$

The distribution function of a redundant system is

$$F(t) = 1 - \frac{\beta_1^2}{\beta_1^2 + t^2} - \int_0^t \frac{1}{1 + \left(\frac{t - \frac{2}{3}y}{\beta_1}\right)^2} \frac{2y}{\left(1 + \frac{y^2}{\beta_1^2}\right)^2 \beta_1^2} dy.$$

Table 19.7. Significance level

Sample size	Significance level (%)
50	8.1
100	6.3
200	5.2

Table 19.8. Power of the test

Sample size \ p	0.1	0.25	0.5	0.65	0.75
50	14	42	60	72	100
100	21	57	73	100	100
200	29	62	100	100	100

Acknowledgments. The authors are grateful to the University Victor Segalen and to “StatXpert” for the financial support.

References

1. Andersen, P.K., Borgan, O., Gill, R.D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer: New York.
2. Bagdonavičius, V. (1978). Testing the hypothesis of the additive accumulation of damages. *Probability Theory and Its Applications*, **23**, 403–408.
3. Bagdonavicius, V., Masiulaityte, I., Nikulin, M.S. (2008). Statistical analysis of redundant systems with one stand-by unit. In: *Mathematical Methods in Survival Analysis, Reliability and Quality of Life*, C. Huber, N. Limnios, M. Mesbah, and M.S. Nikulin (eds.), pp. 184–192. ISTE/WILEY, London.
4. Bagdonavicius, V., Masiulaityte, I., Nikulin, M.S. (2008). Statistical analysis of redundant systems with “warm” stand-by units. *Stochastics: An International Journal of Probability and Stochastic Processes*, **80**, 115–128.
5. Sedyakin, N.M. (1966). On one physical principle in reliability theory. (In Russian). *Technical Cybernetics*, **3**, 80–87.

An Approach to System Reliability Demonstration Based on Accelerated Test Results on Components

Léo Gerville-Réache and Vincent Couallier

Université de Bordeaux, UMR 5251, Bordeaux, France

Abstract: Reliability demonstration is often mandatory for industries that need to prove the quality of their production. In a reliability objective at the global system level, the basic use of standards such as MIL-HDBK-781 or IEC-60300 for reliability demonstration of each sub-system must be adjusted. Practitioners often have to fix high-level reliability goals (around 0.999) on each sub-system to guarantee that system reliability exceeds 0.8. Moreover, the confidence obtained for an overall reliability drops down to non-admissible values because of the multiplicity of testing procedures. We propose an approach, from k accelerated tests, to determine an optimized test time whilst ensuring that the reliability of a component system is greater than a goal value and by avoiding the too conservative method consisting in multiplying the confidence levels.

Keywords and phrases: Accelerated testing, demonstration of system reliability, risk of error, exponential and Weibull distributions

20.1 Introduction

A system is a collection of components arranged to a specific design in order to achieve desired functions. Each component (or sub-system) has its own reliability. The system reliability depends on the reliability of single components and of the structure of the system [9]. A system consisting of k independent components is a series system if the system's reliability is the product reliability of individual components:

$$R_s(t) = \prod_{i=1}^k R_i(t). \quad (20.1)$$

This equality reflects the fact that the system fails as soon as one component fails and that failure times T_1, \dots, T_k of components 1, \dots , k are independent random variables. The lifetime of the system is

$$T_s = \min(T_1, T_2, \dots, T_k).$$

A classical design for a reliability demonstration test (see [5] and [7] for international norms dealing with this issue or [1], Chapter 6 for a Bayesian treatment) is in the methodological framework of experimental design: it may consist in determining the number of units to be used, the test time to apply, in order to design a test plan related to a given demonstration objective. Obviously, this cannot be achieved without making any hypothesis on the underlying distribution since the method relies on the existence of explicit formulas for the unilateral confidence intervals of unknown parameters of the reliability.

For instance, under the assumption that a lifetime T follows an exponential distribution with unknown parameter λ , we often meet the following design ([7, 4]):

Objective: Define a test plan to demonstrate, with confidence γ , that $R(t_0) = P(T > t_0) > R_0$. Denoting by MTTF the Mean Time To Failure defined as the expectation of the lifetime, the objective is equivalent to $\mathbf{E}(T) = MTTF > MTTF_0 = t_0 / \ln(R_0)$ or equivalently to $t_{1-R_0} = \ln(1/(1-p))/\lambda > t_0$ where t_{1-R_0} is the $(1 - R_0)$ -quantile of the distribution for an exponential distribution.

Sampling pattern: Time censoring with replacement fixes the end time TT for observation, observe successively lifetimes T_1, T_2, \dots of units put into test, and replace failed units up to a fixed period of observation TT . By fixing c , the maximum allowable number of failures to be observed (c may be equal to zero) and A_F the acceleration factor of the test, the test time TT for the required objective is

$$TT = \frac{1}{2A_F\lambda_0} \chi_{2c+2}^2(\gamma) = MTTF_0 \frac{\chi_{2c+2}^2(\gamma)}{2A_F}.$$

Demonstration result: Let N be the observed number of failures; if $N \leq c$, then the reliability target is demonstrated with the confidence γ .

With a given confidence γ , the zero-failure demonstration ($c = 0, N_{\text{observed}} = 0$) leads to the minimum test time $TT = MTTF_0 \ln(1/(1-\gamma))/A_F$. This reliability demonstration method has been explored for a long time in the case of a single component ([2, p. 386], [6, 3, 8]). Nevertheless, even if the k sub-systems have successfully passed their demonstration test $\lambda_i < \lambda_i^0$, the targeted system reliability

$$\sum_{i=1}^k \lambda_i < \sum_{i=1}^k \lambda_i^0 = \lambda^0$$

is demonstrated but the global achieved confidence is just lower bounded by γ^k . This value can be very small when k is large: a demonstration at a 90% confidence of the reliability of 20 sub-systems shows the overall reliability at a (at least) 12% confidence.

To overcome this problem, we develop in this chapter an approach in order to determine the optimal test time to validate the global reliability of a series system with a given global confidence from accelerated demonstration tests performed on k

sub-systems. The reliability requirement is defined with respect to a quantile of the system reliability (20.1).

For each component test producing a lower confidence bound for an interval with level $\gamma = 1 - \alpha$, we must infer a lower confidence bound for the interval with level $\gamma = 1 - \alpha$ for the whole system. For ageing sub-systems, it is usual to model the lifetime with a Weibull distribution $W(\eta, \beta)$, where $\beta > 1$ stands for the ageing characteristic ($\beta = 1$ is a particular case allowing modelling of non-ageing component, namely an exponential distribution). See [3, 10, 1] for further details on the Weibull distribution. Using the fact that T^β has an exponential distribution with parameter $\lambda = \eta^{(-\beta)}$, the “zero-failure” demonstration test is available under the assumption that parameters β ’s are known for each sub-systems.

For simplicity we present in this chapter test times for demonstration testing without acceleration. For accelerated testing, the test times presented in the following have to be divided by the acceleration factors of each respective component accelerated test.

20.2 Global Reliability Demonstration from k “Zero-Failure” Component Testing Procedures

For each sub-system, a zero-failure demonstration test is designed and performed. Let TT_1, TT_2, \dots, TT_k be the corresponding test times leading to the demonstration of the reliability targets $t_{P_1}^0, t_{P_2}^0, \dots, t_{P_k}^0$, respectively, for parts $n^\circ 1, n^\circ 2, \dots, n^\circ k$. Assume that each demonstration test has confidence level $1 - \alpha$. If all tests have passed without failure, an equivalent global testing procedure on the system would have given zero systems failure during the time:

$$TT_{sys} = \min(TT_1, TT_2, \dots, TT_k).$$

In case of similar ageing of sub-systems $\beta_1 = \dots = \beta_k = \beta$, the system’s lifetime has a Weibull distribution $T \sim W((\sum_{i=1}^k \eta_i^{-\beta})^{-1/\beta}, \beta_{sys} = \beta)$. Moreover, any p -quantile t_{p-sys} for the system reliability distribution may be computed from the corresponding p -quantile t_{p-i} , $i = 1, \dots, k$ at the component level

$$t_{p-sys} = \left(\sum_{i=1}^k t_{p-i}^{-\beta} \right)^{-\frac{1}{\beta}}. \quad (20.2)$$

For ageing coefficients relatively close to each others and small p^0 (that means p^0 close to zero enough to make inference about the left tail of the distribution only), we can approximate the system reliability by a Weibull distribution with parameter $\beta_{sys} = (\beta_1 + \dots + \beta_k)/k$. By using the fact that $T^{\beta_{sys}}$ is exponentially distributed, a lower bound for the quantile t_{p-sys} may be derived: the test time TT_{sys} thus provides a reliability demonstration for the objective

$$t_{p-sys}^0 = TT_{sys} \left[\frac{\ln(1-p)}{\ln(\alpha)} \right]^{1/\beta_s}$$

with $1 - \alpha$ confidence.

20.2.1 Equal component test times

Let a system be made of k components, each component having a given failure mode following a Weibull distribution with known β_1, \dots, β_k . Assume that “zero-failure” components testing procedures have been passed. Assume also that the test times TT_i , $i = 1, \dots, k$ are equal to a given value TT . Obviously, different reliability targets are demonstrated for the sub-systems.

Suppose that the k tests were conducted successfully. This means that the system operated without failure for $TT_{sys} = TT$ hours. We therefore have

$$t_{p-sys}^0 = TT_{sys} \left[\frac{\ln(1-p)}{\ln(\alpha)} \right]^{1/\beta_s}.$$

As an example, let us take $k = 10$, $\alpha = 0.3$, $TT = 10,000$ hours and ageing parameters β 's given in Table 20.1. k -passed zero-failure demonstration tests have demonstrated that the quantile t_{p-sys} of the overall system is greater than $t_{p-sys}^0 = 3,100$ with a 70% confidence level.

Table 20.1. Ten component demonstration tests with confidence 70%. Equal test times $TT = 10,000$ hours provide different objective values for the $T10\%$ targets

No. of sub-system	Ageing coefficient (Beta)	Number of units/number of admissible failure	T10% at the confidence level 70% to demonstrate	Test time for demonstration
1	2	1/0	2,958	10,000
2	3	1/0	4,439	10,000
3	1.5	1/0	1,971	10,000
4	2	1/0	2,958	10,000
5	2	1/0	2,958	10,000
6	2.5	1/0	3,774	10,000
7	2.5	1/0	3,774	10,000
8	1.5	1/0	1,971	10,000
9	1.8	1/0	2,583	10,000
10	2	1/0	2,958	10,000

This value has to be compared to the one given by the well-known procedure used by the practitioners: suppose that each demonstration testing provides the same t_p for the k sub-systems at confidence $1 - \alpha$. Therefore, by using (20.2), the overall procedure demonstrates the reliability target $t_{p-sys} = k^{-1/\beta} t_p < t_p$ hours for the system at a confidence greater than $(1 - \alpha)^k = 2.8\%$. This is obviously fake and unusable, too conservative to the least.

20.2.2 Equal reliability targets

As system-level reliability demonstration raises the main interest of an integrator, sub-contractors are often asked to perform the unit tests at a component level. The vendor

is then challenged to perform a demonstration test of a reliability as required by the assembler. This customer often imposes an identical reliability goal, whichever the component.

Let us denote by t_p^0 the common reliability target which has been demonstrated at the component's level. According to the β values which may differ from component to component, the demonstration tests have required different test times $(TT_1, TT_2, \dots, TT_k)$.

Let us assume that the k tests were conducted successfully. The overall system has thus operated without failure during $TT_s = \min(TT_1, TT_2, \dots, TT_k)$. We, therefore, infer that the corresponding reliability target at confidence $1 - \alpha$ is

$$t_{p_s}^0 = TT_s \left[\frac{\ln(1-p)}{\ln(\alpha)} \right]^{1/\beta_s}.$$

Table 20.2 shows a numerical application for $k = 10$ sub-systems with component demonstration test times computed for the common component reliability target $T10\% = 5,000$ hours. In that case, the overall reliability target, demonstrated at confidence level 70%, is

$$t_{p_s}^0 = TT_s \left[\frac{\ln(1-p)}{\ln(\alpha)} \right]^{1/\beta_s} = 11,262 * \left[\frac{\ln(0,9)}{\ln(0,3)} \right]^{\frac{1}{2,08}} = 3,491 \text{ hours.}$$

Table 20.2. Ten component demonstration tests with confidence 70%. Equal reliability targets require different test times

No. of sub-system	Ageing coefficient (Beta)	Number units/number admissible failure	of T10% at the confidence level 70% to demonstrate	Test time for demonstration
1	2	1/0	5,000	16,902
2	3	1/0	5,000	11,262
3	1.5	1/0	5,000	25,367
4	2	1/0	5,000	16,902
5	2	1/0	5,000	16,902
6	2.5	1/0	5,000	13,248
7	2.5	1/0	5,000	13,248
8	1.5	1/0	5,000	25,367
9	1.8	1/0	5,000	19,351
10	2	1/0	5,000	16,902

20.2.3 Integrating failed demonstration procedures

The reliability demonstration tests may pass or fail according to a probability that obviously depends on the underlying distributions. The more the number of components (5, 10 or more), the higher the risk to have at least one failure. In that case, the overall

demonstration has failed. Nevertheless, some information has been provided by the test procedure: What can be said about the reliability of the overall system if at least one test has failed?

A failed “zero-failure” demonstration test occurs in case of a failure before the end test time. Statistically, we can compute this information to construct a lower bound confidence interval for the reliability target t_{P_i} . This lower bound corresponds also to a lower bound for an equivalent test without failure, where the test time would be smaller. We thus propose to compute the equivalent “zero-failure” test time leading to the same lower bound confidence interval for t_{P_i} .

Let $t_{P_i \text{ inf}}^0$ be the lower limit (at $1 - \alpha$) confidence interval of $t_{P_i}^0$ obtained after the failed reliability demonstration test on the i th unit. The time test (zero failure) leading to the same $t_{P_i \text{ inf}}^0$ is given by

$$T = \left[\frac{\ln(\alpha)}{\ln(1-p)} (t_{P_i \text{ inf}}^0)^{\beta_i} \right]^{1/\beta_i} = t_{P_i \text{ inf}}^0 \left[\frac{\ln(\alpha)}{\ln(1-p)} \right]^{1/\beta_i}.$$

Assume that in the previous example the 5th component failed at 8,000 hours and thus did not passed its “zero-failure” demonstration test. In that case, $t_{10\%i \text{ inf}}^0 = 1,537$ hours at 70% confidence level. Equivalent time is

$$T = 1,537 * \left[\frac{\ln(0.3)}{\ln(0.9)} \right]^{1/2} = 5,196 \text{ hours.}$$

The method consists in replacing the non-achieved end time of the test by the zero-failure equivalent time, namely 10,000 hours by 5,196 hours for the 5th component in Table 20.1. Therefore $TT_{sys} = \min(TT_1, TT_2, \dots, TT_k) = 5,196$ and

$$t_{p_s}^0 = 5,196 * \left[\frac{\ln(0.9)}{\ln(0.3)} \right]^{\frac{1}{2.08}} = 1,610 \text{ hours (against 3,100 hours).}$$

The “cost” of this failure is expensive! We must therefore plan a system demonstration test that takes into account of this risk.

20.3 Designing a Global Demonstration Test for the Reliability of a Series System Under a Success Probability Constraint

The objective is now to provide a demonstration test for each component that would guarantee at a system’s level that the tests will pass with a given probability. Also, for a fixed failure risk it is important to plan component tests that will maximize the $t_{p_s}^0$ value shown.

20.3.1 Basic principle of the demonstration test planning

By freezing a goal, i.e., a global confidence level α and a “non-demonstration” risk α_{N-D} for the system reliability demonstration, we aim at designing each component

“zero-failure” test with a given risk of non-demonstration, that is, the probability that all tests pass is $1 - \alpha_{N_D}$. We can deduce first the risk of non-demonstration component by component

$$\alpha_{iN_D} = 1 - (1 - \alpha_{N_D})^{1/k}.$$

Note that for any component, sub-system or system whose lifetime has a Weibull distribution, the probability α_{N_D} of failing a demonstration test depends on the value of the reliability objective t_p^0 , the confidence level α and the true value t_p of the target. Indeed

$$\alpha_{N_D} = 1 - \exp(\ln(\alpha)(t_p^0/t_p)^\beta)$$

for “zero-failure” demonstration testing (see Figure 20.1). We can also compute the required true target t_p to guarantee that the demonstration reliability testing has the probability $(1 - \alpha_{N_D})$ to be passed

$$t_p = t_p^0 \left(\frac{\ln(\alpha)}{\ln(1 - \alpha_{N_D})} \right)^{1/\beta} \text{ hours.}$$

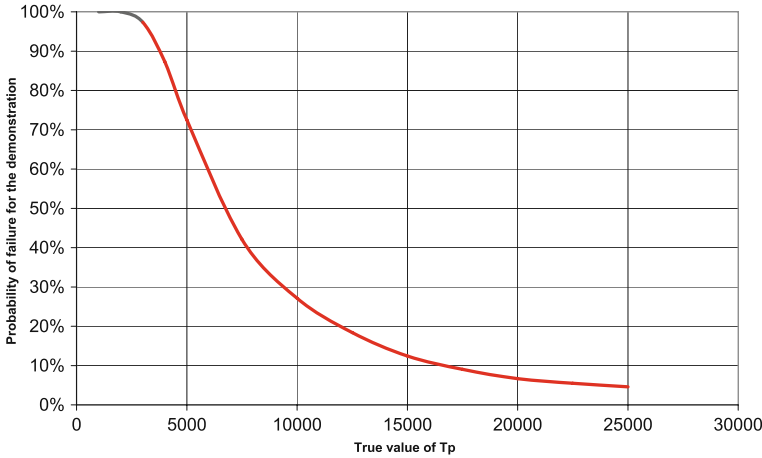


Figure 20.1. Probability of failure of a “zero-failure” demonstration with 70% confidence and $\beta = 2.08$

For instance, let us take a global reliability objective $t_{10\%_s}^0 = 5,000$ and a confidence level $1 - \alpha = 70\%$. Assuming that $\beta_{sys} = 2.08$, a zero-failure demonstration test will have probability 0.5 to succeed for $t_{10\%_s} = 6,520$ hours. The test time must be equal to

$$T = 5,000 * \left[\frac{\ln(0.3)}{\ln(0.9)} \right]^{1/2.08} = 16,129 \text{ hours}$$

to satisfy the test requirement. The risk of non-demonstration for each component must be $\alpha_{iN_D} = 1 - (1 - 0.5)^{1/10} = 0.067$ which allows us to compute for each component a lower bound for the true value of $t_{10\%_s}$ (see Table 20.3, 5th column).

Table 20.3 shows, for example, that the real t_p of the 5th sub-system must be greater than 20,000 hours to have a probability of non-demonstration less than 0.067.

Table 20.3. Component demonstration times and minimum true $t_{10\%}$ for guaranteeing a 50% global probability of success for 0/1 and 3/1 demonstration plans

No. of sub-system	Beta	$t_{10\%}$ at the confidence level 70% to demonstrate	Test time for demonstration trial 1/0	Minimum value of true $t_{10\%}$ trial 1/0	Test time for demonstration trial 3/1	Minimum value of true $t_{10\%}$ trial 3/1
1	2	4,771	16,129	20,000	14,791($\times 3$)	12,000
2	3	7,160	16,129	18,500	15,223($\times 3$)	13,000
3	1,5	3,179	16,129	21,500	14,370($\times 3$)	11,500
4	2	4,771	16,129	20,000	14,791($\times 3$)	12,000
5	2	4,771	16,129	20,000	14,791($\times 3$)	12,000
6	2,5	6,087	16,129	19,500	15,049($\times 3$)	12,500
7	2,5	6,087	16,129	19,500	15,049($\times 3$)	12,500
8	1,5	3,179	16,129	21,500	14,370($\times 3$)	11,500
9	1,8	4,167	16,129	21,000	14,649($\times 3$)	11,800
10	2	4,771	16,129	20,000	14,791($\times 3$)	12,000

This assumption for the real t_p can be unrealistic. In that case, we have to propose another design for this sub-system as for the others.

20.3.2 “ n/j failures” demonstration test

In order to pass the demonstration test, we can design a “ n/j failures” demonstration test (n and $j < n$ are values fixed by the user). In this test, for each sub-system, n units are tested without replacement. The demonstration test is successfully passed for a sub-system if the number of failures during the test does not exceed j .

The total test time T for a “ n/j failures” demonstration test is equal to $T = T' \times n$ where T' satisfies

$$\alpha = \sum_{i=0}^j \frac{n!}{i!(n-i)!} \left[1 - e^{-(T'/\eta)^\beta} \right]^i \left[e^{-(T'/\eta)^\beta} \right]^{n-i},$$

with

$$\eta = \frac{t_p^0}{[-\ln(1-p)]^{1/\beta}}.$$

Solving this binomial equation with respect to T' requires numerical computation or dedicated reliability software. The total test time T is always greater than the “zero-failure” test time for the same reliability objective and confidence level but the required real value of the target for guaranteeing a given probability of passing the test is reduced.

In order to illustrate this, we propose in Table 20.3 a “3/1 failures” demonstration test for each sub-system. In this case, the true t_p of the sub-system number 5 must be greater than 12,000 to have a probability of non-demonstration less than 0.067.

20.4 Conclusion

The reliability demonstration is featured by the estimation of a lower confidence bound. This limit is often obtained out from a “no defect” trial, with or without acceleration.

Demonstrating the reliability at the system level based on a bunch of unit reliability demonstrations requires an operating time-based approach for each subset.

The idea is then to translate each unit lower bound into an “equivalent zero-defect” time, resulting in a single system bound. A demonstration test failure is always due to one outage over. Hence it becomes necessary to program unit tests depending on a priori reliability values that are already known on sub-systems.

Table 20.3 shows such an example, where for instance subset #5 will operate 16,129 hours without failing to demonstrate a $t_{10\%}$ life of 4,771 hours. In order to reach 93.3% probability of success for such a trial, $t_{10\%}$ life must be higher than 20,000 hours, which can be considered unrealistic. In this case a good alternative can be to perform a 3/1 trial that will succeed with an equivalent probability, yet with $B_{10} = 12,000$ hours. This last trial design is undoubtedly more in line with the assessed a priori reliability made on system #5.

This chapter could not have been written if we had not had the opportunity to benefit from support and resource of the company StatXpert based in France. We kindly thank this company and its director for their cooperation and strong involvement in finding and sharing with us interesting and innovating industrial reliability issues.

This approach for a reliability demonstration of a system incorporates all the technical demonstration of an accelerated component-level test. The additional challenge is to design component tests under constraints given by an overall reliability goal for the system. The risk of non-demonstration and the notion of “equivalent time” are then at the core of the planning. Most importantly, the true reliability of each component must be taken as essential for predicting the probability of a successful demonstration.

References

1. Abernethy, R.B. (2004). *The New Weibull Handbook*. Abernethy, North Palm Beach, FL.
2. Coccozza-Thivent, C. (1997). *Processus stochastiques et fiabilité des systèmes*. Springer, Berlin.
3. Gnedenko, B.V. and Ushakov, I.A. (1995). *Probabilistic Reliability Engineering*. John Wiley & Sons, New York.
4. IEC 61124 (1997). Reliability testing – Compliance tests for constant failure rate and constant failure intensity, IEC Standard.
5. IEC 60300 (2001). Dependability management, Part 3-5: Reliability centered maintenance, IEC Standard.
6. Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*. John Wiley & Sons, New York.
7. MIL-HDBK-781 (1997). *Handbook for Reliability Test Methods, Plans, and Environments for Engineering, Development Qualification, and Production*. John Wiley & Sons, New York.

8. Pages, A. and Gondran, M. (1980). *Fiabilité des systèmes*. Collection de la Direction des Etudes et Recherches d'Electricité de France, 39. Editions Eyrolles. XXII, Paris.
9. Rausand, M. and Hoyland, A. (2004). *System Reliability Theory: Models and Statistics Methods*, Second Edition. John Wiley & Sons, New York.
10. Rinne, H. (2009). *The Weibull Distribution: A Handbook*. CRC Press, Boca Raton, FL. xxiv.

Survival Function Estimation

Robust Versus Nonparametric Approaches and Survival Data Analysis

Catherine Huber

Université Paris Descartes, Paris, France; U1018 INSERM, Villejuif, France

Abstract: W.Q. Meeker and L.A. Escobar famous book *Statistical Methods for Reliability Data* [26] is a very well-known reference for all engineers and researchers who are interested in reliability problems. W.Q. Meeker has a long experience in modeling and solving degradation problems with the most complex features, so that it is a pleasure for me to participate in this volume in his honor. For a long time, survival data analysis and reliability studies walked their way parallel without much interpenetration. But nowadays, impulsed by several people among which Bagdonavicius and Nikulin [3, 4], one is aware of the multiple links between reliability and survival analysis, acknowledging still for some specificities. Many parametric models are available, as well as large classes of models like accelerated models, mainly in use in reliability, and extended Cox model [11], that are the favorite for survival data which have parametric as well as semi-parametric versions. In the recent years, there was an increasing interest for purely nonparametric approaches. Their advantage is that they are supposed to be able to adjust any possible data set through a vast class of regular functions. The drawbacks are first that there is generally a lack of easy interpretation for practical purposes and second that proving the required properties of the inference procedures like consistency, for example, is not trivial when censoring and truncation are present (Huber et al. [18, 19]), which is often true in the medical field. The result is that the justification of some proposed procedures is only through some simulations without any theoretical proof of their properties. What I would like to emphasize in this chapter is the need for a robust approach to survival data analysis, which means using robust procedures for flexible parametric models, still valid when the observations follow a model close, but not exactly equal, to the assumed model [22]. This always seemed to me a good compromise between a purely parametric approach and a purely nonparametric one. One of the most interesting chapter on robustness in survival analysis is the one by Kosorok et al. [24], which tackles the case of a possibly misspecified frailty model.

Keywords and phrases: Robustness, Hellinger distance, influence curve, contamination, neighborhoods, Cox model, accelerated model, censoring, Kaplan-Meier, Prohorov distance

21.1 Introduction

The numerous and important works of W.Q. Meeker [25, 26] on accelerated and degradation models have long been and still are an extremely useful reference for engineers working on reliability problems and also for researchers working in this field. Actually, as long as applications are concerned, reliability and survival analysis rely mostly on two favorite types of models and their extensions: the accelerated models for the former and the Cox model for the latter. Those two general types of models have been extensively studied and shown to be linked in Bagdonavicius and Nikulin [3]. They have parametric and semi-parametric versions. However, since several years, the purely nonparametric approach seems to be considered the best by many authors, in spite of the fact that it usually does not give rise to easy interpretation as is the case when one holds on a parameter. In the medical field, the survival data experience both censoring and truncation Finkelstein et al. [12], H. Frydman [13] and Huber and Nikulin [20], which make it nontrivial to derive consistent estimators (Huber et al. [18, 19]). But, it seems to me that complementary to those diverse approaches, there is one that could be investigated more in detail: it is the robust approach. Sir David Cox himself, so famous (among many other reasons) for his proportional hazard (PH) model in use all over the world in medicine applications, said once to me that in his model, whose nice feature is that the basic hazard is free to be “any” hazard, this free basic hazard could be replaced by a sufficiently flexible parametric function without losing much in most situations. So this means that flexible parametric models could be most of the time sufficient. But then, the statistical inference should be robust as one can predict with probability one that the observed data will not follow exactly the model. This is true even when the model is correct for the phenomenon itself, because of errors and unavoidable rounding errors. In Section 21.2, we recall some basic examples of high instability of performances of optimal statistical procedures for parametric models. In Section 21.3, we compare the nonparametric and the robust approach and recall the definition of some measures of robustness. In Section 21.4, following [28] we give a measure of robustness in survival data for the Kaplan–Meier estimator. Finally, in Section 1.5 we conclude that in the field of survival analysis, there is still a need to construct robust estimators for flexible parametric models.

21.2 Motivation for Robustness

We recall in this section some famous examples of acute sensitivity of optimal parametric procedures to a slight discrepancy from the assumed model. We also show that the nonparametric approach may also suffer from this kind of discrepancy.

21.2.1 Instability of usual tests and estimators

Translation parameter

The oldest and most classical example of the sensitivity of optimal estimators to a slight perturbation of the assumed law is due to Tukey [31]. It was the root of the theory of robustness.

We know that in the translation model $F^n(\cdot - \theta)$, $\theta \in \mathbb{R}$, when F is the distribution function of the normal law on \mathbb{R} , $N(0, \sigma^2)$, \overline{X}_n is an efficient estimator of θ . But, if F is contaminated by a more dispersed law, let $F_\varepsilon = (1 - \varepsilon)N(\theta, \sigma^2) + \varepsilon N(\theta, 9\sigma^2)$ for a small value of ε smaller than 0.1, the efficiency of \overline{X}_n decreases very fast, even if this contamination is undetectable on the observed sample.

Actually, if the fraction of contamination is 1% and the sample size is 1000, there will be, in the mean, 10 observations arising from the contaminating law, but among those observations, only those that are outside the interval $[-2.5\sigma; +2.5\sigma]$, where the two laws are very similar, will alert about a possible contamination. Only 40% of those will lie outside: denoting F as both the probability and the corresponding distribution function when there is no possible confusion, we get

$$F_\varepsilon(|X| > 2.5\sigma) = F(|X| > (2.5/3)\sigma) \simeq 0.4.$$

We can thus expect about two observations to the right and two observations to the left of the interval $[-2.5\sigma; +2.5\sigma]$ coming from the contaminating law while there is also in the mean six observations to the right and 6 to the left from the main law. Thus, unless one or several of those four values are exceptionally high, it will not be possible to detect whether the normal model is contaminated or not. Nevertheless, the efficiency of \overline{X}_n decreases rapidly when the fraction of contamination increases from 0 to 0.10.

Contamination fraction	0.00	0.02	0.05	0.10
Empirical mean efficiency	1.00	0.90	0.80	0.70

Otherwise, if one removes a fraction 2α of the observations, the $[\alpha n]$ smallest and $[\alpha n]$ greatest, and takes the average of the remaining ones (where $[\alpha]$ means the greatest integer smaller than α), obtaining thus the so-called α -truncated mean, the resulting loss of efficiency for the strict normal model is very small. For example, for $\alpha = 0.01$, the loss of efficiency for the strictly normal model is 0.4%, while the gain for ε varying between 0 and 0.1 may go up to 12%.

Tukey obtained the following result: any optimal estimator for any value of $\varepsilon \in [0.01; 0.10]$ has an efficiency greater than 0.96 for any contamination between 0 and 0.10, so that any of those estimators is better than the usual mean which has an efficiency decreasing down to 0.7 on this same interval for ε .

Scale parameter

Let (X_1, X_2, \dots, X_n) be a sample of a real variable X whose distribution function is $F(\frac{\cdot - \theta}{\sigma})$, where θ and σ are unknown. We consider two unbiased estimators of σ based on the following measures of dispersion:

$$d_n = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

$$s_n = \left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{1/2}$$

There was a controversy between Eddington who was in favor of the first and Fisher who preferred the second. While the second is optimal for a normal law, the first is

more robust against a contamination as is shown by a calculus of the relative asymptotic efficiency of D_n with respect to S_n for the contaminated preceding model:

$$\begin{aligned} ARE_{D_n/S_n}(\varepsilon) &= \lim_{n \rightarrow +\infty} \frac{Var(S_n)/(E(S_n))^2}{Var(D_n)/(E(D_n))^2} \\ &= \frac{\frac{3(1+80\varepsilon)}{(1+8\varepsilon)^2} - 1}{4\left(\frac{\pi(1+8\varepsilon)^2}{2(1+2\varepsilon)^2} - 1\right)} \end{aligned}$$

Contamination fraction	0.00	0.001	0.002	0.01	0.05	0.50
ARE of D_n/S_n	0.876	0.948	1.016	1.016	1.439	1.017

One can see that a contamination of 0.002 is sufficient to ensure that the mean absolute deviation is more efficient than the root of the mean square deviation.

Comparing student and Wilcoxon tests

Let (X_1, \dots, X_{n_1}) be a sample of a random variable X with distribution function F_1 , (Y_1, \dots, Y_{n_2}) a sample of a random variable Y with distribution function F_2 , and $n = n_1 + n_2$. We want to test $H_0 : F_1 = F_2$ against $H_1 : F_2 = F_1(\cdot - \theta)$, $\theta > 0$. If we know that F_1 is a gaussian law, $N(\theta_0, \sigma^2)$, where θ_0 and σ are unknown parameters, Student test, based on T , is known to be uniformly most powerful:

$$T = \frac{(\bar{Y} - \bar{X})(n_1 + n_2 - 2)^{1/2}}{(\sum (Y_i - \bar{Y})^2 + \sum (X_i - \bar{X})^2)^{1/2} (1/n_1 + 1/n_2)^{1/2}},$$

while if we do not know F_1 , we can use Wilcoxon test, one of the simplest rank tests of the null hypothesis. H_0 is rejected if the sum of ranks W of the Y 's among the pooled samples ordered in increasing order exceeds some value:

$$W = \sum_{i=1}^{n_2} \left(\sum_{j=1}^{n_2} 1_{Y_i \geq Y_j} + \sum_{j=1}^{n_1} 1_{Y_i \geq X_j} \right).$$

Rank tests are based on a “free” (independent of F_1) statistic under H_0 , so that they have a constant level for all laws with continuous distribution function F_1 . The Wilcoxon test is locally the most powerful among the rank tests for the logistic distribution. Let $n = n_1 + n_2$ and $n_1/n \xrightarrow{n \rightarrow \infty} \lambda$ for some $\lambda \in (0; 1)$. One can show that the asymptotic relative efficiency of Wilcoxon test with respect to Student test at F_1 with density f_1 is equal to

$$e_{F_1}(W/T) = 12\sigma^2 \left(\int f_1^2(x) dx \right)^2.$$

When $F_1 = N(0, \sigma^2)$, this quantity is equal to $3/\pi = 0.955$, so that the loss of efficiency is small, which means that only 5% more observations are needed in order to get the same performance with Wilcoxon. This result, though it is asymptotic, holds

also for small values of n . As an example, for $n_1 = n_2 = 3$, $e = 0.979$. Moreover, its minimal value can be proved to be equal to 0.864 for the distribution F_0 having density $f_0(x) = \frac{3}{20\sqrt{5}}(5 - x^2)\mathbb{I}_{[-\sqrt{5}; +\sqrt{5}]}(x)$

$$\inf_{F_1} \{e_{F_1}(W/T)\} = 0.864.$$

Those considerations, very favorable to the nonparametric Wilcoxon test as compared to the parametric Student test (too much adapted to the strictly normal case), could let us think that rank tests are rather insensitive to perturbations of the initial assumed model, which is false as will be shown in the next paragraph.

21.2.2 Sensitivity of rank test

Rank test for the comparison of two samples is very stable as far as the null hypothesis $H_0 : F_1 = F_2$ is concerned since the level is independent of the value of F_1 .

Let ψ be a test based on the linear rank statistic S_n defined through the score function a_n :

$$S_n = \sum_{i=1}^{n_2} a_n(R(Y_i)) \equiv \sum_{i=1}^{n_2} a_n(R_i)$$

where $R_i = R(Y_i)$, the rank of Y_i when ordering all X 's and Y 's altogether by increasing order. Then reject H_0 , against $H_1 : F_2 = F(\cdot - \theta)$, $\theta > 0$, when $S_n > s_n$, so that $\psi_n = 1_{S_n > s_n}$. The score function ϕ that defines $a_n(R_i) = \phi(R_i/(n+1))$ is increasing and regular enough so that S_n is asymptotically normal. Let us now assume that there is some uncertainty about the law of the two samples so that hypotheses H_0 and H_1 become, respectively:

$$H'_0 : \mathcal{L}(X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}) \in H_n = \{\otimes_{i=1}^{n_1} G_i : \|G_i - F_1\|_\infty \leq \varepsilon\}$$

$$H'_1 : \mathcal{L}(X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}) \in K_n = \{\otimes_{i=1}^{n_1} G_i \otimes_{i=1}^{n_2} G'_i : \|G_i - F_1\|_\infty \leq \varepsilon,$$

$$\|G'_i - F_1(\cdot - \theta)\|_\infty \leq \varepsilon, \theta > 0\}.$$

Let, for fixed ε_0 and θ_0 , $\varepsilon = \varepsilon_0/\sqrt{n}$ and $\theta = \theta_0/\sqrt{n}$. Then, denoting U as the uniform variable on $[0; 1]$, $n_1 = \lambda n + o(n)$, and Φ as the distribution function of the standardized normal law, the asymptotic level and power are, respectively, equal to

$$\alpha_\infty = 1 - \liminf_n \Phi(\Phi^{-1}(1 - \alpha) + \frac{E_{F^n}(S_n) - \sup_{F_n \in H_n} E_{F_n}(S_n)}{[n\lambda(1-\lambda)\text{var}(\Phi(U))]^{1/2}})$$

$$\beta_\infty = 1 - \limsup_n \Phi(\Phi^{-1}(1 - \alpha) + \frac{E_{F^n}(S_n) - \inf_{F_n \in K_n} E_{F_n}(S_n)}{[n\lambda(1-\lambda)\text{var}(\Phi(U))]^{1/2}}).$$

As soon as function ϕ that defines the score is not bounded, level and power converge, respectively, to 1 and 0, as can be seen from choosing $G_i = F$, $i = 1, \dots, n_1$ and $G'_i = (1 - \varepsilon_0/\sqrt{n})F + (\varepsilon_0/\sqrt{n})1_{[a_n, +\infty)}$, $i = 1, \dots, n_2$, where $a_n = (1 - 1/n^2)$. It follows that rank tests based on unbounded score functions may experience instability. Most locally most powerful tests against parametric alternatives analog to H_1 , whose score functions are given by $\phi = g'/g \circ G^{-1}$, are in this situation as can be seen, for example, for the van der Waerden test locally most powerful in the gaussian case. Other cases of perturbation have been investigated like approximate independence instead of strict independence in assumptions H_0 and H_1 [2].

21.3 Robustness Concepts

21.3.1 Robust versus nonparametric approach

In the preceding section, we saw that optimality of estimators and tests for a specified parametric model is seriously damaged when the observations suffer from even very small perturbations of the model. Actually, it is always true that observed samples cannot fit exactly a given parametric model, even if the model is true for the phenomenon under study, due to errors in measurement and more intrinsically because rounding effects are unavoidable. Huber [16] and Hampel [15] were among the first, after Tukey, who took care of this problem systematically. But, time passing by, one could observe a certain drift from the initial objective, leading to nonparametric and adaptive methods. While, on the opposite side, many people focus on the study of outliers, assuming that, except a very small number of them (sometimes reduced to one as in Balakrishnan [5]), called outliers, the observations follow exactly the assumed model. Another direction which is the framework common to all those studies is the following:

- A parametric model fulfilling regularity conditions, given through a theory about a random phenomenon. This model is the kernel of the statistical model that will be used.
- A statistical model including the parametric model. In general, this will be a neighborhood of the parametric model for a distance d on the set of all probabilities; d is chosen in order to take into account the expected deviations of the law of the observations with respect to the law of the phenomenon itself.
- A criterion of the minimax type, generally asymptotic, to choose an optimal test or estimator for the parameters of the underlying parametric model.

We could sketch the difference between the initial robustness studies by Hampel [14, 15], Huber [22], Jaeckel [23] and the subsequent studies by Stone [30], Beran [8, 9], Millar [27], and Bickel [10] as follows: Let $\hat{\theta}_n$ be a function of the sample (X_1, \dots, X_n) estimating the parameter θ , $\theta \in \mathbb{R}^p$. We have now four generally distinct probabilities:

- P_θ , the law of X the variable under study.
- P , the law of the corresponding actually observed variable. This means that the law of the observations $\mathcal{X} = (X_1, \dots, X_n)$ is close to P_θ^n .
- P_n , the empirical law of \mathcal{X} : $P_n = \frac{1}{n} \sum \delta_{X_i}$.
- $P_{\hat{\theta}_n}$ that may be considered as an estimator of P_θ .

The first point of view says that we are not interested in P . Only the parameter θ of the model is the aim to be reached, in spite of the deviations of P with respect to P_θ^n . The estimator will then have to meet the following three requirements:

- We should keep a good efficiency, as close as possible to the optimal one in case $P \equiv P_\theta^n$. This loss in efficiency is the price we accept to pay in order to fulfill the two other requirements.
- Its performances should not vary much in terms of mean quadratic error for example when P lies in a specified neighborhood of P_θ^n .
- No disaster occurs when going further from P_θ^n .

The second point of view leaves behind the multiplicity of those antagonistic objectives in order to recover a classical statistical decision framework. We are thus led to consider that P is interesting in itself. Let us, for example, consider [30] who obtains an estimator of the center of symmetry of a symmetric law on \mathbb{R} that is asymptotically efficient for any symmetrical law on \mathbb{R} . Originally, there was the estimation of the mean of a normal law $N(\theta, \sigma^2)$ on \mathbb{R} , whose optimal estimator is the very unstable empirical mean \bar{X}_n . But θ is also the median of this normal law and it is also its symmetry center. If, for all symmetrical law Q , the functional $T(Q)$ is defined as the symmetry center of Q , Stone gives an optimal solution to the following problem: estimate $T(Q)$, $Q \in \mathcal{S}$. This is a nonparametric problem covering the robustness problem: estimate

$$T(Q) : Q \in \{(1 - \varepsilon)N(\theta, \sigma^2) + \varepsilon R(\cdot - \theta), \sigma \in \mathbb{R}^+, R \in \mathcal{S}_0, \varepsilon \in [0; \infty],$$

where \mathcal{S}_0 is the set of laws symmetrical with respect to 0.

But as we aim at estimating $T(P)$ in an optimal way, $T(P)$ should be equal to $T(P_\theta)$, which *allows only symmetrical contaminations*, while it is unrealistic to assume that all expected errors will be symmetrical.

The second example deals with the closest neighborhood approach: $P_{\hat{\theta}_n}$ is defined as the closest neighbor of P_n in the parametric model for the distance which defines the deviations with respect to the model. The risk being defined through this same distance between P and $P_{\hat{\theta}_n}$, the choice of the distance plays a central role.

21.3.2 Regularity of the parametric model

Let H be a family of distributions on a metric separable complete space E with its Borel sigma-field B , mainly \mathbb{R}^k with the euclidian distance, indexed by the points of an open convex set Θ of an euclidian space, mainly \mathbb{R}^p :

$$H = \{P_\theta : \theta \in \Theta \subset \mathbb{R}^p\}$$

assumed regular, which means well parametrized and such that the maximum likelihood estimator is optimal:

- Θ is an open convex subset of \mathbb{R}^p , $\theta \longrightarrow P_\theta$, $\theta \in \Theta$ is a one-to-one mapping and P_θ is absolutely continuous with respect to a sigma-finite measure μ , or else in some cases, mutually absolutely continuous. We denote $f_\theta = dP_\theta/d\mu$ and define the Hellinger distance h of two probabilities P and Q and assume that $|t|$ being the euclidian norm:

$$\begin{aligned} h^2(P, Q) &= \frac{1}{2} \int (\sqrt{dP} - \sqrt{dQ})^2 \\ h^2(P_{\theta+t}, P_\theta) &= o(|t|^2). \end{aligned}$$

- For every $\theta \in \Theta$, there exists a function $\xi(\theta) \in L^2(P_\theta)$ such that

$$\|\sqrt{f_{\theta+t}} - \sqrt{f_\theta} - t, \xi_\theta\|_2^2 = o(|t|^2)$$

and the matrix information $I(\theta) = 4E_\mu(\xi_\theta \xi'_\theta)$ exists and is non-singular.

It is convenient to express the regularity conditions on the underlying parametric model in terms of Hellinger distance for several reasons. First, the Hellinger distance is easy to deal with when considering tensorial power of measures:

$$\begin{aligned} h^2(P, Q) &= 1 - \int \sqrt{dP} \sqrt{dQ} \\ h^2(P^{\otimes n}, Q^{\otimes n}) &= 1 - (1 - h^2(P, Q))^n = nh^2(P, Q) + o(h^2(P, Q)). \end{aligned}$$

Second, it gives an easy lower bound for the sum of first and second kinds of errors of a test of P against Q for any test ϕ :

$$\begin{aligned} E_P(\phi) + E_Q(1 - \phi) &\geq \int dP \wedge dQ \geq \frac{1}{2} (\int \sqrt{dP} \sqrt{dQ})^2 \\ &= \frac{1}{2} (1 - h^2(P, Q))^2 \geq \frac{1}{2} \exp(-4h^2(P, Q)) \end{aligned}$$

as long as $h^2(P, Q)$ is smaller than $1/2$. And this problem of testing two simple hypotheses is fundamental both for general test problems and for estimation problems.

21.3.3 Extension of the underlying model

The expected deviations with respect to the central model may be modeled by a contamination model as in the preceding sections:

$$V(H) = \{P : \exists \theta : P = (1 - \varepsilon)P_\theta + \varepsilon Q, \quad Q \in \mathcal{P}'\}.$$

where \mathcal{P}' is a subset of \mathcal{P} , or by a neighborhood based on a specified distance on the probability laws.

$$V(H) = \{P : \exists \theta : d(P, P_\theta) \leq \varepsilon\}$$

Let P and Q be two probabilities on $(\mathbb{R}, \mathcal{B})$ with respective distribution functions F and G . Possible distances on \mathcal{P} , the set of all probabilities on $(\mathbb{R}, \mathcal{B})$, to define $V(H)$ are the following:

$d_K(P, Q) = \sup_t F(t) - G(t) $	Kolmogorov distance
$d_v(P, Q) = \sup_{B \in \mathcal{B}} P(B) - Q(B) $	Total variation distance
$d_L(P, Q) = \inf(\varepsilon : G(t) \leq F(t + \varepsilon)) \quad t \in \mathbb{R},$	Levy distance
$\pi(P, Q) = \inf(\varepsilon : Q(B) \leq P(B^\varepsilon) + \varepsilon; \quad B \in \mathcal{B})$	Prohorov distance
$h(P, Q) = (\frac{1}{2} \int (\sqrt{dP} - \sqrt{dQ})^2)^{1/2}$	Hellinger distance
$d_\nu(F, G)^2 = \int (F - G)^2 d\nu; \quad \nu \text{ sigma-finite}$	$L^2(\nu)$ distance
$d_{LB}(P, Q) = \sup_{\psi \in \mathcal{L}} \int \psi dP - \int \psi dQ $	Lipschitz Bounded distance,

where \mathcal{L} is a set of Lipschitz functions: $|\psi(x) - \psi(y)| \leq d(x, y)$, or, finally, by a subset of \mathcal{P} that contains the initial model and is specified through a condition, like, for example, the subset of symmetrical laws, or the subset of laws verifying a specified Sobolev condition. Among the possible distances mentioned above, the most adequate to take into account both gross errors and rounding errors is undoubtedly the Prohorov distance, recommended by F.R. Hampel. Levy's distance has the same property but it is limited to laws on \mathbb{R} while Prohorov distance is more general.

21.3.4 Measures of robustness

The definition of robustness hereafter as well as the two quantitative measures of robustness that follow are due to F.R. Hampel [14, 15]. Roughly speaking, a sequence of statistics T_n is robust at the parametric model H if it is equicontinuous in law, for the

weak convergence of measures, at every point in H . More precisely, if $T_n, n \in \mathbb{N}$ is a sequence of statistics, functions of the empirical law $P_n = (1/n) \sum \delta_{X_i}$, a one-to-one functional of the observed sample $\mathbf{X} := (X_1, \dots, X_n)$ defined up to a permutation, we have the following definition:

Definition 1 (robustness). *The sequence $(T_n), n \in \mathbb{N}$, is robust at P if for any $\varepsilon > 0$, there exists a $\delta > 0$ such that for any Q_n such that $\pi(P^n, Q_n) \leq \delta$, then $\pi(\mathcal{L}(T(\mathbf{X})|P^n), \mathcal{L}(T(\mathbf{X})|Q_n) \leq \varepsilon$.*

Very often, the sequence T_n is the restriction of a functional T , defined on a subset of the set of all probabilities \mathcal{M} on (E, \mathcal{B}) , to the set \mathcal{M}_n of the discrete probabilities whose atoms are multiples of $1/n$, with values in \mathbb{R}^p with the euclidian distance d . Under those conditions, the robustness of T_n at point P is equivalent to the continuity of T at point P for the weak convergence of measures.

Then, for a robust functional T , there are two quantitative measures of robustness. The influence curve, $IC_{T|F}(x)$, is a measure of the influence of an outlier at point x on the functional T of F :

$$IC_{T|F}(x) = \lim_{t \downarrow 0} \frac{T((1-t)F + t\delta_x) - T(F)}{t}. \quad (21.1)$$

The influence curve can be used to derive an approximation of the difference $T(G) - T(F)$:

$$T(G) - T(F) = \int IC_{T|F}(u) dG(u). \quad (21.2)$$

Sensitivity of T at F , $\sigma(T|F)$, is defined as the maximum absolute value of the influence curve, while breakdown point measures how far from the assumed model one can go without breaking the properties of T :

Definition 2. *If $\varepsilon_{T|F}^*(\delta) = \inf\{\varepsilon : \pi(F, G) \leq \delta \Rightarrow |T(F) - T(G)| \leq \varepsilon\}$ is the maximal asymptotic bias, the breakdown point of T at point F is defined as*

$$\delta^*(T|F) = \sup\{\delta : \varepsilon_{T|F}^*(\delta) < \infty\}. \quad (21.3)$$

21.4 Robustness in Survival Analysis

21.4.1 Measure of robustness of Kaplan–Meier estimator

Framework: the right censored case

We are interested in estimating the distribution function F of a positive random variable X , which is the time to onset of some event called the “survival time.” But, due to right censoring, X is observed only if it is smaller or equal to some random variable C with distribution function G . Otherwise, C is observed instead of X . C is called the right censoring variable. Finally, the data is the pair T, D where $T = \min(X, C) := X \wedge C$ is the observed time and $D = \mathbb{1}\{X \leq C\}$ is the censoring indicator. We have a sample of n (T_i, D_i) . Subjects for which $D = 1$ are not censored. The distribution functions F and G are assumed to have no common jump. Associated to F , generally, but not

always, assumed to be continuous, are the following five functions. The density f , the hazard rate λ , the cumulated hazard Λ , and the survival function $S = 1 - F$:

$$\begin{aligned}\lambda(t) &= \frac{f(t)}{S(t)} \\ \Lambda(t) &= \int_0^t \lambda(u) du \\ S(t) &= \exp -\{\Lambda(t)\}.\end{aligned}$$

Any of the five functions F , S , f , λ , or Λ identify the distribution of X . Instead of considering that the data is the pairs $(T_i, D_i, i = 1, \dots, n)$, one may use the point process (Andersen et al., [1]) representation of the data, namely the point process $N(t)$ and the at-risk process $Y(t)$

$$\begin{aligned}N(t) &= \sum_{i=1}^n \mathbb{1}\{T_i \leq t, D_i = 1\} := \sum_{i=1}^n N_i(t) \\ Y(t) &= \sum_{i=1}^n \mathbb{1}\{T_i \geq t\} := \sum_{i=1}^n Y_i(t).\end{aligned}$$

Considering now the case $n = 1$, to avoid the subscript i for the two processes N and Y relative to a unique subject i , we can write:

$$\begin{aligned}P(dN(t) = 1 | N(t^-) = 0) &= \lambda(t)dt \\ P(dN(t) = 1 | N(t^-) = 1) &= 0 \\ &\text{otherwise written} \\ P(dN(t) = 1 | N(t^-)) &= \lambda(t)dt \text{ with probability } S(t) \\ &= 0 \text{ with probability } 1 - S(t).\end{aligned}$$

The intensity $i(t)$ and cumulative intensity $I(t)$ of process $N(t)$ at time t are random and equal to

$$\begin{aligned}i(t) &= Y(t)\lambda(t) \\ I(t) &= \int_0^t Y(u)\lambda(u)du = \Lambda(t \wedge T)\end{aligned}$$

and the difference $M = N(t) - I(t)$ is a martingale, which is the main feature provided in the point process approach of survival data.

Peterson representation of Kaplan–Meier estimator

The Kaplan–Meier estimator \hat{S}_{KM} of the survival function S is a generalization to right censored data of the usual empirical survival function $1 - F_n(t) = \sum_{i=1}^n (\mathbb{1}\{X_i \geq t\})$:

$$\hat{S}_{KM}(t) = \prod_{T_{(j)} \leq t} \left(1 - \frac{dN(T_{(j)})}{Y(T_{(j)})}\right),$$

where $T_{(j)}$ are the order statistics of the sample T_1, \dots, T_n . This nonparametric estimator has several remarkable properties. First, it is the unique coherent estimator of S . A coherent estimator \tilde{S} is defined as having the following property:

$$\tilde{S}(t) = \frac{1}{n} \left[\sum_{i=1}^n \mathbb{1}\{T_i \geq t\} + \sum_{i=1}^n \mathbb{1}\{T_i < t, D_i = 0\} \frac{\tilde{S}(t)}{\tilde{S}(T_i)} \right].$$

Second, it suffices that there is no common discontinuity of F and G for \hat{S}_{KM} to be a consistent estimator of S . A way of proving most properties of \hat{S}_{KM} is to express it by Peterson formula hereafter. Let H be the distribution function of T and S_0 and S_1 the following subdistribution functions:

$$\begin{aligned}
S_0 &= P(T \geq t, D = 0), \\
S_1 &= P(T \geq t, D = 1), \\
S_2 &= S_0 + S_1 \text{ the survival function of the observed time } T.
\end{aligned}$$

Then $1 - H = S_0 + S_1$.

Lemma 1 (Peterson functional). *Let us assume that F and G , the distributions of the survival X and the right censoring variable C , have no common discontinuities. Then the survival function S of X may be expressed as*

$$S(t) = \exp \left\{ \int_0^t \frac{dS_1}{S_0 + S_1} + \sum_{u < t, S_0(u^+) = S_0(u^-)} \log \left[\frac{S_0(u^+) + S_1(u^+)}{S_0(u^-) + S_1(u^-)} \right] \right\}. \quad (21.4)$$

The first term is integration on continuity intervals of S_1 while the second one is a sum on discontinuity points of S_1 . Let us define W the corresponding functional of S_0 and S_1 :

$$S(t) = W(S_0, S_1, t). \quad (21.5)$$

From Equation (21.4), one can derive a Peterson expression for the cumulated hazard Λ :

$$T(S_0, S_1, t) = \int_0^t \frac{dS_1}{S_0 + S_1} + \sum_{0 \leq u \leq t, S_0(u^+) = S_0(u^-)} \log \left\{ \frac{S_0(u^+) + S_1(u^+)}{S_0(u^-) + S_1(u^-)} \right\}. \quad (21.6)$$

The empirical subdistributions corresponding to S_0 and S_1 are denoted as, respectively, \hat{S}_0 and \hat{S}_1 :

$$\begin{aligned}
\hat{S}_0(t) &= \sum \frac{1}{n} (\mathbb{1}\{T_i \geq t, D_i = 0\}) \\
\hat{S}_1(t) &= \sum \frac{1}{n} (\mathbb{1}\{T_i \geq t, D_i = 1\}).
\end{aligned}$$

Then the following lemma holds.

Lemma 2.

$$\hat{S}_{KM} = W(\hat{S}_0, \hat{S}_1, t). \quad (21.7)$$

Peterson expression of Kaplan–Meier estimator is very useful. For example, in order to prove the almost sure convergence of \hat{S}_{KM} toward S , it is sufficient to remark that the empirical \hat{S}_0 and \hat{S}_1 converge uniformly in t almost surely to S_0 and S_1 , respectively, by Glivenko Cantelli theorem, and the functional W is a continuous functional of its arguments for the sup norm. The asymptotic normality as well results from an extension of Donsker's theorem:

Theorem 1. *If X and C are independent and have no common discontinuity, then*

$$\sqrt{n}(\hat{S}_{KM}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} Z, \quad (21.8)$$

where Z is a gaussian process with covariance function:

$$\text{cov}(Z(t_1), Z(t_2)) = S(t_1)S(t_2) \int_0^{t_1 \wedge t_2} \frac{dF(u)}{[1 - F(u)]^2 [1 - G(u)]} \quad (21.9)$$

Corollary 1.

$$\hat{S}_{KM}(t) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(S(t), \frac{S^2(t)}{n} \int_0^t \frac{dS_1(u)}{(1 - (S_2(t))^2)}), \quad (21.10)$$

with a consistent estimator $\hat{V}(t)$ of the variance $V(t)$ of $\hat{S}_{KM}(t)$:

$$\hat{V}(t) = \hat{S}^2(t) \int_0^t \frac{dN(u)}{(Y(u^-))(Y(u^-) - dN(u))}. \quad (21.11)$$

Measure of robustness of the Kaplan–Meier estimator

Definition 3 (Fréchet derivative). \mathcal{B} and \mathcal{C} being two normed vectorial spaces, let T be a functional from \mathcal{B} to \mathcal{C} , F and H be elements of \mathcal{B} , dT_F a linear functional from \mathcal{B} to \mathcal{C} , and R defined as the difference

$$R(F + H) = T(F + H) - T(F) - dT_F \circ H.$$

Then T is Fréchet differentiable at F if and only if

$$\frac{\|R(F + H)\|_{\mathcal{C}}}{\|H\|_{\mathcal{B}}} \xrightarrow[\|H\|_{\mathcal{B}} \rightarrow 0]{} 0. \quad (21.12)$$

From the preceding paragraph, the Kaplan–Meier estimator is a functional W (21.7), and the cumulated hazard a functional T (21.6) of two arguments, \hat{S}_0 and \hat{S}_1 , which are empirical subdistribution functions. Thus we need to consider two Frechet derivatives, one for each of the two arguments.

Here, \mathcal{B} is the space of subsurvival functions whose norm is the sup norm $\|H\|_{\mathcal{B}} = \|H\|_{\infty}$, and $\hat{S}(t) = \exp \{T(\hat{S}_0, \hat{S}_1, t)\}$ is a functional of two arguments both in \mathcal{B} .

Theorem 2. *The functional $T(S_0, S_1, t)$ is Fréchet differentiable at every point of $\mathcal{B} \otimes \mathcal{B}$ and at every point t such that $t < \infty$, $S_0(t) > 0$, and $S_1(t) > 0$.*

Let $T(G_1, G_2)$ be a bivariate functional. The von Mises development of T around the point (F_1, F_2) can be written as

$$\begin{aligned} T(G_1, G_2) &= T(F_1, F_2) + \int IC_1(T|F_1, F_2)(y) d(G_1 - F_1)(y) \\ &\quad + \int IC_2(T|F_1, F_2)(y) d(G_2 - F_2)(y) + o(\|G_1 - F_1\| + \|G_2 - F_2\|). \end{aligned}$$

The two influence curves IC_1 and IC_2 are defined by

$$\begin{aligned} \frac{\partial}{\partial \varepsilon} [T(F_1 + \varepsilon(G_1 - F_1), F_2 + \delta(G_2 - F_2))]_{(\varepsilon=0, \delta=0)} &= \int IC_1(T|F_1, F_2)(y) d(G_1 - F_1)(y) \\ \frac{\partial}{\partial \delta} [T(F_1 + \varepsilon(G_1 - F_1), F_2 + \delta(G_2 - F_2))]_{(\varepsilon=0, \delta=0)} &= \int IC_2(T|F_1, F_2)(y) d(G_2 - F_2)(y) \end{aligned}$$

Here, T is a function, thus IC_1 and IC_2 have as arguments S_0 the “sub censoring” and S_1 the “sub survival”, t the current point, and s the additional point.

The two influence curves of Kaplan–Meier

The Kaplan–Meier estimator is a step function that jumps only at points where a “death” occurred. The additional observation may be a death or a censoring. In the first case, we have an additional jump and a modification of the other jumps, while in the second case, only the second effect takes place. The influence curves are first computed on the estimated cumulated hazard rate \hat{A} , using its Peterson representation (21.6). S_0 and S_1 are left continuous nonincreasing functions from \mathbb{R}^+ to $[0; \alpha]$ for some $0 < \alpha \leq 1$. Then

$$\begin{aligned}\hat{A}(t) &= -\log(\hat{S}_{KM}(t)) = -T(\hat{S}_0, \hat{S}_1, t) \\ &= -\sum_{0 \leq u \leq t, \hat{S}_0(u^+) = \hat{S}_0(u^-)} \log\left\{ \frac{\hat{S}_0(u^+) + \hat{S}_1(u^+)}{\hat{S}_0(u^-) + \hat{S}_1(u^-)} \right\}.\end{aligned}$$

As \hat{S}_0 and \hat{S}_1 are step functions, the first term in (21.6) is canceled. Let us compute the influence curve of T at point (S_0, S_1) :

$$\begin{aligned}T(S_0 + \delta(\hat{S}_0 - S_0), S_1 + \varepsilon(\hat{S}_1 - S_1), t) &= \int_0^t \frac{d(1-\varepsilon)S_1(s)}{[S_0 + \delta(\hat{S}_0 - S_0) + S_1 + \varepsilon(\hat{S}_1 - S_1)](s)} \\ &\quad + \sum_{0 \leq s < t} \frac{[S_0 + \delta(\hat{S}_0 - S_0) + S_1 + \varepsilon(\hat{S}_1 - S_1)](s^+)}{[S_0 + \delta(\hat{S}_0 - S_0) + S_1 + \varepsilon(\hat{S}_1 - S_1)](s^-)}.\end{aligned}$$

In the second term the summation takes place on the discontinuity points of \hat{S}_1 . Thus

$$\begin{aligned}\frac{\partial}{\partial \varepsilon} [T]_{\varepsilon=0, \delta=0} &= -\int_0^t \frac{dS_1(s)}{[S_0 + S_1](s)} - \int_0^t \frac{[\hat{S}_1 - S_1](s) dS_1(s)}{[(S_0 + S_1)(s)]^2} \\ &\quad + \sum_{0 \leq s < t} \frac{\hat{S}_1(s^+) - S_1(s^-)}{(S_1 + S_0)(s)} \\ \frac{\partial}{\partial \delta} [T]_{\varepsilon=0, \delta=0} &= \int_0^t \frac{[\hat{S}_0 - S_0](s) dS_1(s)}{[(S_0 + S_1)(s)]^2}.\end{aligned}$$

Let us notice first that for the second derivative, there is no jump term as the jump terms are associated with the jumps of \hat{S}_1 and second that

$$[\hat{S}_1 - S_1](s) = -\int_s^\infty d(\hat{S}_1 - S_1)(u).$$

Then we get

$$\begin{aligned}\int_0^t \frac{[\hat{S}_1 - S_1](s) dS_1(s)}{[(S_0 + S_1)(s)]^2} &= -\int_0^\infty \int_0^{s \wedge t} \frac{dS_1(u)}{[(S_0 + S_1)(u)]^2} d(\hat{S}_1 - S_1)(u) \\ \sum_{0 \leq s < t} \frac{\hat{S}_1(s^+) - S_1(s^-)}{[(S_1 + S_0)(s)]^2} &= \int_0^t \frac{d(\hat{S}_1 - S_1)(s)}{(S_0 + S_1)(s)} + \int_0^t \frac{dS_1(s)}{(S_0 + S_1)(s)},\end{aligned}$$

which give the two influence curves of functional T :

$$\begin{aligned}IC_1(T|S_0, S_1, t)(s) &= \int_0^{s \wedge t} \frac{dS_1(u)}{[(S_0 + S_1)(u)]^2} + \frac{\mathbb{1}\{s < t\}}{(S_0 + S_1)(s)} \\ IC_2(T|S_0, S_1, t)(s) &= \int_0^{s \wedge t} \frac{dS_1(u)}{[(S_0 + S_1)(u)]^2}.\end{aligned}$$

In the very usual special case of exponential survival with parameter 1 and independent uniform censoring on $[0; a]$, the calculus of the two IC, is explicit and gives

$$\begin{aligned}
IC_1(T|S_0, S_1, t)(s) &= e^{-t} \left[\frac{a}{a-s} e^s \cdot \mathbb{1}\{s \leq t\} - \int_0^{t \wedge s} \frac{ae^{-u}}{a-u} du \right] \\
IC_2(T|S_0, S_1, t)(s) &= -e^{-t} \int_0^{t \wedge s} \frac{ae^{-u}}{a-u} du,
\end{aligned}$$

where one can see that the two are identical for $s > t$ and that, for $s < t$, the influence is increasing in absolute value with increasing s and is negative for IC_2 (additional censoring at s) and positive for IC_1 (additional death at s).

21.4.2 Perspectives

Influence curves measure the degree of robustness of a statistic, meant to be an estimator or a test, against the presence of outliers. In order to derive an optimal robust procedure, one has to choose a compromise between the loss of efficiency at the strict parametric model and the protection (influence curve and breakdown point) against deviations. In the case of right censoring, Basu et al. [7] developed a robust estimation, adapting the robust density power divergence methodology (Basu et al. [6]). It seems, however, that there is still work to be done for the general case of interval censoring and truncation. Flexible parametric models can be found in the works of Meeker [26] and Bagdonavicius and Nikulin [3] involving generalized Weibull and Gamma distributions that could replace the nonparametric approach that we consider in Huber et al. [18] for interval censored and truncated data first studied by Turnbull [32]. Robust procedures adapted to these models should be developed. In order to do so, one has to use minimax optimality properties and search for a least favorable submodel inside a neighborhood of the strict flexible parametric model, for which the optimal procedure has minimax properties for the whole neighborhood. Devices to do so are available, for example, in Slud and Vonta [29] and Huber and Strassen [16].

References

1. Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N.: Statistical Models based on Counting Processes. Springer Series in Statistics (1993)
2. Arra, J.: Tests robustes pour le problème des deux échantillons. Thèse, Paris 11, Orsay (1979).
3. Bagdonavicius, V. and Nikulin, M.S.: Accelerated Life Models. Modeling and Statistical Analysis. Monographs on Statistics and Applied Probability. Chapman and Hall (2002)
4. Bagdonavicius, V. and Nikulin, M.S.: Statistical analysis of survival and reliability data with multiple crossings of survival functions. Comptes Rendus de l'Académie des Sciences de Paris, Ser. I, **340**, 377–382 (2005)
5. Balakrishnan, N.: Permanents. Order statistics, outliers, and robustness. Rev. Mat. Complut., **20**(1), 7–107 (2007)
6. Basu, A., Harris, I.R., Hjort, N.L. and Jones, M.C.: Robust and efficient estimation by minimising a density power divergence. Biometrika, **85**(3), 549–559 (1998).
7. Basu, S., Basu, A. and Jones, M.C.: Robust and efficient parametric estimation for censored survival data. AISM, **58**, 341–355 (2006)
8. Beran, R.: An efficient and robust adaptive estimator of location. Ann. Statist., **5**, 445–463 (1977)
9. Beran, R.: Minimum Hellinger distance estimates for parametric models. Ann. Statist., **6**, 292–313 (1978)

10. Bickel, J.P.: Quelques aspects de la statistique robuste. Lecture notes in mathematics, **876**, Ecole d't de Saint-Flour (1979)
11. Cox, D.R., Oakes, D.: Analysis of Survival Data. Chapman and Hall (1984)
12. Finkelstein, D.M., Moore, D.F. and Shoenfeld, D.A.: A proportional hazards model for truncated AIDS data. Biometrics, **49**, 731–740 (1993)
13. Frydman, H.: A note on nonparametric estimation of the distribution function from interval-censored and truncated observations. JRSS, B, **56(1)**, 71–74 (1994)
14. Hampel, F.R.: Robust estimation: a condensed partial survey. Zeitschrift für Wahrscheinlichkeitstheorie, **27**, 87–104 (1973)
15. Hampel, F.R.: The influence curve and its role in robust estimation. JASA, **62**, 1179–1186 (1974)
16. Huber, P.J. and Strassen, V.: Minimax tests and the Neyman Pearson lemma for capacities. Ann. Statist., **1**, 251–263 (1973)
17. Huber-Carol, C. and Vonta F.: Frailty models for arbitrarily censored and truncated data. Lifetime Data Anal., **10**, 369–388 (2004)
18. Huber, C., Solev, V. and Vonta, F.: Estimation of density for arbitrarily censored and truncated data. In: Nikulin, M.S. (ed) Probability, Statistics and Modeling in Public Health. Springer, Berlin Heidelberg New York, pp. 246–265 (2006)
19. Huber-Carol, C., Solev, V. and Vonta, F.: Frailty models for censored and truncated data. In: Vonta, F. (ed), BIOSTAT2006, Proc. Intern. Conf. Statist. Meth. Biomed. Tech. Syst., pp. 13–20, 2006.
20. Huber, C. and Nikulin M.S.: Extended cox and accelerated models in reliability with general censoring and truncation. In: Vonta, F. (ed) Statistical Models and Methods for Biomedical and Technical Systems. Birkhauser, pp. 3–21 (2008)
21. Huber-Carol, C., Solev, V. and Vonta, F.: Interval censored and truncated data: rate of convergence of NPMLE of the density. J. Statist. Plan. Infer., **139**, 1734–1749 (2009)
22. Huber, P.J.: Robust Statistics. John Wiley and Sons (1981)
23. Jaeckel, L.A.: Some flexible estimates of location. Ann. Math. Statist., **42**, 1540–1552 (1971)
24. Kosorok, M., Lee, B. and Fine, J.: Robust inference for univariate proportional hazards frailty regression models. Ann. Statist, **32(4)**, 1448–1491 (2004)
25. Meeker, W.Q. and Escobar, L.A.: A review of recent research and current issues in accelerated testing. Int. Statist. Rev., **61**, 147–168 (1993)
26. Meeker, W.Q. and Escobar, L.: Statistical Methods for Reliability Data. John Wiley and Sons (1998)
27. Millar, P.W.: Robust estimation via minimum distance methods. Probability Theory and Related Fields, Springer **55**, **1**, 73–89 (1981)
28. Reid, N.: Influence functions for censored data. Ann. Statist., **9**, 78–92 (1981)
29. Slud, E.V. and Vonta, F.: Consistency of the NPML estimator in the right-censored transformation model. Scand. J. Statist., **31**, 21–41 (2005)
30. Stone, C.: Adaptive maximumlikelihood estimation of a location parameter. Ann. Statist., **3**, 187–195 (1975)
31. Tukey, J.W.: A survey of sampling from contaminated distributions. Contributions to Probability and Statistics. I. Olkin, Ed., Stanford University Press, Stanford, Calif.
32. Turnbull, B.W.: The empirical distribution function with arbitrarily grouped, censored and truncated data. J. Roy. Stat. Soc., **38**, 290–295 (1976)

Modelling Recurrent Events for Repairable Systems Under Worse Than Old Assumption

G. Babykina^{1,2} and V. Couallier²

¹ Cemagref, Cestas Cedex, France

² IMB, University of Bordeaux, Bordeaux, France

Abstract: The objective of the work is to model the failure process of a repairable system under “worse than old”, or harmful repairs, assumption. The proposed model is founded on the counting process probabilistic approach and interprets harmful repairs as the accumulation of failures on the same system. Increase in the conditional intensity is rather induced by the number of previous repair actions than by time contrarily to virtual age models. The LEYP model is defined and some comparison with existing imperfect repair models is given. The explicit form of likelihood function is provided. A covariate-dependent model is defined in order to take the effect of internal or external factors, which may be constant or time dependent, into account. After a description of the estimation procedure for left-truncated and right-censored data using a multiple systems data set, we provide some useful formulae for prediction of the number of failures in a future period. An application to data from the water distribution system of the city of Oslo (Norway) is given.

Keywords and phrases: LEYP model, counting processes, repairable systems, imperfect repairs, recurrent events, time-dependent covariates, maximum likelihood estimation

22.1 Introduction

Ageing of a repairable system, or of a set of repairable systems, includes several un-exclusive notions aimed to model and analyse the increasing intensity of failures or repairs.

The well-known term of recurrent events is usually used in this framework to model the multiple maintenance actions performed on systems rejuvenated by a technical operation, such as repair action after failure, replacement action after failure, replacement action after design change, or replacement action due to a conditional maintenance policy. The ageing (or wear out) of a repairable system is manifested by the increase in the frequency of repair actions undertaken after each failure. That implies more costly restorations of the system’s operating condition and increasing expected number of interventions by time unit.

In this context the most popular model is the non-homogeneous Poisson process (NHPP) for which the corresponding statistical estimation method is long explored [2, 3, 10]. The repairs are then assumed minimal in the sense that after restoration the system is in the same state as just before failure. In other terms the repair action just brings the system back to functioning thus defining an As Bad As Old effect (ABAO). This model characterizes ageing phenomenon only if the corresponding failure intensity is a strictly increasing function of time.

As the system deteriorates, the durations between failures gradually stochastically decrease and the expected number of events, also called the mean cumulated number of failures (MCF), has a convex shape. Its derivative, the rate of occurrence of failures (ROCOF), is then an increasing function of time [16].

The harmful maintenance assumption is opposed to that of perfect maintenance. The successive times of repairs form the i.i.d. sequence of random variables referred to as renewal process, and each maintenance restores the initial condition of a system (condition at time $t = 0$). The repair effect is then supposed As Good As New (AGAN). In this case ageing of a system, or of a set of systems, is not an appropriate notion. Moreover, it has for long been shown that the ROCOF converges to a constant [18].

In some cases, the state of the system just after the repair in terms of the intensity of failure may even be worse than just before the repair (Worse Than Old effect, WTO). This phenomenon can be observed when some units of a set of systems accumulate failures over time, since first failures embrittle the system and increase the probability of future failures.

The modellization is then based on the conditional intensity of the counting process for failures which correspond neither to the Poisson process nor to a renewal process. Numerous models aiming to describe the effect of imperfect maintenance on the conditional intensity of the process have been proposed (see [17, 5, 14, 6] for a review). Most of those use a concept of virtual age by defining the conditional intensity at current time as a function of the system's virtual age.

In the remainder, the covered models only describe recurrent events of repair action caused by failure (corrective maintenance). Assuming that the duration of reparations is neglected, that is, failure times are confounded with repair times, we denote by $(T_i)_{i \geq 1}$ the sequence of successive failure times of one system and $N(t) = \sum_{i \geq 1} \mathbb{1}_{T_i \leq t}$ the associated counting process with $N(0) = 0$.

In relation to the impact of repair on system's condition, some useful models are available, namely

- The NHPP model with increasing intensity function used for As Bad As Old effect of the repair.
- The Renewal model used for As Good As New effect of the repair.
- Generalized renewal model used to model the imperfect maintenance.

The stochastic intensity of a counting process is defined by [1]

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{1}{dt} \mathbb{P}(N(t+dt) - N(t) = 1 \mid \mathcal{H}_{t-}),$$

where \mathcal{H}_{t-} denotes the history of the process up to, but not including, time t . Note that this notation is equivalent to

$$\mathbb{E}[dN(t) | \mathcal{H}_{t-}] = \lambda(t)dt.$$

First, let us introduce the initial deterministic intensity function λ_0 which is the hazard function of the first failure time T_1

$$\lambda_0(t) = \lim_{dt \rightarrow 0} \frac{1}{dt} \mathbb{P}(T_1 \in [t, t + dt] | T_1 > t) = \frac{f_1(t)}{R_1(t)},$$

where f_1 and R_1 are the density function and the reliability function of T_1 , respectively. Obviously, λ_0 does not depend on the effect of the repair action and the conditional intensity $\lambda(\cdot)$ coincides with $\lambda_0(\cdot)$ on the random interval $[0, T_1]$.

As for As Bad As Old models, the stochastic intensity is reduced to $\lambda(t) = \lambda_0(t) \quad \forall t$, and the T_i 's represent the jump points of a non-homogeneous Poisson process. The repair actions have clearly no impact except for the fact that at each failure time, the system falls in a non-functioning state and the instantaneous repair just brings the system back to a functioning state without changing the ageing process. Wear out explicitly requires here that $\lambda_0(\cdot)$ is an increasing function of time. For instance, a power shape for the intensity, $\lambda_0(t) = \alpha t^\beta$ with $\beta > 1$, corresponds to the hypotheses of Weibull-distributed time to the first failure (T_1) with increasing failure rate (IFR) and of minimal repair effect. Note that when T_1 is exponentially distributed ($\beta = 1$), the minimal repair assumption gives an homogeneous Poisson process with intensity constant in time and thus does not model wearing phenomenon. Note also that an NHPP with decreasing intensity models reliability growth instead of wear out [3, 9]. In this case, a minimal repair is advantageous since it does not rejuvenate the system and allows age accumulation.

In case of As Good As New models, the intensity is nicely expressed by $\lambda(t) = \lambda_0(t - T_{N(t)}) \quad \forall t$, where $T_N(t)$ is the time to failure just before t . Hence $(t - T_{N(t)})$ is the time elapsed since the last failure and the trajectories of λ show a total rejuvenation of the system at each repair time. Thus ageing is not modelled and the ROCOF function converges to $1/\mathbb{E}[T_1]$ as t tends to infinity. Note also that in case of exponentially distributed time to first failure, AGAN and ABAO models are equivalent.

Models of imperfect repair, that is, a repair effect between ABAO and AGAN, apply the concept of virtual age [13, 5, 11, 12]. These models can account for the Worse Than Old assumption as a particular case.

In this chapter, we shall focus on a new model (LEYP model) introduced by Le Gat [7] where the repair action has a Worse Than Old effect (WTO). The increasing of the conditional intensity is rather induced by the number of previous repair actions than by time. The LEYP model is defined and some comparison with existing imperfect repair models is given in Section 22.2. We introduce in Section 22.3 a covariate-dependent model in order to take into account the effect of internal or external factors. Constant and time-dependent covariates are used. Section 22.4 describes the estimation procedure for left-truncated and right-censored data using a multiple systems data set and provides some useful formulae for predicting the number of failures in a future period. An application to data from the water distribution system of the city of Oslo (Norway) is given in Section 22.5.

22.2 A New Model of Imperfect Repair: The LEYP Model

In this chapter we propose a new model of imperfect repair under the following heuristics:

- It may be difficult to quantify the effect of the repair by an amount proportional to the virtual current age or to the current intensity of the process.
- In some cases the number of repair actions up to the current time has a heavier impact on failure intensity than ageing.
- In case of constant baseline intensity $\lambda_0(\cdot) = \lambda_0$, the models covered in Section 22.1 cannot catch the repair effect. More specifically, the ABAO assumption is accounted for by using an homogeneous Poisson model, equivalent to the AGAN model with exponential inter-occurrence times between failures $\mathcal{E}(\lambda_0)$. The virtual age models [5] in their turn are not identifiable for constant baseline intensity. Thus a model with identifiable impact of the repair action even for exponential distribution of the first failure time is needed.

The LEYP model assumes that the conditional intensity evolves as

$$\mathbb{E}[dN(t) | \mathcal{H}_{t-}] = \mathbb{E}[dN(t) | N(t-) = j] = (1 + \alpha j) \lambda_0(t) dt, \quad (\alpha > 0). \quad (22.1)$$

Originally, models of conditional intensity of counting process where the conditioning is with respect to the number of past events were developed for birth processes and related processes. The Yule process has intensity [19]

$$\mathbb{E}[dN(t) | \mathcal{H}_{t-}] = \mathbb{E}[dN(t) | N(t-) = j] = \alpha j, \quad (\alpha > 0).$$

The proposed model is an extension of this one (LEYP: linear extended Yule process [7]).

Equation (22.1) can be seen as the probability that the system undergoes a failure in an infinitesimal time interval $[t, t + dt]$ given that j failures had occurred up to time t . Note that for $\alpha = 0$, Equation (22.1) is reduced to a non-homogeneous Poisson process and therefore models ABAO repair effects. The LEYP model with $\lambda_0(t) = e^{\beta_0} \delta t^{\delta-1}$ can be seen as a generalization of the Poisson process with power intensity, referred to as Weibull process (see for instance [20]), name which may be confusing and explained by the Weibull distributed time to first failure.

Assuming that $\alpha > 0$ means that the repair action is harmful, that is, the system condition after the repair is worse than that just before failure (Worse Than Old). Even in case of constant baseline failure rate λ_0 , some ageing phenomenon is modelled by the parameter $\alpha > 0$. Indeed, the conditional intensity λ trajectories are step functions jumping at failure times $(T_i)_{i \geq 1}$, see Figures 22.1 and 22.2.

Precisely, $\lambda(t) = (1 + \alpha i) \lambda_0(t)$, for $t \in [T_i, T_{i+1})$ and, in the simplest case of $\lambda_0(\cdot) = \lambda_0$, we have $T_1 \sim \mathcal{E}(\lambda_0)$, $D_2 = T_2 - T_1 \sim \mathcal{E}((1 + \alpha) \lambda_0)$, \dots , $D_k = T_k - T_{k-1} \sim \mathcal{E}((1 + k\alpha) \lambda_0)$. Incidentally, the ageing phenomenon is due to unsatisfactory maintenances which increase the failure rate of an exponential time to next repair.

For simplicity of presentation, in the remainder of the chapter we shall consider that the baseline intensity is a power function

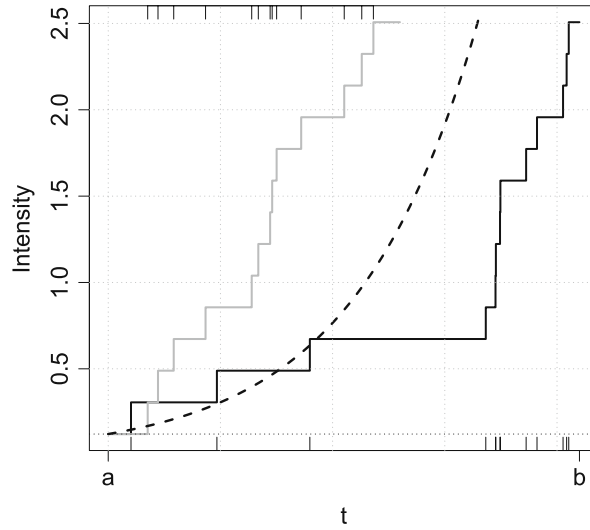


Figure 22.1. Conditional intensities of two sets of simulated times (*plain lines*), ROCOF function (*thick dashed line*), initial intensity (*thin dashed line*) for $\alpha = 1.5$ and $\delta = 1$. $\lambda(t) = (1 + \alpha i)e^{-2.1} \delta t^{\delta-1}$

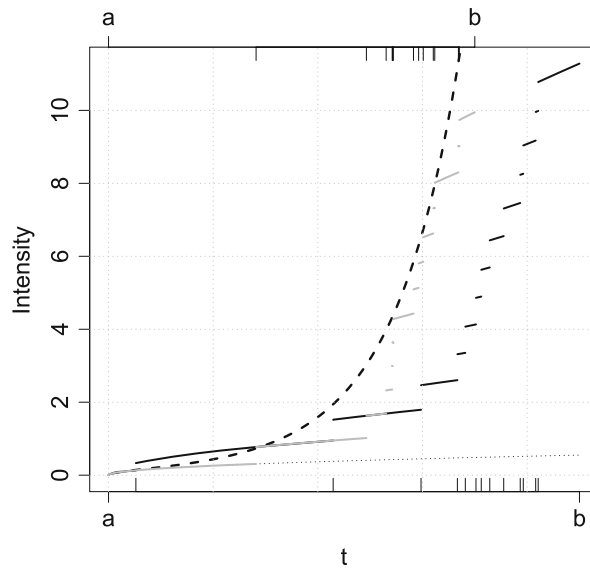


Figure 22.2. Conditional intensities of two sets of simulated times (*plain lines*), ROCOF function (*thick dashed line*), initial intensity (*thin dashed line*) for $\alpha = 1.5$ and $\delta = 1.5$. $\lambda(t) = (1 + \alpha i)e^{-2.1} \delta t^{\delta-1}$

$$\lambda_0(t) = e^{\beta_0} \delta t^{\delta-1}.$$

Let us denote by $\Lambda_0(t) = \int_0^t \lambda_0(s) ds$ the cumulative baseline intensity and $\boldsymbol{\theta} = \{\beta_0, \alpha, \delta\}$ the vector of parameters to estimate.

22.2.1 Some useful properties

References [7, 8] prove that the number of failures in the time period $[a, b]$ conditional to the number of previously occurred failures follows the *Negative Binomial* distribution defined by

$$[N(b) - N(a) \mid N(a-) = m] \sim \mathcal{NB} \left(\alpha^{-1} + m, e^{\alpha[\Lambda_0(b) - \Lambda_0(a)]} \right). \quad (22.2)$$

Equation (22.2) is equivalent to

$$P[N(b) - N(a) = k \mid N(a-) = j] = \frac{\Gamma(\alpha^{-1} + j + k)}{\Gamma(\alpha^{-1} + j)k!} \left(\frac{\mu(a)}{\mu(b)} \right)^{\alpha^{-1} + j} \left(1 - \frac{\mu(a)}{\mu(b)} \right)^k, \quad (22.3)$$

where $\mu(u) = e^{\alpha\Lambda(u)}$.

We also have

$$[N(c) - N(b) \mid N(b-) - N(a) = m] \sim \mathcal{NB} \left(\alpha^{-1} + m, \frac{e^{\alpha\Lambda_0(b)} - e^{\alpha\Lambda_0(a)} + 1}{e^{\alpha\Lambda_0(c)} - e^{\alpha\Lambda_0(a)} + 1} \right) \quad (22.4)$$

and thus the counting process $N(t)$ has the marginal distribution

$$[N(t)] \sim \mathcal{NB} \left(\alpha^{-1}, e^{-\alpha\Lambda_0(t)} \right).$$

By denoting the parameters of the *Negative Binomial* distribution given by Equation (22.4) as $k = \alpha^{-1} + m$ and $p = \frac{e^{\alpha\Lambda_0(b)} - e^{\alpha\Lambda_0(a)} + 1}{e^{\alpha\Lambda_0(c)} - e^{\alpha\Lambda_0(a)} + 1}$, the conditional expectation of the number of events likely to occur in the interval $[c, b]$, knowing the process history, is given by

$$\mathbb{E}[N(c) - N(b) \mid N(b-) - N(a) = m] = k \times \frac{1-p}{p}. \quad (22.5)$$

The MCF and the ROCOF functions are easily obtained as

$$MCF(t) = \mathbb{E}[N(t)] = \frac{1}{\alpha} (e^{\alpha\Lambda_0(t)} - 1)$$

and

$$ROCOF(t) = \frac{d}{dt} \mathbb{E}[N(t)] = \lambda_0(t) e^{\alpha\Lambda_0(t)}.$$

As α tends to zero, it is straightforward to show that these equations are generalizations of those defined for the Poisson model (ABAO effect): $\mathbb{E}[N(t)] = \Lambda_0(t)$ and $ROCOF(t) = \lambda_0(t)$.

22.3 Taking Covariates into Account

Model (22.1) implies that the system's failure intensity entirely results from ageing (for $\delta > 1$) and from occurred failures (for $\alpha > 0$). It can, however, be envisaged that individual characteristics and/or external stress factors also influence the system's fragility. In this case, the covariates effect may be taken into account when modelling the failure intensity [15]. The resulting model is given by (22.6)

$$\mathbb{E}[dN(t) | N(t-) = j, \mathbf{Z}(s), 0 \leq s < t] = (1 + \alpha j) \delta t^{\delta-1} e^{\mathbf{Z}'(t)\boldsymbol{\beta}} dt \quad (22.6)$$

with \mathbf{Z} the $(p+1)$ -dimensional vector of p fixed or time-dependent covariates and $\boldsymbol{\beta} = \{\beta_0, \beta_1, \dots, \beta_p\}$ the coefficients associated with covariate effects.

Thus the vector of parameters to estimate becomes $\boldsymbol{\theta} = \{\alpha, \delta, \boldsymbol{\beta}\}$.

22.4 Statistical Estimation and Data Description

We suppose that the data concerns n systems. The data collection covers a certain calendar time period $[S, E]$ called the observational window, S and E being, respectively, the starting and the ending date of observation.

For each system i ($i = 1, \dots, n$) the installation date, the date of discharge, the individual characteristics and the times of m_i failures within the observational window, $\{t_j\}_{1 \leq j \leq m_i}$, are available.

We assume that the beginning of the observational window does not necessarily coincide with the system's installation date. That is, the complete event history is not systematically observed. The data is therefore said to be left-truncated.

Let a_i and b_i stand for the age of the system i at the beginning and at the end of observational window, respectively. We consider that

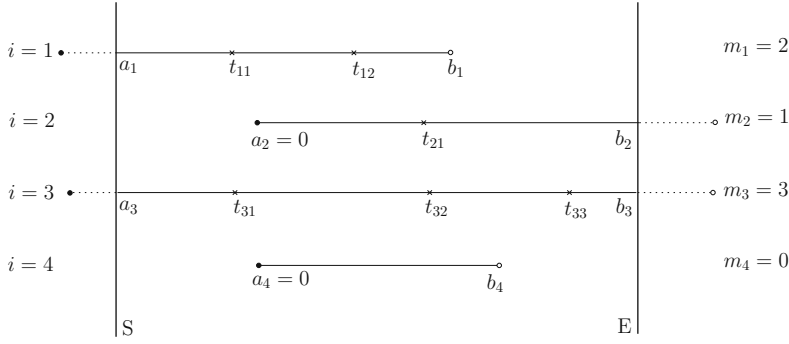
- $a_i = 0$ if system i is installed within $[S, E]$;
- $t_{i(m_i+1)} = b_i$;
- $t_{i0} = a_i$.

Figure 22.3 illustrates the data collection procedure.

The general likelihood function for the failure process of system i depending on the parameter vector $\boldsymbol{\theta}$, as defined in Section 22.3, with m_i failures at times t_{ij} within $[a_i, b_i]$ is expressed by [1]

$$L_i(\boldsymbol{\theta}) = \prod_{j=1}^{m_i} \mathbb{E}[dN(t_{ij}) | \mathcal{N}_{t_i-}] \prod_{j=0}^{m_i} \prod_{t \in t_{ij}, t_{i(j+1)}} \mathcal{P}_{t \in t_{ij}, t_{i(j+1)}} (1 - \mathbb{E}[dN(t_{ij}) | \mathcal{N}_{t_i-}]). \quad (22.7)$$

The product integral term $\prod_{t \in t_{ij}, t_{i(j+1)}} (1 - \mathbb{E}[dN(t_{ij}) | \mathcal{N}_{t_i-}])$ stands for the joint probability of the non-occurrence of failures in each infinitesimal time interval between the j th and $(j+1)$ th failures of the system i . Thus, (22.7) is the likelihood of m_i failures at times t_{ij} and of $m_i + 1$ failure-free periods within the observational window.

**Figure 22.3.** Data collection schema

The likelihood (22.7) can then be expressed by [4]

$$L_i(\boldsymbol{\theta}) = \left(\prod_{j=1}^{m_i} \lambda(t_{ij}) \right) \times \exp \left(- \sum_{j=0}^{m_i} \int_{t_{ij}}^{t_{i(j+1)}} \lambda(u) du \right). \quad (22.8)$$

Writing the explicit form of the likelihood requires defining the conditional probability of failure in $[t, t + dt]$ knowing the history of the process, that is, the covariates and the number of previous observed failures. The latter supposes taking the left-truncated structure of the data into account.

Using (22.8) and (22.3) we can define the likelihood of m_i failures of the observational system i as

$$L_i(\boldsymbol{\theta}) = \alpha^{m_i} \frac{\Gamma(\alpha^{-1} + m_i)}{\Gamma(\alpha^{-1})} \frac{\prod_{j=1}^{m_i} e^{\alpha \Lambda_0(t_{ij})} \lambda_0(t_{ij})}{(e^{\alpha \Lambda_0(b_i)} - e^{\alpha \Lambda_0(a_i)} + 1)^{\alpha^{-1} + m_i}}.$$

The complete failure likelihood of an n sample of systems is defined as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n L_i(\boldsymbol{\theta}).$$

The log-likelihood function to maximize is given by

$$\begin{aligned} \ln L(\boldsymbol{\theta}) &= \sum_{i=1}^n m_i \ln \alpha + \ln \Gamma(\alpha^{-1} + m_i) - \ln \Gamma(\alpha^{-1}) \\ &\quad - (\alpha^{-1} + m_i) \ln(e^{\alpha \Lambda_0(b_i)} - e^{\alpha \Lambda_0(a_i)} + 1) \\ &\quad + \sum_{j=1}^{m_i} (\ln \lambda_0(t_{ij}) + \alpha \Lambda_0(t_{ij})). \end{aligned}$$

The estimated parameters of the model (22.6) can be used to predict the number of future failures of a system, given its individual characteristics and the failure history.

Let us suppose that the system i is observed in a given time period $[a_i, b_i]$, that its event history is left-truncated where applicable and that we aim at predicting the number of failures that the system is likely to undergo in a future time period $[b_i, c_i]$. This framework is illustrated in Figure 22.4.

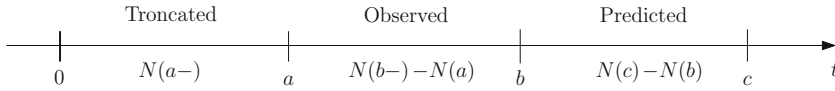


Figure 22.4. Observation and prediction schema

The conditional failure distribution of interest in this case is given by (22.4) and the expected number of future failures is calculated using (22.5).

22.5 Numerical Example

In this section we propose to use the LEYP model to analyse the failure data of the water supply network.

22.5.1 Data description

The available data set concerns the water distribution system of the city of Oslo (Norway). Due to heterogeneous behaviour of different materials in terms of failure process, the material-stratified analysis is usually carried out in practice. Therefore the subset of the grey cast iron pipes is used for estimations.

The data collection covers the 31-year period from 01/01/1976 to 31/12/2006 and concerns 5,111 breaks of 20,476 pipes. Covariates such as the length, the diameter are available as well as the failure dates. The general format of raw data is presented in Tables 22.1 and 22.2 for individual pipe characteristics and for failure history, respectively (for simplicity purpose only the pipe length is taken into account as covariate).

Table 22.1. Raw pipes data format

Ident	Installation	End of service	Length (m)
1	01/01/1937	NA	28.8
2	01/01/1890	20/11/1985	60.0
3	01/01/1940	11/11/1991	7.12
...
20, 475	01/01/1930	NA	70.6
20, 476	01/01/1905	10/05/2000	100.0

Table 22.2. Raw failure data format

Ident	Failure date
1	12/03/1976
1	18/06/1996
...	...
20, 476	12/03/1976
20, 476	18/06/1996
20, 476	10/05/2000

Since the grey cast iron is an old material produced until 1965 and extensively replaced from then onwards, the pipes in the sample are laid from 1851 to 1964 and gradually came up to the end of service since the 1990s (see Figures 22.5 and 22.6 for illustration).

The general trend for failure rate increasing with age is broken by the pipes laid during the post-war period (see Figures 22.7 and 22.8 for illustration).

The remarkable peak of break rate in 1996 (see Figure 22.9) is explained by climatic conditions, which may have a direct effect on failure process (pipes are vulnerable

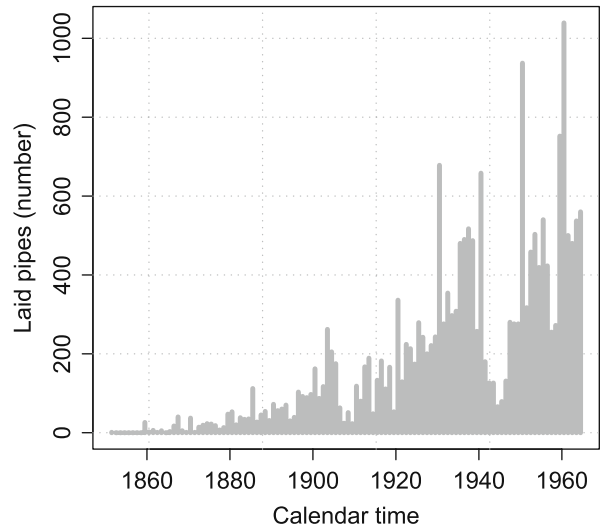


Figure 22.5. Bar chart of installation dates

to cold weather) as well as the indirect one – more resources for leak detection are allocated, resulting in more observed failures.

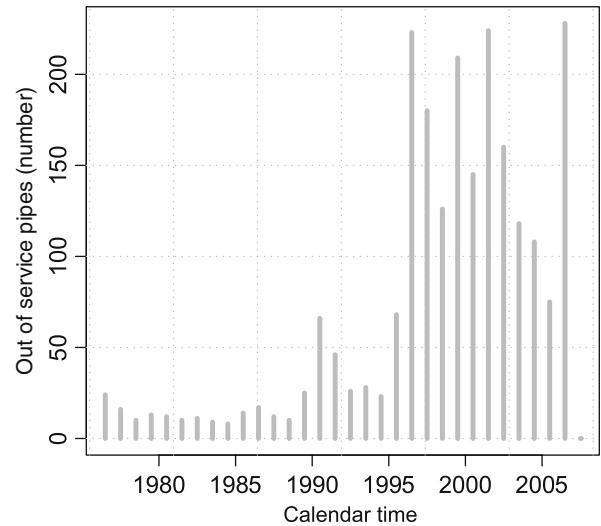


Figure 22.6. Bar chart of replacement dates

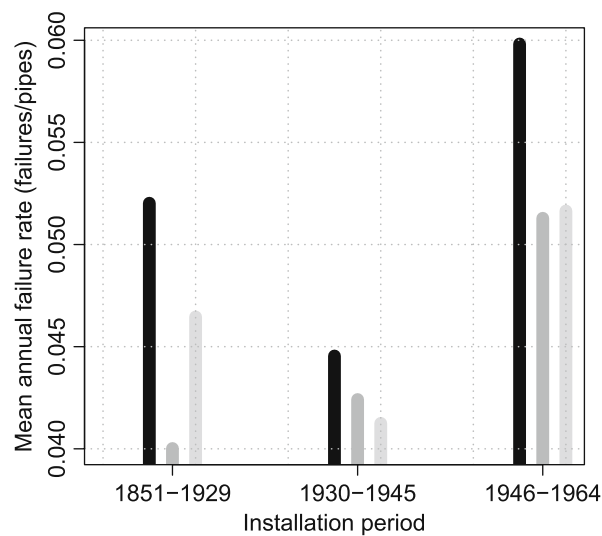


Figure 22.7. Mean annual failure rate per pipe during 1976–1985: *black*, 1986–1995: *darkgrey*, 1996–2006: *lightgrey*

The accumulation of failures on the same pipes (see Table 22.3) motivates the LEYP utilization.

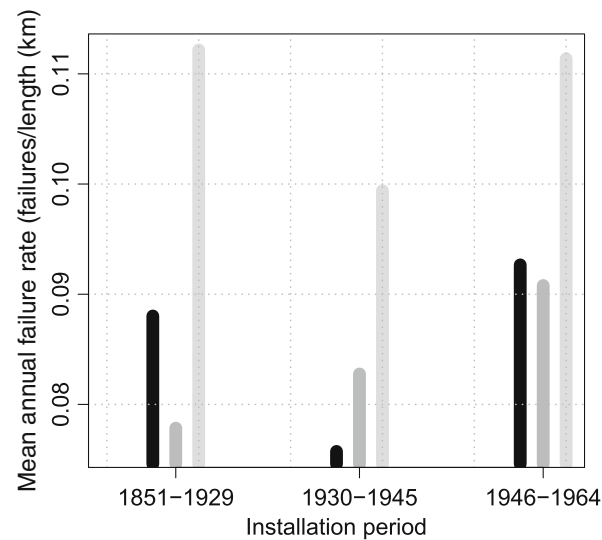


Figure 22.8. Mean annual failure rate per km of length during 1976–1985: *black*, 1986–1995: *darkgrey*, 1996–2006: *lightgrey*

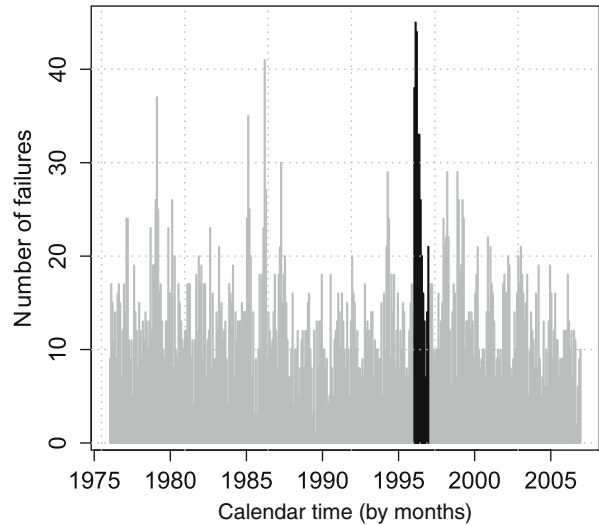


Figure 22.9. Number of monthly failures over the period 1976–2006, the failures of the year 1996 in *black*

Table 22.3. Distribution of the number of failures

No. of failures	0	1	2	3	4	5	6	7	8	9	10
No. of pipes	16, 792	2, 719	682	176	71	20	8	2	1	4	1
% of pipes	82.01	13.28	3.33	0.86	0.35	0.10	0.04	0.01	<0.01	0.02	<0.01

22.5.2 Parameter estimation

The model parameter estimates are obtained using the failure data from the period 01/01/1976 to 31/12/2003 (estimation period). The data covering the period from 01/01/2004 to 31/12/2006 (prediction period) is used to validate the results.

We include the logarithm of the pipe’s length as a fixed continuous covariate and the 1996 year indicator as a dummy time-dependent covariate. The model defined by (22.1) becomes

$$\mathbb{E} [dN(t) \mid N(t-) = j, \mathbf{Z}(t)] = (1 + \alpha j) \delta t^{\delta-1} e^{\beta_0 + \beta_1 \times \ln(\text{length}) + \beta_2 \times \mathbb{1}_{1996}(t)} dt.$$

Let us denote by s_i the age of the pipe i at 01/01/1996 and e_i its age at 31/12/1996 to define

$$\mathbb{1}_{1996}(t) = \begin{cases} 1 & \text{if } t \in [s_i, e_i] \\ 0 & \text{otherwise} \end{cases}.$$

For comparison purposes, the estimations are performed using the reduced observational period (1976–2003) and the whole failure data (1976–2006), in the presence of the time-dependent covariate (complete model) and for the reduced fixed-covariate model.

The inverse of the empirical negative Hessian to compute the variance of the estimator is used assuming the asymptotic normality of the maximum likelihood estimate.

The statistical significance of the coefficients θ_i is assessed by means of the Wald test of the hypothesis

$$\begin{aligned} H_0 : \theta_i &= \theta_0 \\ H_1 : \theta_i &\neq \theta_0. \end{aligned}$$

The parameter estimates are presented in Table 22.4.

The following conclusions can be derived from the obtained results:

- No significant change can be stated in the general trend of the failure occurrence between the reduced and the complete estimation periods.
- All else being equal and at each time point, past failures increase the instantaneous propensity to fail, and thus, the breaks are accumulated on the same pipes ($\hat{\alpha} > 0$). The ABAO assumption is therefore rejected in favour of WTO effect of the repair actions.
- Ageing, the natural factor of increasing break rate, does not appear to condition the failure process in the analysis ($\hat{\delta}$ is not significantly different from 1). This can be partially explained by the post-war installation period phenomenon covered in Section 22.5.1.

Table 22.4. Estimation results

Parameter	Estimation period	Model	Estimate	95% CI	P-value	θ_0
$\hat{\alpha}$	01/01/1976–31/12/2003	Reduced	1.78	[1.60 1.97]	< 0.000001	0
		complete	1.77	[1.59 1.95]	< 0.000001	0
	01/01/1976–31/12/2006	Reduced	1.72	[1.55 1.89]	< 0.000001	0
		complete	1.71	[1.54 1.88]	< 0.000001	0
$\hat{\delta}$	01/01/1976–31/12/2003	Reduced	1	[0.93 1.07]	0.999997	1
		complete	1	[0.93 1.07]	0.999966	1
	01/01/1976–31/12/2006	Reduced	1	[0.93 1.06]	0.999967	1
		complete	1	[0.93 1.07]	0.999991	1
$\hat{\beta}_0$	01/01/1976–31/12/2003	Reduced	−6.92	[−7.26 − 6.58]	< 0.000001	0
		complete	−6.97	[−7.31 − 6.62]	< 0.000001	0
	01/01/1976–31/12/2006	Reduced	−6.97	[−7.30 − 6.63]	< 0.000001	0
		complete	−7.01	[−7.35 − 6.67]	< 0.000001	0
$\hat{\beta}_1$	01/01/1976–31/12/2003	Reduced	0.43	[0.40 0.47]	< 0.000001	0
		complete	0.44	[0.40 0.47]	< 0.000001	0
	01/01/1976–31/12/2006	Reduced	0.44	[0.41 0.48]	< 0.000001	0
		complete	0.45	[0.41 0.48]	< 0.000001	0
$\hat{\beta}_2$	01/01/1976–31/12/2003	complete	0.63	[0.51 0.74]	< 0.000001	0
	01/01/1976–31/12/2006	complete	0.66	[0.55 0.78]	< 0.000001	0

- The pipes's length is a risk factor ($\hat{\beta}_1 > 0$). That is, all else being equal, the instantaneous relative risk (RR) of failure for two pipes with lengths l_1 and l_2 , respectively, is given by

$$RR = \frac{\lambda(t, l_1)}{\lambda(t, l_2)} = \left(\frac{l_1}{l_2} \right)^{\hat{\beta}_1}.$$

- All else being equal, the failure rate is about twice higher during 1996 ($e^{\hat{\beta}_2}$).

22.5.3 Predictions

Equation (22.5) is applied to use the parameter estimates from 01/01/1976 to 31/12/2003 data to predict the number of failures likely to occur in every future 10-day period (prediction period: 01/01/2004 – 31/12/2006). The estimated number is then compared to the number of failures actually occurred. The cumulated number of real and predicted failures is illustrated in Figure 22.10. Figure 22.11 gives the prediction error. The estimates provide quite accurate failure rate predictions: ± 15 failures per 10-day period. Taking the time-dependent covariate into account (complete model) improves the forecast quality.

We notice, however, the cyclic behaviour of the prediction error, which can correspond to seasonal variations in empirical failure rate. This phenomenon can be accounted for by including a continuous time-dependent covariate such as the coldest monthly air or water temperature.

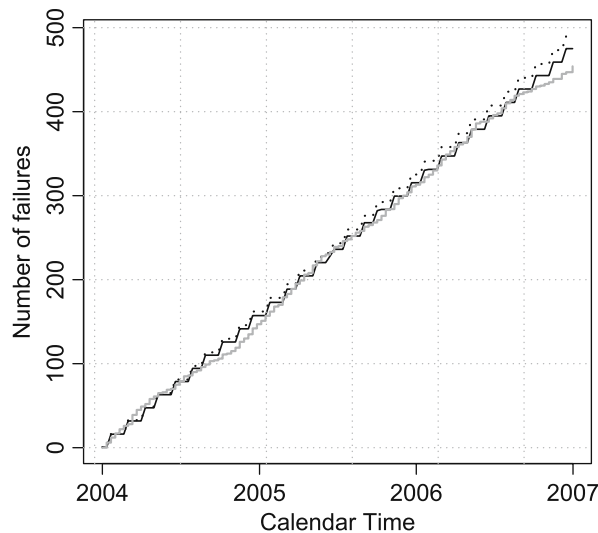


Figure 22.10. Cumulat number of observed failures (in *grey*), predicted failures using the complete model (*plain black line*) and the reduced model (*dotted line*)

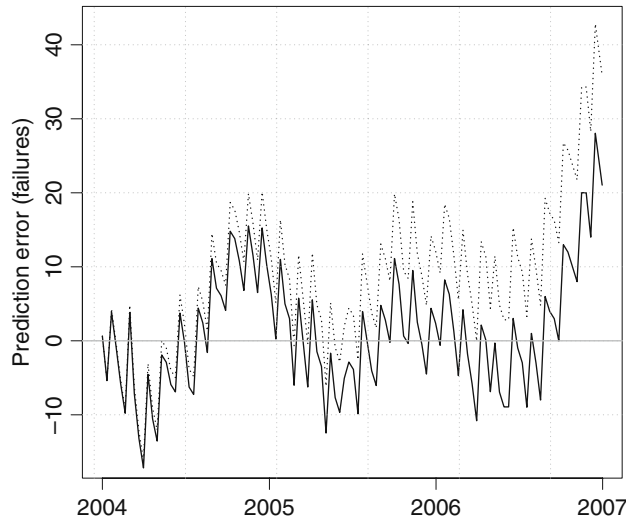


Figure 22.11. Prediction error in 10-day periods (real failures – predicted failures) calculated using the complete model (*plain line*) and reduced model (*dotted line*)

22.6 Conclusion

A new imperfect repair model has been defined for which the stochastic intensity is driven by the cumulated number of previous failures. The model accounts for time-dependent covariate and thus is adapted to dynamic environment. Truncated or censored data can be envisaged as well as several systems with different start and discharge times. The statistical parameter estimation allows predicting the number of future events of a system or of a set of systems. Further work is needed to develop some goodness-of-fit tests. Unlike the well-known virtual age models, LEYP, as defined in the chapter, does not allow to directly take the advantageous effect of repair into account. Defining a new way of modelling the repeated previous failure effects would overcome this drawback.

References

1. Andersen, P.K., Borgan, O., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Process*. Springer-Verlag.
2. Crow, L.H. (1982). Confidence interval procedures for the Weibull process with applications to reliability growth. *Technometrics*, **24**, 67–72.
3. Crow, L.H. (1990). Evaluating the reliability of repairable systems. In: *Proceedings of the Reliability and Maintainability Symposium*, pp. 275–279.
4. Coccozza-Thivent, C. (1997). *Processus stochastiques et fiabilité des systèmes*. Springer.

5. Doyen, L. and O. Gaudoin, O. (2004). Classes of imperfect repair models based on reduction of failure intensity or virtual age. *Reliability Engineering and System Safety*, **84**, 45–56.
6. Finkelstein, M.S. (2000). Modeling a process of non-ideal repair. In: *Recent Advances in Reliability Theory. Methodology, Practice, and Inference*, pp. 41–53. Birkhäuser, Statistics for Industry and Technology.
7. Le Gat, Y. (2009). Extending the Yule process to model recurrent failures of pressure pipes. Submitted to *Journal of Statistical Planning and Inference*.
8. Le Gat, Y. (2009). *Une extension du processus de Yule pour la modélisation stochastique des vnements rcurrents*. PhD thesis, ENGREF.
9. Gaudoin, O. and Ledoux, J. (2007). *Modlisation alatoire en fiabilit des logiciels*. Hermès - Lavoisier.
10. Høyland, A. and Rausand, M. (1994). *System Reliability Theory: Models and Statistical Methods*. John Wiley & Sons, New York.
11. Jack, N. (1998). Age-reduction models for imperfect maintenance. *IMA Journal of Management Mathematics*, **9**, 347–354.
12. Kahle, W. (2007). Optimal maintenance policies in incomplete repair models. *Reliability Engineering and System Safety*, **92**, 563–565.
13. Kijima, M. (1989). Some results for repairable systems with general repair. *Journal of Applied Probability*, **26**, 89–102.
14. Krivtsov, V. (2007). Recent advances in theory and applications of stochastic point process models in reliability engineering. *Reliability Engineering and System Safety*, **92**, 549–551.
15. Lawless J.F. (2000). Dynamic analysis of failures in repairable systems and software. In: *Statistical and Probabilistic Models in Reliability*, pp. 341–351. Birkhäuser, Boston.
16. Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*. Wiley Series in Probability and Mathematical Statistics, Applied Section. New York, 680p.
17. Pham, H. and Wang, H. (1996). Imperfect maintenance. *European Journal of Operational Research*, **94**, 425–438.
18. Rigdon, S.E. and Basu, A.P. (2000). *Statistical Methods for the Reliability of Repairable Systems*. John Wiley & Sons, New York.
19. Ross, S. (1996). *Stochastic Processes*. John Wiley & Sons, New York.
20. Weckman, G.R., Shell, R.L., and Marvel, J.H. (2001). Modeling the reliability of repairable systems in the aviation industry. *Computers & Industrial Engineering*, **40**, 51–63.

Survival Models for Step-Stress Experiments With Lagged Effects

N. Kannan¹, D. Kundu², and N. Balakrishnan³

¹ Department of Management Science and Statistics, The University of Texas at San Antonio, San Antonio, Texas, USA

² Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India

³ Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada

Abstract: In this chapter, we consider models for experiments in which the stress levels are altered at intermediate stages during the exposure. These experiments, referred to as step-stress tests, belong to the class of accelerated models that are extensively used in reliability and life-testing applications. Models for step-stress tests have largely relied on the cumulative exposure model (CEM) discussed by Nelson. Unfortunately, the assumptions of the model are fairly restrictive and quite unreasonable for applications in survival analysis. In particular, under the CEM the hazard function has discontinuities at the points at which the stress levels are changed.

We introduce a new step-stress model where the hazard function is continuous. We consider a simple experiment with only two stress levels. The hazard function is assumed to be constant at the two stress levels and linear in the intermediate period. This model allows for a lag period before the effects of the change in stress are observed. Using this formulation in terms of the hazard function, we obtain the maximum likelihood estimators of the unknown parameters. A simple least squares-type procedure is also proposed that yields closed-form solutions for the underlying parameters. A Monte Carlo simulation study is performed to study the behavior of the estimators obtained by the two methods for different choices of sample sizes and parameter values. We analyze a real data set and show that the model provides an excellent fit.

Keywords and phrases: Accelerated testing, step-stress model, cumulative exposure model, lagged effects, hazard function, bootstrap method

23.1 Introduction

In many applications in the physical and biomedical sciences, experimental conditions may change during the exposure duration. For example, travellers to high mountainous regions or mountain climbers frequently experience acute mountain sickness (AMS) or

altitude sickness. To avoid/decrease symptoms, it is recommended that individuals acclimatize by slowly increasing their elevation. Tables for altitude climbers and deep-sea divers are available that specify the length of time at each altitude. This staged ascent/descent provides protection by “delaying” or eliminating the onset of symptoms.

In a recent article by Greven et al. [7], the authors examine the effects of water contamination on fish, in particular on swimming performance. The underlying hypothesis was that fish exposed to toxins may exhibit a lower threshold for fatigue. Fatigue was induced by increasing the stress, in this case water velocity, at fixed time points during the experiment. Stress tests are also routinely used to assess cardiac function, with the speed and elevation of the treadmill being increased progressively during the experiment.

In the first example, the stress levels were increased with the hope that a gradual increase would significantly prolong the time to symptoms when compared to a direct ascent/descent to the final altitude. In the remaining two examples, the experiment was designed to reduce the time to fatigue by gradually increasing the stress levels.

All the examples describe what are commonly termed “step-stress tests” in the reliability literature. Step-stress experiments are a particular type of accelerated test routinely used in life-testing experiments. With an increased emphasis on quality and reliability, many products manufactured today have extremely long times to failure on the average. The products, however, must undergo rigorous testing to determine the effect of different stress factors on the reliability and performance. In accelerated testing, the items or units are subjected to higher stress levels than normal. This increased stress induces shorter failure times. Using a model relating the stress levels and failure distributions, it is often possible to determine the properties of the underlying failure distribution under normal operating conditions.

In a standard step-stress experiment, all individuals or items are subject to an initial stress level. The stress is gradually increased at pre-specified times during the exposure. The stress factor may refer to the dose of a drug, elevation of the treadmill, temperature, voltage, and pressure. It is surprising that there are very limited applications of step-stress tests in the survival literature: most references on accelerated testing are found in the reliability literature.

Key references in the broad area of accelerated testing include the books by Nelson [16], Meeker and Escobar [13], and Bagdonavicius and Nikulin [2]. Step-stress models have been studied quite extensively in the literature using the CEM formulation discussed earlier by Nelson [15], [16]. The reader may refer to the works of DeGroot and Goel [6], Miller and Nelson [14], Bai et al. [3], Khamis and Higgins [12], Xiong [19], Xiong and Milliken [20], Gouno and Balakrishnan [8], Bagdonavicius et al. [1], Gouno et al. [9], Han et al. [10], and Balakrishnan et al. [5] for some other important developments on step-stress testing. Balakrishnan [4] has recently provided a synthesis of exact inferential results and optimal accelerated tests in the context of exponential step-stress models.

In Section 23.2, we will formally introduce the CEM and discuss its properties and limitations. We introduce a new model that accounts for the possibility that units will not show immediate effects of the stress change. We model this “lag” effect using a piecewise continuous hazard function. The new model includes the CEM as a limiting case. We discuss both likelihood and least squares-type estimators and assess their performance using Monte Carlo simulations. We use the model to analyze a real data set and show that the model provides an excellent fit.

23.2 Model Description

The standard formulation for step-stress experiments uses the cumulative exposure model discussed by Nelson [16]. The CEM relates the survival distribution of the units at one stress level to the next level. A key assumption of the model is that the residual life of the experimental units depends only on the cumulative exposure they have experienced, with no memory of how this exposure was accumulated.

If we look at the assumptions of the model in terms of the hazard function, we notice that it translates to discontinuities at the points at which the stress levels are changed. In other words, the effect of the change in stress level is instantaneous. This is clearly not reasonable for most applications: the effect of a change in stress will produce an increase in the risk but it seems quite likely that one will observe a lag or latency period. The assumption of an instantaneous jump in the hazard function, though unrealistic leads to a more simple and tractable model.

We consider a more realistic step-stress model that accounts for these latency periods using a piecewise hazard model. This new model reduces to Nelson's model under certain limiting conditions.

23.2.1 Step-stress models with latency

Consider a simple experiment wherein the stress is changed only once during the exposure duration. We assume that the hazard function associated with the initial and elevated stress levels is constant. All n individuals or items are exposed to the same initial stress level x_1 . Subjects are continuously monitored, and at some prespecified time point τ_1 , the level of the stress is increased to x_2 . The effect of increasing the stress is not seen immediately: we assume that there is a known latency period δ before the effects are completely observed. In the interval $[\tau_1, \tau_2]$ where $\tau_2 = \tau_1 + \delta$, the hazard slowly increases.

The piecewise hazard function is assumed to have the following form:

$$h(t) = \begin{cases} \theta_1, & 0 < t < \tau_1; \\ a + bt, & \tau_1 \leq t < \tau_2; \\ \theta_2, & t \geq \tau_2, \end{cases} \quad (23.1)$$

where $\tau_2 > \tau_1$ and both time points are known. The parameters a and b in the model are chosen to ensure the hazard function $h(t)$ is continuous. Therefore, a and b satisfy

$$a + b\tau_1 = \theta_1, \quad (23.2)$$

$$a + b\tau_2 = \theta_2. \quad (23.3)$$

The hazard is constant in the intervals $[0, \tau_1]$ and $[\tau_2, \infty)$. In the interval (τ_1, τ_2) , the hazard changes linearly. The parameter b measures how quickly the effects of the increased stress are observable. In this model, we assume that both τ_1 and τ_2 are known. We can easily extend this model to the case when both these constants are unknown.

We call this the cumulative risk model (CRM) to emphasize the accumulated effects of the stress on the lifetimes. Using the definition of the hazard function in (23.1), the survival function (SF), $S(t)$ may be written as

$$S(t) = \begin{cases} e^{-(a+b\tau_1)t}, & 0 < t < \tau_1; \\ e^{-at - \frac{b(t^2 + \tau_1^2)}{2}}, & \tau_1 \leq t < \tau_2; \\ e^{-(a+b\tau_2)t - \frac{b}{2}(\tau_1^2 - \tau_2^2)}, & t \geq \tau_2. \end{cases}$$

The corresponding probability density function (PDF), $f(t)$, is

$$f(t) = \begin{cases} (a + b\tau_1)e^{-(a+b\tau_1)t}, & 0 < t < \tau_1; \\ (a + bt)e^{-at - \frac{b(t^2 + \tau_1^2)}{2}}, & \tau_1 \leq t < \tau_2; \\ (a + b\tau_2)e^{-(a+b\tau_2)t - \frac{b}{2}(\tau_1^2 - \tau_2^2)}, & t \geq \tau_2. \end{cases}$$

The cumulative hazard (CH) function is

$$H(t) = \begin{cases} (a + b\tau_1)t & 0 < t < \tau_1; \\ (a + b\tau_1)\tau_1 + \frac{1}{2}(t - \tau_1)^2b & \tau_1 \leq t < \tau_2; \\ (a + b\tau_1)\tau_1 + \frac{1}{2}(\tau_2 - \tau_1)^2b + (t - \tau_2)(a + b\tau_2), & t \geq \tau_2. \end{cases} \quad (23.4)$$

If we assume that the hazard functions are constant at both stress levels x_1 and x_2 , then the cumulative hazard function for Nelson's CEM is

$$H(t) = \begin{cases} \theta_1 t & 0 < t < \tau_1; \\ \theta_1 \tau_1 + \theta_2(t - \tau_1) & t \geq \tau_1. \end{cases}$$

It is clear that, in (23.4), if we let $\delta = (\tau_2 - \tau_1) \rightarrow 0$ with $\delta b \rightarrow (\theta_2 - \theta_1)$, then the CRM converges to the above CEM formulation of Nelson.

To compare Nelson's CEM with this new model, we plot the cumulative hazard for the two models. A graph is provided in Figure 23.1 for $\tau_1 = 1, \tau_2 = 3$. The cumulative hazard function for both models is identical up to τ_1 . At τ_1 , the CEM is linear with a slope θ_2 , while the CRM is quadratic. After τ_2 , the CRM is linear as well.

In the next section, we derive the estimators of a and b for the CRM using the maximum likelihood approach.

23.3 Maximum Likelihood Estimators for the CRM

Let

$$t_1 < \dots < t_{n_1} < \tau_1 < t_{n_1+1} < \dots < t_{n_1+n_2} < \tau_2 < t_{n_1+n_2+1} < \dots < t_n \quad (23.5)$$

be the ordered failure times. Here n_1 , n_2 , and $n_3 = n - (n_1 + n_2)$ denote the number of failures that occur before τ_1 , between τ_1 and τ_2 , and beyond τ_2 , respectively.

The likelihood function as a function of a and b is given by

$$\begin{aligned} l(a, b) = & (a + b\tau_1)^{n_1} e^{-(a+b\tau_1)\sum_{i \in I_1} t_i} \times \prod_{i \in I_2} (a + bt_i) e^{-\sum_{i \in I_2} (at_i + \frac{bt_i^2}{2})} e^{-\frac{n_2 b \tau_1^2}{2}} \\ & \times (a + b\tau_2)^{n_3} e^{-(a+b\tau_2)\sum_{i \in I_3} t_i} \times e^{-\frac{n_3 b}{2}(\tau_1^2 - \tau_2^2)}, \end{aligned} \quad (23.6)$$

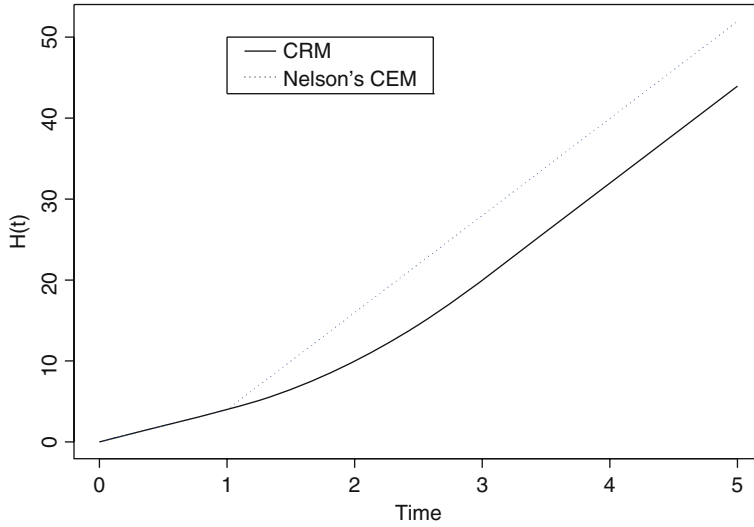


Figure 23.1. Cumulative hazard functions of cumulative exposure and cumulative risk models

where $I_1 = \{1, \dots, n_1\}$, $I_2 = \{n_1+1, \dots, n_1+n_2\}$, and $I_3 = \{n_1+n_2+1, \dots, n\}$. We use the convention that $\sum_{i \in I_j}(\cdot) = 0$ if I_j is an empty set. Note that if $n_1 > 0, n_2 > 0, n_3 > 0$, then $(n_1, n_2, t_{n_1+1}, \dots, t_{n_1+n_2}, \sum_{i \in I_1} t_i, \sum_{i \in I_2} t_i)$ is a complete sufficient statistic.

The following observations are clear from the form of the likelihood function in (23.6):

- If $n_2 > 0$, then the MLE's of a and b always exist.
- If $n_2 = 0$, then the MLE's of a and b exist if both $n_1 > 0$ and $n_3 > 0$.
- If $n_2 = 0$ and either one of the two remaining counts is zero, then, the MLE's of a and b do not exist.

The derivation of the maximum likelihood estimators is provided in the appendix, along with the expressions for the observed Fisher information. We note that the MLE of b can be obtained as the solution of the nonlinear equation

$$-1 + \frac{n_1}{n - bK + b\tau_1 T} + \frac{n_3}{n - bK + bT\tau_2} + \sum_{i \in I_2} \frac{1}{n - bK + bTt_i} = 0. \quad (23.7)$$

Here, $T = \sum_{i=1}^n t_i$ and

$$K = \tau_1 \sum_{i \in I_1} t_i + \frac{1}{2} \sum_{i \in I_2} t_i^2 + \frac{1}{2} n_2 \tau_1^2 + \tau_2 \sum_{i \in I_3} t_i - \frac{1}{2} n_3 (\tau_2^2 - \tau_1^2). \quad (23.8)$$

Once the MLE of b , \hat{b} , is obtained as the solution of (23.7), then the MLE of a , say \hat{a} , is simply

$$\hat{a} = \frac{n - \hat{b}K}{T}. \quad (23.9)$$

Remark: Suppose $n_2 = 0$. In this case, we can easily see from (23.7) that

$$\hat{b} = \frac{nK - T(n_1\tau_1 + n_3\tau_2)}{(K - T\tau_1)(K - T\tau_2)}. \quad (23.10)$$

Remark: Suppose the ratio $\frac{\theta_2}{\theta_1}$ is known. The case of the known ratio is often intuitive in a reliability setting. We may have a physical model that relates the stress levels and the corresponding mean failure times. This is equivalent to saying that the ratio $\frac{a}{b}$ is a known quantity, say α . In this case,

$$\hat{b} = \frac{n}{K + \alpha T} \quad \text{and} \quad \hat{a} = \alpha \hat{b}. \quad (23.11)$$

23.4 Least Squares Estimators

In this section, we propose an alternative method for estimating a and b that does not require solving a nonlinear equation. This method is based on least squares and yields estimators that are in explicit form; one may refer to Johnson et al. [11] and Swain et al. [18] for details on this least squares estimation technique. Since the cumulative hazard function for the CRM is simple, it is natural to obtain a and b by minimizing the least squares distance between the empirical cumulative hazard function and the fitted hazard function with respect to the unknown parameters.

The estimators of a and b , say \tilde{a} and \tilde{b} , can be obtained by minimizing

$$\sum_{i=1}^n \left(H(t_i) - \hat{H}(t_i) \right)^2 \quad (23.12)$$

with respect to the unknown parameters a and b . Here, $H(t_i)$ and $\hat{H}(t_i)$ are the cumulative hazard function and the estimated cumulative hazard function, respectively. We use the standard nonparametric estimator given by

$$\hat{H}(t_i) = -\ln(\hat{S}(t_i)) = \ln n - \ln(n - i + 1) = c_i \quad (\text{say}). \quad (23.13)$$

The cumulative hazard function $H(t)$ has a simple form given in (23.4) using which \tilde{a} and \tilde{b} can be obtained as follows:

$$\begin{aligned} \tilde{a} &= \frac{(\sum_{i=1}^n v_i^2)(\sum_{i=1}^n c_i t_i) - (\sum_{i=1}^n t_i v_i)(\sum_{i=1}^n c_i v_i)}{(\sum_{i=1}^n t_i^2)(\sum_{i=1}^n v_i^2) - (\sum_{i=1}^n t_i)(\sum_{i=1}^n v_i)}, \\ \tilde{b} &= \frac{(\sum_{i=1}^n t_i^2)(\sum_{i=1}^n c_i v_i) - (\sum_{i=1}^n t_i v_i)(\sum_{i=1}^n c_i t_i)}{(\sum_{i=1}^n t_i^2)(\sum_{i=1}^n v_i^2) - (\sum_{i=1}^n t_i)(\sum_{i=1}^n v_i)}, \end{aligned}$$

where

$$v_i = \begin{cases} \tau_i t_i, & 1 \leq i \leq n_1; \\ \frac{1}{2}(\tau_1^2 + t_i^2), & n_1 \leq i \leq n_1 + n_2; \\ \frac{1}{2}(\tau_1^2 - \tau_2^2) + \tau_2 t_i, & i > n_3. \end{cases}$$

Remark: If $\frac{a}{b} = \alpha$ with α known, then

$$\tilde{b} = \frac{\alpha \sum_{i=1}^n c_i t_i + \sum_{i=1}^n c_i v_i}{\sum_{i=1}^n (\alpha t_i + v_i)^2} \quad \text{and} \quad \tilde{a} = \alpha \tilde{b}. \quad (23.14)$$

23.5 Data Analysis

In this section, we analyze the data from Greven et al. [7]. In the paper, the authors consider the swimming performances of two groups of fish. The first group consisted of offspring where the parents were exposed to a contaminant. The second group was a control group with the same characteristics except for the lack of exposure to the chemical. The fish were placed in a chamber where the flow rate of the water was maintained at 15 cm/sec. After 90 minutes, the flow rate was increased every 20 seconds by 5 cm/sec. The time at which the fish could no longer maintain its position in the chamber was recorded which is the fatigue time. We assumed for simplicity that there was only one stress level. The empirical cumulative hazard suggests that the first increase has a significant effect: the effect of subsequent increases is not evident.

For each data set, we obtained the MLEs and the LSEs based on the cumulative risk model. Since τ_2 is unknown, we used a discrete optimization method to obtain an estimate from the data. The results of the two analyses are given in Tables 23.1 and 23.2. The tables also provide the approximate and bootstrap (parametric) confidence intervals for the two estimators.

Table 23.1. Estimates and confidence intervals for the control group

Methods	\hat{a}	\hat{b}	$\hat{\theta}_1$	$\hat{\theta}_2$
MLE	-0.23500	0.00218	0.00421	0.10424
($\hat{\tau}_2 = 156$)	(-0.48422, -0.11501)	(0.00103, 0.00443)	(0.00212, 0.00922)	(0.05712, 0.20786)
	(-0.42762, -0.04238)	(0.00043, 0.00393)	(0.00100, 0.00860)	(-0.08811, 0.29827)
LSE	-0.23469	0.00216	0.00325	0.06598
($\hat{\tau}_2 = 139$)	(-0.45200, -0.07468)	(0.00074, 0.00405)	(0.00044, 0.00710)	(0.02713, 0.12058)
	(-0.41597, -0.05341)	(0.00051, 0.00381)	(0.00060, 0.00522)	(-0.07433, 0.20542)

Table 23.2. Estimates and confidence intervals for the test group

Methods	\hat{a}	\hat{b}	$\hat{\theta}_1$	$\hat{\theta}_2$
MLE	-0.67057	0.00612	0.00248	0.03307
($\hat{\tau}_2 = 115$)	(-1.29324, -0.30603)	(0.00282, 0.01177)	(0.00127, 0.00684)	(0.01824, 0.06235)
	(-1.14949, -0.19165)	(0.00177, 0.01047)	(0.00053, 0.00473)	(-0.11280, 0.17926)
LSE	-1.34638	0.01226	0.00214	0.02666
($\hat{\tau}_2 = 112$)	(-2.55898, -0.54740)	(0.00497, 0.02276)	(0.00003, 0.00491)	(0.01350, 0.04754)
	(-2.29977, -0.39299)	(0.00359, 0.02093)	(0.00046, 0.00398)	(-0.15589, 0.20937)

Data for the Control Group: 83.50, 91.00, 91.00, 97.00, 107.00, 109.50, 114.00, 115.41, 128.61, 133.53, 138.58, 140.00, 152.08, 155.10.
Data for the Test Group : 91.00, 93.00, 94.00, 98.20, 115.81, 116.00, 116.50, 117.25, 126.75, 127.50, 154.33, 159.50, 164.00, 184.14, 188.33.

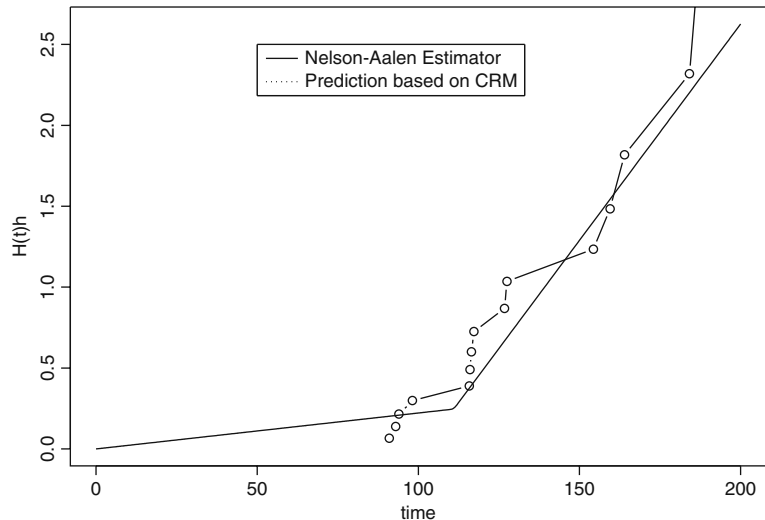


Figure 23.2. Empirical and predicted CH functions: test group

We computed the Kolmogorov–Smirnov (KS) distance between the empirical distribution function and the estimated CDF based on the CRM. For the control group, the KS distance using the MLEs (LSEs) was 0.3302 (0.2639), with corresponding p values of 0.0736 (0.2422). For the test group, the KS distance using the MLEs (LSEs) was 0.2128 (0.1946) with corresponding p values of 0.4564 (0.5726). These results reveal that the CRM model does provide an excellent fit to the data. Figure 23.2 provides an overlay plot of the empirical cumulative hazard function and the predicted CH function based on the least squares fit of the CRM. The model is clearly able to isolate the “change point”. We can say that there is approximately a 2-minute delay before the effects of the flow rate increase are really observable. This observation also provides some rationale for considering the case of a single stress level. The overlay plot does not show the quadratic nature of the cumulative hazard function: if we magnify the region from 90 minutes to 95 minutes, that trend is visible.

23.6 Simulation Results

To assess the performance of the MLEs and the least squares-type estimators, a simulation study was conducted. We considered different sample sizes and different values for the parameters. For each combination of sample size and parameter values, the bias

and the square root of the mean squared errors (RMSE's) are reported for both estimators in Tables 23.3–23.10. All the simulations are performed using RAN2 random deviate generator of Press et al. [17].

Each pair of tables uses the same parameter values: the second table in the pair assumes a fixed ratio. All the results are based on 1,000 replications. For one simulation with $n = 25$, $\tau_1 = 100$, $\tau_2 = 150$, $\theta_1 = 0.01$ and $\theta_2 = 0.02$, we also provide the sampling distribution of the MLEs of a , b , θ_1 , and θ_2 . These distributions are presented in Figures 23.3–23.6 and these figures display that the distributions are skewed and not close to normal. We would, therefore, recommend against using confidence intervals based on normality unless the sample sizes are large.

Table 23.3. $\tau_1 = 100$, $\tau_2 = 150$, $\theta_1 = 1/100$, $\theta_2 = 1/50$

n	Methods	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{MSE}(\hat{\theta}_1)$	$\text{MSE}(\hat{\theta}_2)$
25	MLE	0.01139	0.04032	0.00285	0.03266
	LSE	0.01090	0.02989	0.00254	0.02340
50	MLE	0.01084	0.02633	0.00190	0.01232
	LSE	0.01090	0.02105	0.00193	0.00972
75	MLE	0.01066	0.02362	0.00147	0.00744
	LSE	0.01080	0.01983	0.00158	0.00634
100	MLE	0.01052	0.02219	0.00134	0.00546
	LSE	0.01067	0.01918	0.00143	0.00478

Table 23.4. Known Ratio

n	Methods	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{MSE}(\hat{\theta}_1)$	$\text{MSE}(\hat{\theta}_2)$
25	MLE	0.01047	0.02093	0.00226	0.00452
	LSE	0.00914	0.01828	0.00228	0.00456
50	MLE	0.01025	0.02050	0.00151	0.00302
	LSE	0.00941	0.01882	0.00162	0.00323
75	MLE	0.01016	0.02032	0.00120	0.00240
	LSE	0.00941	0.01882	0.00132	0.00265
100	MLE	0.01013	0.02026	0.001102	0.00205
	LSE	0.00960	0.01919	0.00113	0.00227

Table 23.5. $\tau_1 = 100$, $\tau_2 = 150$, $\theta_1 = 1/100$, $\theta_2 = 1/25$

n	Methods	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{MSE}(\hat{\theta}_1)$	$\text{MSE}(\hat{\theta}_2)$
25	MLE	0.01086	0.06454	0.00273	0.03688
	LSE	0.01029	0.05022	0.00256	0.02597
50	MLE	0.01040	0.05234	0.00194	0.02020
	LSE	0.01028	0.04411	0.00198	0.01578
75	MLE	0.01031	0.04959	0.00151	0.01569
	LSE	0.01029	0.04337	0.00161	0.01262
100	MLE	0.01021	0.04728	0.00128	0.01256
	LSE	0.01023	0.04242	0.00137	0.01078

Table 23.6. Known ratio

n	Methods	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{MSE}(\hat{\theta}_1)$	$\text{MSE}(\hat{\theta}_2)$
25	MLE	0.01041	0.04164	0.00227	0.00910
	LSE	0.00907	0.03630	0.00231	0.00923
50	MLE	0.01018	0.04074	0.00149	0.00594
	LSE	0.00934	0.03737	0.00162	0.00648
75	MLE	0.01011	0.04045	0.00115	0.00460
	LSE	0.00946	0.03783	0.00131	0.00523
100	MLE	0.01005	0.04021	0.00100	0.00398
	LSE	0.00952	0.03810	0.00115	0.00461

Table 23.7. $\tau_1 = 100$, $\tau_2 = 125$, $\theta_1 = 1/100$, $\theta_2 = 1/50$

n	Methods	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{MSE}(\hat{\theta}_1)$	$\text{MSE}(\hat{\theta}_2)$
25	MLE	0.01088	0.03175	0.00264	0.01652
	LSE	0.01016	0.02565	0.00250	0.01101
50	MLE	0.01060	0.02669	0.00195	0.00949
	LSE	0.01029	0.02343	0.00194	0.00682
75	MLE	0.01035	0.02491	0.00151	0.00720
	LSE	0.01015	0.02266	0.00160	0.00547
100	MLE	0.01028	0.02386	0.00136	0.00558
	LSE	0.01013	0.02215	0.00143	0.00441

Table 23.8. Known ratio

n	Methods	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{MSE}(\hat{\theta}_1)$	$\text{MSE}(\hat{\theta}_2)$
25	MLE	0.01046	0.02091	0.00223	0.00447
	LSE	0.00906	0.01812	0.00224	0.00447
50	MLE	0.01026	0.02052	0.00151	0.00302
	LSE	0.00940	0.01879	0.00163	0.00326
75	MLE	0.01019	0.02038	0.00123	0.00245
	LSE	0.00954	0.01908	0.00136	0.00271
100	MLE	0.01015	0.02029	0.00105	0.00210
	LSE	0.00962	0.01923	0.00118	0.00236

Table 23.9. $\tau_1 = 100$, $\tau_2 = 125$, $\theta_1 = 1/100$, $\theta_2 = 1/25$

n	Methods	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{MSE}(\hat{\theta}_1)$	$\text{MSE}(\hat{\theta}_2)$
25	MLE	0.01072	0.04723	0.00254	0.02006
	LSE	0.01012	0.03806	0.00234	0.01742
50	MLE	0.01063	0.04355	0.00198	0.01248
	LSE	0.01052	0.03670	0.00197	0.01158
75	MLE	0.01042	0.04268	0.00147	0.01025
	LSE	0.01045	0.03731	0.00158	0.01009
100	MLE	0.01029	0.04180	0.00132	0.00839
	LSE	0.01039	0.03720	0.00144	0.00899

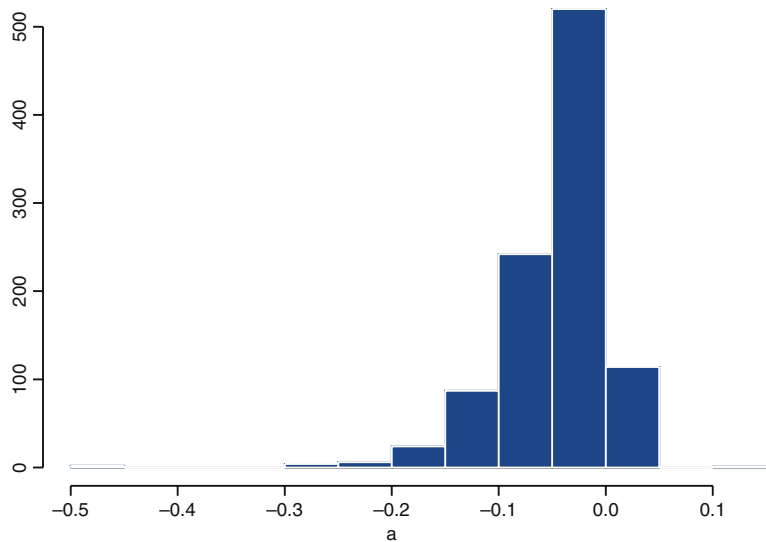


Figure 23.3. Sampling distribution of the MLE of a

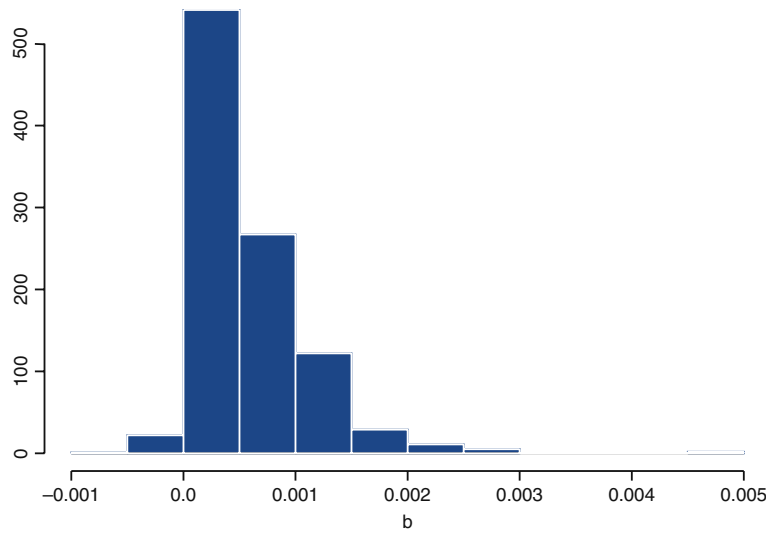


Figure 23.4. Sampling distribution of the MLE of b

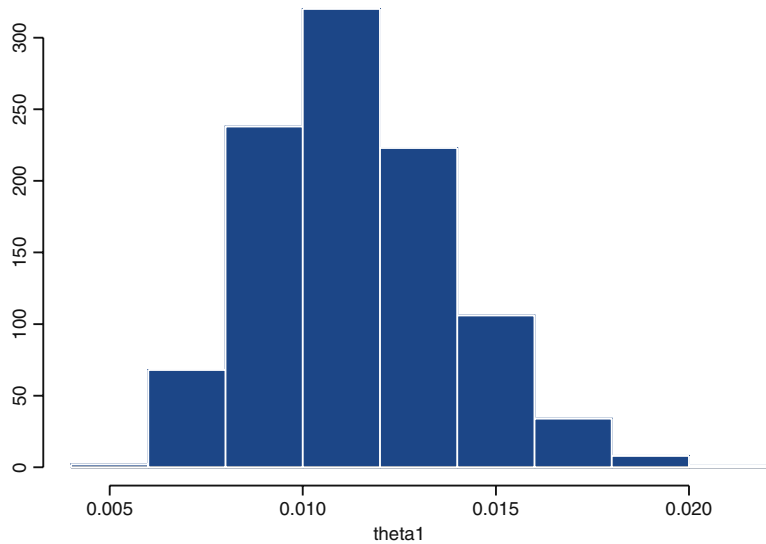


Figure 23.5. Sampling distribution of the MLE of θ_1

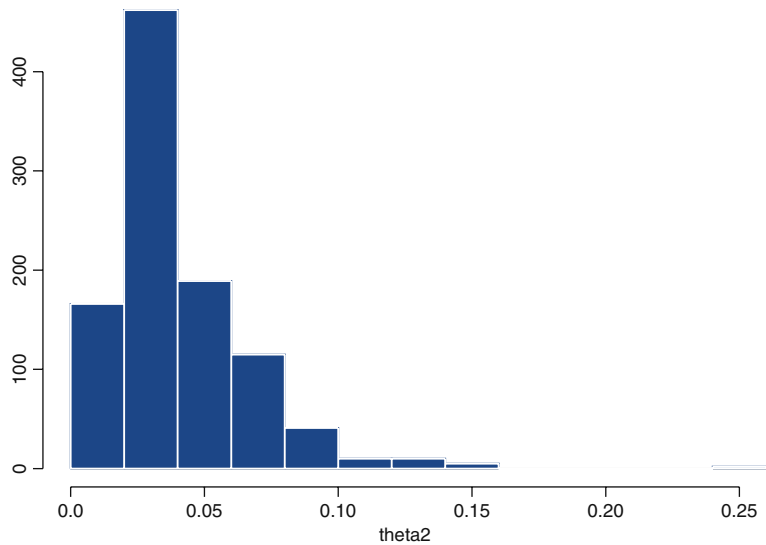


Figure 23.6. Sampling distribution of the MLE of θ_2

Table 23.10. Known ratio

n	Methods	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{MSE}(\hat{\theta}_1)$	$\text{MSE}(\hat{\theta}_2)$
25	MLE	0.01039	0.04157	0.00226	0.00903
	LSE	0.00908	0.03631	0.00231	0.00923
50	MLE	0.01023	0.04090	0.00158	0.00633
	LSE	0.00936	0.03745	0.00171	0.00684
75	MLE	0.01016	0.04065	0.00126	0.00502
	LSE	0.00950	0.03798	0.00139	0.00556
100	MLE	0.01011	0.04043	0.00104	0.00416
	LSE	0.00956	0.03823	0.00120	0.00480

From the tables, we can observe that, as expected, the average bias and the MSEs decrease as the sample size increases. When the ratio $\frac{\theta_2}{\theta_1}$ is known the performance of the estimators is better. The RMSE for the estimator of $\hat{\theta}_2$ for a known ration of 4 is twice the RMSE of $\hat{\theta}_2$ for a known ration of 2. This can be seen from the form of the estimators. When the sample size is small, the LSE outperforms the MLE, but as the sample size increases, the MLE clearly is superior. For fixed τ_1 , if τ_2 increases, the estimates of θ_2 are not as good, although the performance of the estimator of θ_1 does not change significantly. This is intuitive since, as τ_2 increases, the number of failures n_3 will decrease. For fixed τ_1 , τ_2 , and θ_1 , when θ_2 increases the performances of the estimators deteriorate.

23.7 Conclusions

This new model for a step-stress experiment was motivated by the limitations of Nelson's CEM. Our formulation in terms of a piecewise hazard function provides a more realistic model for biomedical applications and reduces to the standard CEM of Nelson in the limiting case.

We have developed both likelihood estimators as well as estimators based on least squares. The performance of these estimators has been evaluated using a simulation study. The least-squares estimators, obtained in closed form, perform extremely well for small samples. In addition to the simulation results, we have used a real data set to fit the new model. The model is extremely flexible and provides an excellent fit to the fatigue data.

Appendix

The likelihood equations

Here we show that the MLE of b can be obtained as the solution of (23.7). Note that the log-likelihood function can be written from (23.6) as

$$L(a, b) = \ln l(a, b) = -aT - bK + n_1 \ln(a + b\tau_1) + \sum_{i \in I_2} \ln(a + bt_i) + n_3 \ln(a + b\tau_2).$$

Therefore, \hat{a} and \hat{b} can be obtained by solving the equations

$$\frac{\partial L}{\partial a} = 0 \quad \text{and} \quad \frac{\partial L}{\partial b} = 0. \quad (23.15)$$

From (23.15), we obtain

$$a \times \frac{\partial L}{\partial a} + b \frac{\partial L}{\partial b} = 0. \quad (23.16)$$

After simplification, we obtain

$$a = \frac{n - bK}{T}. \quad (23.17)$$

Substituting this expression of a from (23.17) in $\frac{\partial L}{\partial a} = 0$, we obtain the equation in (23.7).

Fisher information matrix

In this subsection, we obtain the explicit expressions for elements of the observed Fisher information matrix. We can use the asymptotic normality property of the MLEs to construct approximate confidence intervals of a and b for large n .

Let $O(a, b) = O_{ij}(a, b)$, $i, j = 1, 2$, denote the observed Fisher information matrix of a and b . Then

$$\begin{aligned} O_{11} &= \frac{n_1}{(\hat{a} + \hat{b}\tau_1)^2} + \sum_{i \in I_2} \frac{1}{(\hat{a} + \hat{b}t_i)^2} + \frac{n_3}{(\hat{a} + \hat{b}\tau_2)^2} \\ O_{22} &= \frac{n_1\tau_1^2}{(\hat{a} + \hat{b}\tau_1)^2} + \sum_{i \in I_2} \frac{t_i^2}{(\hat{a} + \hat{b}t_i)^2} + \frac{n_3\tau_2^2}{(\hat{a} + \hat{b}\tau_2)^2} \\ O_{12} &= \frac{n_1\tau_1}{(\hat{a} + \hat{b}\tau_1)^2} + \sum_{i \in I_2} \frac{t_i}{(\hat{a} + \hat{b}t_i)^2} + \frac{n_3\tau_2}{(\hat{a} + \hat{b}\tau_2)^2}. \end{aligned}$$

The approximate variances of \hat{a} and \hat{b} can then be obtained through the observed information matrix as

$$V_a = \frac{O_{22}}{O_{11}O_{22} - O_{12}^2} \quad \text{and} \quad V_b = \frac{O_{11}}{O_{11}O_{22} - O_{12}^2}. \quad (23.18)$$

Using the observed asymptotic variances of \hat{a} and \hat{b} , approximate confidence intervals can be easily constructed.

References

1. Bagdonavicius, V., Gerville-Reache, L., and Nikulin, M.S. (2002), Parametric inference for step-stress models, *IEEE Transactions on Reliability*, **51**, 27–31.

2. Bagdonavicius, V., and Nikulin, M.S. (2002), *Accelerated Life Models: Modeling and Statistical Analysis*, Chapman and Hall/CRC Press, Boca Raton, Florida.
3. Bai, D.S., Kim, M.S., and Lee, S.H. (1989), Optimum simple step-stress accelerated life test with censoring, *IEEE Transactions on Reliability*, **38**, 528–532.
4. Balakrishnan, N. (2009), A synthesis of exact inferential results for exponential step-stress models and associated optimal accelerated left-tests, *Metrika*, **69**, 351–396.
5. Balakrishnan, N., Kundu, D., Ng, H.K.T., and Kannan, N. (2007), Point and interval estimation for a simple step-stress model with Type-II censoring, *Journal of Quality and Technology*, **9**, 35–47.
6. DeGroot, M. and Goel, P.K. (1979), Bayesian estimation and optimal design in partially accelerated life testing, *Naval Research Logistics Quarterly*, **26**, 223–235.
7. Greven, S., Bailer, J., Kupper, L.L., Muller, K.E., and Craft, J.L. (2004). A parametric model for studying organism fitness using step-stress experiments, *Biometrics*, **60**, 793–799.
8. Gouno, E. and Balakrishnan, N. (2001), Step-stress accelerated life test, In *Handbook of Statistics*, Vol. 20, Editors, Balakrishnan, N., and Rao, C.R., North-Holland, Amsterdam, The Netherlands.
9. Gouno, E., Sen, A., and Balakrishnan, N. (2004), Optimal sept-stress test under progressive Type-I censoring, *IEEE Transactions on Reliability*, **53**, 383–393.
10. Han, D., Balakrishnan, N., Sen, A., and Gouno, E. (2006), Corrections on “Optimal stpe-stress test under progressive Type-I censoring,” *IEEE Transactions on Reliability*, **55**, 613–614.
11. Johnson, N.L., Kotz, S., and Balakrishnan, N. (1994), *Continuous Univariate Distributions – Vol. 1*, 2nd edition, John Wiley & Sons, New York.
12. Khamis, I.H. and Higgins, J.J. (1998), A new model for step-stress testing, *IEEE Transactions on Reliability*, **47**, 131–134.
13. Meeker, W.Q. and Escobar, L.A. (1998), *Statistical Methods for Reliability Data*, John Wiley & Sons, New York.
14. Miller, R. and Nelson, W.B. (1983), Optimum simple step-stress plans for accelerated life testing, *IEEE Transactions on Reliability*, **32**, 59–65.
15. Nelson, W.B. (1980), Accelerated life testing: step-stress models and data analysis, *IEEE Transactions on Reliability*, **29**, 103–108.
16. Nelson, W.B. (1990), *Accelerated Life Testing, Statistical Models, Test Plans and Data Analysis*, John Wiley & Sons, New York.
17. Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1992), *Numerical Recipes in FORTRAN, The Art of Scientific Computing*, 2nd edition, Cambridge University, Cambridge, UK.
18. Swain, J., Venkatraman, S., and Wilson, J. (1988), Least squares estimation of distribution function of Johnson’s translation system, *Journal of Statistical Computation and Simulation*, **29**, 271–297.
19. Xiong, C. (1998), Inferences on a simple step-stress model with type-II censored exponential data, *IEEE Transactions on Reliability*, **47**, 142–146.
20. Xiong, C. and Milliken, G.A. (1999), Step-stress life-testing with random change times for exponential data, *IEEE Transactions on Reliability*, **48**, 141–148.

Estimation of Density on Censored Data

V. Solev

Steklov Mathematical Institute, St. Petersburg, Russia

Abstract: We investigate the problem of approximation of MLE as we need to estimate unknown density on censored data. What the accuracy of such approximation should be if we need to save the same accuracy of estimating?

Keywords and phrases: Estimation of density, censored data, Hellinger consistency

24.1 Introduction

Let X be a random variable with the distribution $P \in \mathcal{P}$ and density $f = f(P) \in \mathcal{F}$. Let Y be a random variable independent of X with the distribution $Q \in \mathcal{Q}$ and density $g = g(Q) \in \mathcal{G}$. We shall use notations $F = F^f, G = G^g$ for distribution functions with densities f and g , respectively.

Censoring consists in the following: instead of X, Y we observe the random vector $W = (Z, \delta)$,

$$Z = X \wedge Y, \quad \delta = \mathbf{1}_{X_j \leq Y_j}. \quad (24.1)$$

The distribution P^W of random vector W has density $p(z, t)$,

$$p(z, t) = [f(z)(1 - G(z))]^t \times [g(z)(1 - F(z))]^{1-t}, \quad (24.2)$$

with respect to the measure $\mu(dz \times dt) = dz \times \nu(dt)$. Here $\nu(dt)$ is the counting measure on $\{0, 1\}$. The distribution of random variable Z has density h

$$r(z) = f(z)(1 - G(z)) + g(z)(1 - F(z)) \quad (24.3)$$

and distribution function $R(z)$,

$$1 - R(z) = (1 - F(z))(1 - G(z)). \quad (24.4)$$

Further we shall assume that random variables X, Y are nonnegative.

The problem consists in the estimating of unknown density $f \in \mathcal{F}$ on the sample $W_1 = (Z_1, \delta_1), \dots, W_n = (Z_n, \delta_n)$ with common density $p(z, t)$.

Kaplan and Meier [3] suggested to construct estimators \hat{F}_n and \hat{G}_n for distribution functions F and G in the following way. Let H_n be the empirical distribution function for the sample Z_1, \dots, Z_n . Then

$$1 - R_n(z) = \left(1 - \hat{F}_n(z)\right) \left(1 - \hat{G}_n(z)\right) \text{ and}$$

$$\hat{F}_n(Z_j + 0) - \hat{F}_n(Z_j - 0) = 0, \text{ if } \delta_j = 0; \quad \hat{G}_n(Z_j + 0) - \hat{G}_n(Z_j - 0) = 0, \text{ if } \delta_j = 1.$$

It is clear that

$$\hat{F}_n(z) = 1 - \prod_{\substack{j: Z_j \leq z \\ 1 \leq j \leq n}} \left(1 - \frac{\delta_{j,n}}{n - j + 1}\right)$$

$$\hat{G}_n(z) = 1 - \prod_{\substack{j: Z_j \leq z \\ 1 \leq j \leq n}} \left(1 - \frac{1 - \delta_{j,n}}{n - j + 1}\right).$$

Here $Z_{1,n} \leq \dots \leq Z_{n,n}$ are the ordered Z_1, \dots, Z_n , and $\delta_{j,n}$ is the δ_k corresponding to $Z_{j,n} = Z_k$.

Bitouzé et al. [1] proved that there exists an absolute constant C such that for any $x > 0$

$$P \left\{ \sqrt{n} \|(1 - G) (\hat{F}_n - F)\|_\infty > x \right\} \leq 2, 5e^{-2x^2 + Cx}.$$

It is impossible to use this excellent result for statistical goals since function G is unknown.

We set

$$\mathcal{K} = \left\{ p(z, t) : p(z, t) = [f(z)(1 - G(z))]^t \times [g(z)(1 - F(z))]^{1-t}, (f, g) \in \mathcal{F} \times \mathcal{G} \right\}. \quad (24.5)$$

Consider the maximum likelihood estimator \hat{p}_n of the unknown density $p \in \mathcal{K}$:

$$\int \ln \frac{\hat{p}_n(z, t)}{p(z, t)} dP_n^W \geq 0, \quad p \in \mathcal{K}, \quad (24.6)$$

where dP_n^W is the empirical measure,

$$P_n^W \{A\} = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_A(W_j).$$

It is evident that

$$\hat{p}_n(z, t) = \left[\hat{f}_n(z)(1 - \hat{G}_n(z)) \right]^t \times \left[\hat{g}_n(z)(1 - \hat{F}_n(z)) \right]^{1-t}, \quad (24.7)$$

for some $\hat{f}_n \in \mathcal{F}$ and $\hat{g}_n \in \mathcal{G}$. Here \hat{F}_n and \hat{G}_n are distribution functions corresponding to densities $\hat{f}_n \in \mathcal{F}$ and $\hat{g}_n \in \mathcal{G}$.

Clearly, the relation (24.6) is equivalent to

$$\int \ln \frac{t\hat{f}_n(z) + (1-t)(1 - \hat{F}_n(z))}{tf(z) + (1-t)(1 - F(z))} dP_n^W \geq 0, \quad f \in \mathcal{F} \quad (24.8)$$

and

$$\int \ln \frac{(1-t)\widehat{g}_n(z) + t(1-\widehat{G}_n(z))}{(1-t)g(z) + t(1-G(z))} dP_n^W \geq 0, \quad g \in \mathcal{G}. \quad (24.9)$$

So, it is possible to estimate unknown densities f and g separately.

Let $H(P, Q)$ be the Hellinger distance between probability distributions P and Q ,

$$H^2(P, Q) = \frac{1}{2} \int \left(\sqrt{f} - \sqrt{q} \right)^2 d\vartheta = h^2(f, q),$$

where f and q are densities of probability distributions P and Q with respect to a measure ϑ . Clearly,

$$h^2(f, q) = 1 - \int \sqrt{fq} d\vartheta. \quad (24.10)$$

van de Geer [6] proved that

$$h^2(\widehat{p}_n, p) \leq \int_{p>0} \sqrt{\frac{\widehat{p}_n}{p}} d(P_n^W - P^W).$$

In order to use this result we need to estimate the probabilities

$$P \left\{ \sup_{p \in \mathcal{K}} \int_{p>0} \sqrt{\frac{\widehat{p}_n}{p}} d(P_n^W - P^W) > x \right\}.$$

Let $(\mathcal{X}, \mathcal{B})$ be a measurable space, μ be a σ -finite measure on \mathcal{B} . For a nonnegative function h , defined on \mathcal{X} , we denote by $V_\delta(h)$ the Hellinger ball of radius δ with center in the point h ,

$$V_\delta(h) = \left\{ f \geq 0 : \frac{1}{\sqrt{2}} \|\sqrt{f} - \sqrt{h}\|_{L_2(d\mu)} \leq \delta \right\}.$$

Here

$$\|\varphi\|_{L_2(d\mu)}^2 = \int_{\mathcal{X}} |\varphi|^2 d\mu.$$

Let \mathcal{D} be a set of densities (with respect to the measure μ) of probability distributions on $(\mathcal{X}, \mathcal{B})$. The covering number $N = N(\delta) = N(\delta, \mathcal{D})$ is the minimum number N such that

$$\mathcal{D} \subset \bigcup_{j=1}^N V_\delta(h_j).$$

The value $H(\delta) = H(\delta, \mathcal{D}) = \log N(\delta, \mathcal{D})$ is called by the Hellinger entropy of \mathcal{D} .

Now we define the bracketing Hellinger metric entropy of the set \mathcal{D} . For a pair of functions $0 \leq f^- \leq f^+$, we denote by $[f^-, f^+]$ the set

$$[f^-, f^+] = \{f : f^-(x) \leq f(x) \leq f^+(x)\}.$$

For a positive δ , let $N_b(\delta, \mathcal{D})$ be the minimum number N such that

$$\mathcal{D} \subset \bigcup_{j=1}^N [f_j^-, f_j^+], \quad \frac{1}{\sqrt{2}} \left\| \sqrt{f} - \sqrt{h} \right\|_{L_2(d\mu)} \leq \delta.$$

The value $H_b(\delta) = H_b(\delta, \mathcal{D}) = \log N_b(\delta, \mathcal{D})$ is called by the bracketing entropy of \mathcal{D} .

Suppose that the parametric set \mathcal{K} is totally bounded in Hellinger distance. It is well known that for regular case minimax rate ε_n (in Hellinger distance) correlated with entropy $H(\varepsilon, \mathcal{K})$ (see in details Van der Vaart and Welner [7]):

$$H(\varepsilon, \mathcal{K}) \asymp n\varepsilon_n^2. \quad (24.11)$$

Huber et al. [2] suppose that for $0 < \delta < 1$ and $\beta > 1$

$$H_b(\delta, \mathcal{F}) \leq C_1 \delta^{-\frac{1}{\beta}}, \text{ and } H_b(\delta, \mathcal{G}) \leq C_2 \delta^{-\frac{1}{\beta}}, \quad (24.12)$$

and proved under some regularity conditions that there exist positive constants $c_1 = c_1(\mathcal{K})$, $c_2 = c_2(\mathcal{K})$, $C = C(\mathcal{K})$ such that

$$\begin{aligned} \mathbb{P}_p \left\{ \int \left(1 - \widehat{G}_n\right) \left(\sqrt{\widehat{f}_n} - \sqrt{f}\right)^2 dx \geq c_1 \left(\frac{1}{n}\right)^{\frac{2\beta}{2\beta+1}} \right\} \\ \leq C \exp\{-c_2 n^{\frac{1}{2\beta+1}}\}. \end{aligned} \quad (24.13)$$

24.2 Approximating of Parametric Set

Here we consider the case as we approximate the parametric set \mathcal{K} by parametric set $\mathcal{K}(n)$. What the accuracy of such approximation should be if we need to save the same accuracy of estimating? We shall use the approach suggested in Wong and Shen [8] (see also Shen [5]).

At the beginning we formulate the following well-known lemma (see Massart [4]).

Lemma 1. *Let X_1, \dots, X_n be independent random elements with common density p with respect to a measure ϑ , P_n be corresponding empirical measure. Then, for any density q of a probability distribution and any positive x*

$$\mathbb{P}_p \left\{ \int \ln \frac{q}{p} dP_n \geq -2h^2(p, q) + 2\frac{x}{n} \right\} \leq e^{-x}. \quad (24.14)$$

Here we use notation \mathbb{P}_p for distribution of random vector (X_1, \dots, X_n) as X_1, \dots, X_n are independent random elements with common density p . In the case, when $x = \frac{1}{2} nh^2(p, q)$, we obtain

$$\mathbb{P}_p \left\{ \int \ln \frac{q}{p} dP_n \geq -\frac{1}{2} h^2(p, q) \right\} \leq e^{-\frac{1}{2} nh^2(p, q)}. \quad (24.15)$$

In other words (see Wong and Shen [8])

$$\mathbb{P}_p \left\{ \prod_{j=1}^n \frac{q}{p}(X_j) \geq e^{-\frac{1}{2} nh^2(p, q)} \right\} \leq e^{-\frac{1}{2} nh^2(p, q)}. \quad (24.16)$$

We consider the family $\{\mathcal{K}(n), n = 1, 2, \dots\}$ of approximating space $\mathcal{K}(n)$ such that

$$\mathcal{K} = \bigcup_{n=1}^{\infty} \mathcal{K}(n), \quad (24.17)$$

where

$$\mathcal{K}(n) \subset \mathcal{K}(n+1), \quad n = 1, 2, \dots \quad (24.18)$$

We plan to investigate the asymptotic behavior of the estimator \hat{p}_n of the unknown density $p \in \mathcal{K}$ defined by

$$\int \ln \frac{\hat{p}_n(z, t)}{q(z, t)} dP_n^W \geq 0, \quad q \in \mathcal{K}(n). \quad (24.19)$$

It is evident that for $c > 0$

$$\mathbb{P}_p \{h(\hat{p}_n, p) \geq \varepsilon\} \leq \mathbb{P}_p \left\{ \sup_{\substack{q \in \mathcal{K}(n) \\ h(q, p) \geq \varepsilon}} \int \ln \frac{q(z, t)}{p(z, t)} dP_n^W \geq -c \right\}. \quad (24.20)$$

So, we need to estimate the probabilities

$$\begin{aligned} & \mathbb{P}_p \left\{ \sup_{\substack{q \in \mathcal{K}(n) \\ h(q, p) \geq \varepsilon}} \int \ln \frac{q(z, t)}{p(z, t)} dP_n^W \geq -c \right\} = \\ & = \mathbb{P}_p \left\{ \sup_{\substack{q \in \mathcal{K}(n) \\ h(q, p) \geq \varepsilon}} \prod_{j=1}^n \frac{q(Z_j, \delta_j)}{p(Z_j, \delta_j)} \geq e^{-cn} \right\}. \end{aligned} \quad (24.21)$$

We assume that there exists positive constant C such that for any $\delta > 0$ and some $\beta > 1$

$$H_b(\delta, \mathcal{K}(n)) \leq C\delta^{-\frac{1}{\beta}}. \quad (24.22)$$

The accuracy of approximation of the parametric set \mathcal{K} by the approximating space $\mathcal{K}(n)$ is defined by

$$\sup_{p \in \mathcal{K}} \inf_{q \in \mathcal{K}(n)} h^2(p, q) \leq C_1 n^{-\frac{2\beta}{2\beta+1}}. \quad (24.23)$$

It is possible to prove (see in details Wong and Shen [8]) that under conditions (24.22) and (24.23) there exist positive constants c_1, c_2, C_* such that

$$\mathbb{P}_p \left\{ \sup_{\substack{q \in \mathcal{K}(n) \\ h(q, p) \geq \varepsilon}} \int \ln \frac{q(z, t)}{p(z, t)} dP_n^W \geq -c_2 \varepsilon^2 \right\} \leq C_* e^{-nc_1 \varepsilon^2}. \quad (24.24)$$

24.3 Hellinger Distance

The following proposition is evident

Lemma 2. *Let f, g be densities of probability distributions in R^1 with distribution functions F, G . Then the function*

$$\begin{aligned} p(z, t) &= [f(z)(1 - G(z))]^t \times [g(z)(1 - F(z))]^{1-t} = \\ &= t [f(z)(1 - G(z))] + (1 - t) [g(z)(1 - F(z))] \end{aligned} \quad (24.25)$$

is density of probability distributions in $\mathcal{X} = R^1 \times \{0, 1\}$ with respect to the measure $\mu(dz \times dt) = dz \times \nu(dt)$. Here $\nu(dt)$ is the counting measure on $\{0, 1\}$.

Now we suppose that we have two of functions

$$p_j(z, t) = [f_j(z)(1 - G_j(z))]^t \times [g_j(z)(1 - F_j(z))]^{1-t}, \quad j = 1, 2.$$

Here f_j, g_j are densities of probability distributions in R^1 with distribution functions F_j, G_j . We plan to compare the Hellinger distance $h^2(p_1, p_2)$,

$$\begin{aligned} h^2(p_1, p_2) &= \\ &= 1 - \int \left[t \sqrt{f_1 f_2} \sqrt{(1 - G_1)(1 - G_2)} + (1 - t) \sqrt{g_1 g_2} \sqrt{(1 - F_1)(1 - F_2)} \right] d\mu, \end{aligned}$$

and the value

$$\int (1 - G_2(z)) \left(\sqrt{f_1(z)} - \sqrt{f_2(z)} \right)^2 dz.$$

Namely we plan to prove the following proposition.

Lemma 3. *The following inequality is valid*

$$h^2(p_1, p_2) \geq \frac{1}{4} \int (1 - G_2(z)) \left(\sqrt{f_1(z)} - \sqrt{f_2(z)} \right)^2 dz. \quad (24.26)$$

PROOF. Since the function $g = (g_1 + g_2)/2$ is density of the probability distribution with distribution function $G = (G_1 + G_2)/2$, then (see lemma 1) functions

$$\begin{aligned} p_1^*(z, t) &= [f_1(z)(1 - G(z))]^t \times [g(z)(1 - F_1(z))]^{1-t}, \\ p_2^*(z, t) &= [f_2(z)(1 - G(z))]^t \times [g(z)(1 - F_2(z))]^{1-t} \end{aligned} \quad (24.27)$$

are densities of probability distributions in $\mathcal{X} = R^1 \times \{0, 1\}$ with respect to the measure $\mu(dz \times dt) = dz \times \nu(dt)$.

From inequalities

$$g = (g_1 + g_2)/2 \geq \sqrt{g_1 g_2}, \quad 1 - G = ((1 - G_1) + (1 - G_2))/2 \geq \sqrt{(1 - G_1)(1 - G_2)}$$

we deduce that

$$\begin{aligned}\sqrt{p_1^* p_2^*} &= t\sqrt{f_1 f_2} (1 - G) + (1 - t)g \sqrt{(1 - F_1)(1 - F_2)} \geq \\ &\geq t\sqrt{f_1 f_2} \sqrt{(1 - G_1)(1 - G_2)} + (1 - t)\sqrt{g_1 g_2} \sqrt{(1 - F_1)(1 - F_2)} = \sqrt{p_1 p_2}.\end{aligned}$$

Therefore,

$$h^2(p_1^*, p_2^*) = 1 - \int \sqrt{p_1^* p_2^*} d\mu \leq 1 - \int \sqrt{p_1 p_2} d\mu = h^2(p_1, p_2). \quad (24.28)$$

Notice that

$$\begin{aligned}h^2(p_1^*, p_2^*) &= \frac{1}{2} \int \left(\sqrt{p_1^*} - \sqrt{p_2^*} \right)^2 d\mu \geq \frac{1}{2} \int (1 - G(z)) \left(\sqrt{f_1(z)} - \sqrt{f_2(z)} \right)^2 dz \geq \\ &\geq \frac{1}{4} \int (1 - G_2(z)) \left(\sqrt{f_1(z)} - \sqrt{f_2(z)} \right)^2 dz.\end{aligned}$$

Thus, from (24.28) we obtain

$$h^2(p_1, p_2) \geq \frac{1}{4} \int (1 - G_2(z)) \left(\sqrt{f_1(z)} - \sqrt{f_2(z)} \right)^2 dz.$$

So, we obtain the inequality (24.25). ■

We shall use the inequality (24.25) in the following situation.

Lemma 4. *Let $\hat{p}_n(z, t)$ be an estimator,*

$$\hat{p}_n(z, t) = \left[\hat{f}_n(z)(1 - \hat{G}_n(z)) \right]^t \times \left[\hat{g}_n(z)(1 - \hat{F}_n(z)) \right]^{1-t},$$

$p(z, t)$ be the true density,

$$p(z, t) = [f(z)(1 - G(z))]^t \times [g(z)(1 - F(z))]^{1-t},$$

then

$$h^2(\hat{p}_n, p) \geq \frac{1}{4} \int \left(1 - \hat{G}_n(z) \right) \left(\sqrt{\hat{f}_n(z)} - \sqrt{f(z)} \right)^2 dz. \quad (24.29)$$

24.4 Main Result

Let X be a random variable with density $f \in \mathcal{F}$ and distribution function F ; Y be independent of X random variable with density $g = g(Q) \in \mathcal{G}$ and distribution function F . Instead of X, Y we observe the random vector $W = (Z, \delta)$,

$$Z = X \wedge Y, \quad \delta = \mathbf{1}_{X_j \leq Y_j}, \quad (24.30)$$

with density $p(z, t)$,

$$p(z, t) = [f(z)(1 - G(z))]^t \times [g(z)(1 - F(z))]^{1-t}, \quad (24.31)$$

with respect to the measure $\mu(dz \times dt) = dz \times \nu(dt)$. Here $\nu(dt)$ is the counting measure on $\{0, 1\}$. The parametric set \mathcal{K} is defined by

$$\mathcal{K} = \left\{ p(z, t) : p(z, t) = [f(z)(1 - G(z))]^t \times [g(z)(1 - F(z))]^{1-t}, (f, g) \in \mathcal{F} \times \mathcal{G} \right\}. \quad (24.32)$$

The parametric set \mathcal{K} is approximated by the set $\mathcal{K}(n)$ such that there exists positive constant C such that, for any $\delta > 0$ and some $\beta > 1$

$$H_b(\delta, \mathcal{K}(n)) \leq C\delta^{-\frac{1}{\beta}}. \quad (24.33)$$

We assume that

$$\mathcal{K}(n) = \left\{ p(z, t) : p(z, t) = [f(z)(1 - G(z))]^t \times [g(z)(1 - F(z))]^{1-t}, (f, g) \in \mathcal{F}(n) \times \mathcal{G}(n) \right\}. \quad (24.34)$$

The accuracy of approximation of the parametric set \mathcal{K} by the approximating space $\mathcal{K}(n)$ satisfies the condition

$$\sup_{p \in \mathcal{K}} \inf_{q \in \mathcal{K}(n)} h^2(p, q) \leq C_1 n^{-\frac{2\beta}{2\beta+1}}. \quad (24.35)$$

The estimator \hat{p}_n belongs to $\mathcal{K}(n)$ and is defined by

$$\int \ln \frac{\hat{p}_n(z, t)}{q(z, t)} dP_n^W \geq 0, \quad q \in \mathcal{K}(n). \quad (24.36)$$

The estimator \hat{p}_n can be represented in the form

$$\hat{p}_n(z, t) = [\hat{f}_n(z)(1 - \hat{G}_n(z))]^t \times [\hat{g}_n(z)(1 - \hat{F}_n(z))]^{1-t}, \quad (24.37)$$

where $\hat{f}_n \in \mathcal{F}(n)$ and $\hat{g}_n \in \mathcal{G}(n)$.

Theorem 1. Under conditions (24.33) and (24.33) there exist such positive constants c_1, c_2, C such that

$$\begin{aligned} \mathbb{P}_p \left\{ \int \left(1 - \hat{G}_n \right) \left(\sqrt{\hat{f}_n} - \sqrt{f} \right)^2 dx \geq c_1 \left(\frac{1}{n} \right)^{\frac{2\beta}{2\beta+1}} \right\} \\ \leq C \exp \left\{ -c_2 n^{\frac{1}{2\beta+1}} \right\}. \end{aligned} \quad (24.38)$$

Here density $p(z, t)$ is defined by (24.31) and belongs to \mathcal{K} .

PROOF. It follows from (24.24) and (24.20) that for appropriately chosen ε , $\varepsilon = c_0 n^{-\frac{\beta}{2\beta+1}}$,

$$\mathbb{P}_p \{h(\hat{p}_n, p) \geq \varepsilon\} \leq C_* e^{-nc_1 \varepsilon^2}. \quad (24.39)$$

Thus, we can obtain (24.38) from (24.29). ■

Acknowledgment. The investigation was supported in part by grants RFBR 08-01-00692, NSH 638.2008.1., and RFBR-DFG 09-01-91331.

References

1. Bitouzé, D., Laurent, B., and Massart, P. (1999). A Dvoretzky-Keifer-Wolfowitz type inequality for the Kaplan-Meier estimator, *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, **35**, 735–763.
2. Huber, C., Solev, V., and Vonta, F. (2007). Maximum likelihood estimator: the nonparametric approach. In: *Probability and Statistics, 11*. I. Ibragimov and V. Sudakov (eds.) pp. 220–228, Steklov Institute of Mathematics, St. Petersburg RAS.
3. Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.
4. Massart, P. (2003). *Concentration Inequality and Model Selection*, Springer-Verlag, New York.
5. Shen, N.X. (1997). On methods sieves and penalization, *The Annals of Statistics*, **2**, 2555–2591.
6. Van de Geer, S. (1993). Hellinger-consistency of certain nonparametric maximum likelihood estimators, *The Annals of Statistics*, **21**, 14–44.
7. Van der Vaart, A.W. and Welner, J.A. (1996). *Weak Convergence and Empirical Processes*, Springer-Verlag, New York.
8. Wong, W.H. and Shen, N.X. (1995). Probability inequalities and convergence rates of sieve mles, *The Annals of Statistics*, **2**, 339–362.

Competing Risk and Chaotic Systems

Toward a Test for Departure of a Trajectory from a Neighborhood of a Chaotic System

M. LuValle

OFS Laboratories, Somerset, NJ

Abstract: Many real systems have been identified as potentially chaotic (e.g., trains of heart beats, records of multiple populations in an ecological system, weather and climate data). In such cases, when “healthy” is defined by the measured trajectory being confined to a “strange attractor” in a “phase space,” how can we define and measure degradation? The simplest course is to define degradation as when the attractor changes. The earliest warning of such a change would be when a trajectory leaves an attractor, without an external forcing, but there seems little written on the problem of identifying whether a short trajectory belongs to a given attractor (at least to within a neighborhood defined by stochastic disturbances and uncertainty in the parameters of the defining equations).

This chapter explores the problem of testing short trajectories in the neighborhood of the Lorenz attractor. Although there are some conditions necessary for developing a working test, it turns out that the Lorenz attractor is perhaps one of the kindest examples to explore, being easily visualized as well as having significant work done on its statistical properties. Further work with more complex attractors is necessary before this approach can be considered general. However, the proposed test shows the potential of not only detecting degradation, but of being a general partial goodness of fit test for a short trajectory to a neighborhood of an attractor, even for high-dimensional attractors. In addition, the construction of the test suggests a dimension reduction scheme useful for exploring deficiencies in the attractor in describing the trajectory.

Keywords and phrases: Deterministic chaos, false discover rate, goodness of fit test, short trajectory

25.1 Introduction

Many natural systems have been potentially identified as chaotic, in particular, traces of heartbeats [8], records of multiple populations in ecological systems [9], and weather and climate systems [28]. What defines a system as chaotic is that its long-term behavior is confined to a compact manifold within some (possibly infinite-dimensional) phase

space, but (confined within the manifold), starting at any time at any two points a small distance ε apart, the distance between the two points as they evolve with time grows exponentially.

Much of the literature on statistical inference methods applied to chaos has to do with

1. identifying whether a system is chaotic, random, or a mixture of the two, as in, [13, 17, 2, 19, 12], and [7].
2. estimating invariant aspects of the attractor, such as the Lyapunov exponent, embedding dimension, and Hausdorff dimension (defined for example in [5]). Papers attacking this problem include [13] and [19]. Perhaps one of the most interesting results is that even though the attractor may be defined in a m -dimensional space, it is possible to construct a “pseudo attractor” by taking data from as few as one margin over time, and choose $2m + 1$ delays T_1, \dots, T_{2m+1} so that the pseudo attractor has the same Lyapunov exponent. An interesting approach to the choice of delays and (when available) which dimensions to use is developed in [18].

In this chapter we consider a different problem. In particular we assume that the system under study is known to be chaotic, and the dynamics are known up to some small error in the parameters describing the dynamics. The question is can we detect a departure from said dynamics in observing a short trajectory with noise. The short answer, at least with the Lorenz attractor, is yes. The approach is particularly intuitive to the statistician. We observe a short trajectory, with quantified statistical measurement error. We ‘spin up’ (calculate in the original phase space) several versions of the attractor in the neighborhood of uncertainty of the nominal parameters we believe govern the dynamics (the problem of adding in stochastic disturbances is left for another chapter). The initial conditions are set by an early point in the observed trajectory. Then we take a later section in the observed trajectory and project each point onto the set of points defined by the collection of calculated attractors (projection is to the nearest point). We take the difference of the point and the projection and test the component of the difference orthogonal to the approximate manifold described by the attractors to see if it is significantly different from 0. We do this for a number of points in the trajectory and apply false discovery rate theory to see if the trajectory appears significantly separated from the group of manifolds defined by the spun-up attractors. So this looks like a local version of the kind of orthogonal projection and testing familiar from linear models. There are, however, a few differences arising from the need to define a distribution under the null hypothesis that the system is correctly specified.

- The null distribution has contributions from the measurement error, the sampling of the neighborhood of the assumed parameters in generating the collection of attractors, and the underlying chaos in the system. There is some hope that the ergodics of the chaotic system [18, 1, 16] result in a null distribution with good properties. Interestingly, with the current setup the null hypothesis apparently matches a χ^2 distribution with 1 degree of freedom very closely. However, an alternative method of determining the null distribution is also included because this specific result is strongly dependent on the nature of the attractor and even more on the measurement error.
- The specification of the starting point, the size of the neighborhood around the assumed driving parameters, and the relative size of the measurement error have significant effect on the test. Initial attempts to construct this test using random

starting points for the null attractor set reliably resulted in false rejection of the null hypothesis with too high a proportion of spun-up attractors. Fixing the starting point as an early part of the trajectory to be tested helped this somewhat, but even then, with relatively low levels of measurement noise, the probability of an ‘extreme’ null attractor being created was too large. Significant bifurcation has been observed in the Lorenz system [25], though with parameter values far from where the sampling done in this chapter takes place.

- The starting point specification is important also from the point of view that chaotic systems may have pseudo-attractors which are apparently stable for long periods [5], but which then devolve to the true attractors. This may be some of the reasons for the large departures observed in the initial simulations.

We consider both the cases in which the full phase space is observed for the whole trajectory and when the only observation in the full phase space is the point setting the initial value for the trajectory. In the latter case, the comparison has to be made with respect to reconstructed attractors. Both seem to offer reasonable approaches to testing for the example considered. An interesting problem is constructing a starting point using only a reconstructed attractor. From Taken’s theorem [27] there is an invertible continuous mapping from the original phase space to the reconstruction space when the reconstruction is done properly, so in principle, the starting point can be defined from the trajectory in a reconstructed space. Demonstration of this is left as an applied research problem.

In the remainder of this section, we provide a short introduction to the terminology of chaos and to the particulars of the Lorenz attractor and the alternative we will consider. In the next section I introduce the test statistics and state the conjectures that support this work. Indications will be given of the proofs, along with emphasis of what unproven conjectures remain. In the third section, results from the simulated example are given, and in the fourth section directions for further work are explored.

25.1.1 Terminology and the Lorenz attractor

Terminology will follow that used in [5]. An attractor in dynamic systems is a set of points such that the long-term behavior of the dynamic system falls on that set of points as long as the initial conditions fall within its basin of attraction. A fixed point is an example of an attractor. A strange attractor is such that (locally) the attractor has a sensitive dependence on initial conditions, so a small variation in initial conditions (from x to $x + \varepsilon$) grows exponentially with time. In other words the strange attractor is a compact manifold representing the “stable states” of the solution (stable in the sense that points are not reached outside the orbits in the manifold). However, this collection of stable states is in no way periodic or quasi-periodic. Often such attractors have some fractal nature including fractional box dimension.

An example of a simple nonlinear dynamic system, with a three-dimensional phase space and a non-trivial strange attractor, is the Lorenz system [15]. It is pictured in Figure 25.1 in a particularly useful projection and is defined by the system of Equation (25.1).

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -\sigma x_1 + \sigma x_2 \\ -x_1 x_3 + r x_1 - x_2 \\ x_1 x_2 - b x_3 \end{pmatrix} \quad (25.1)$$

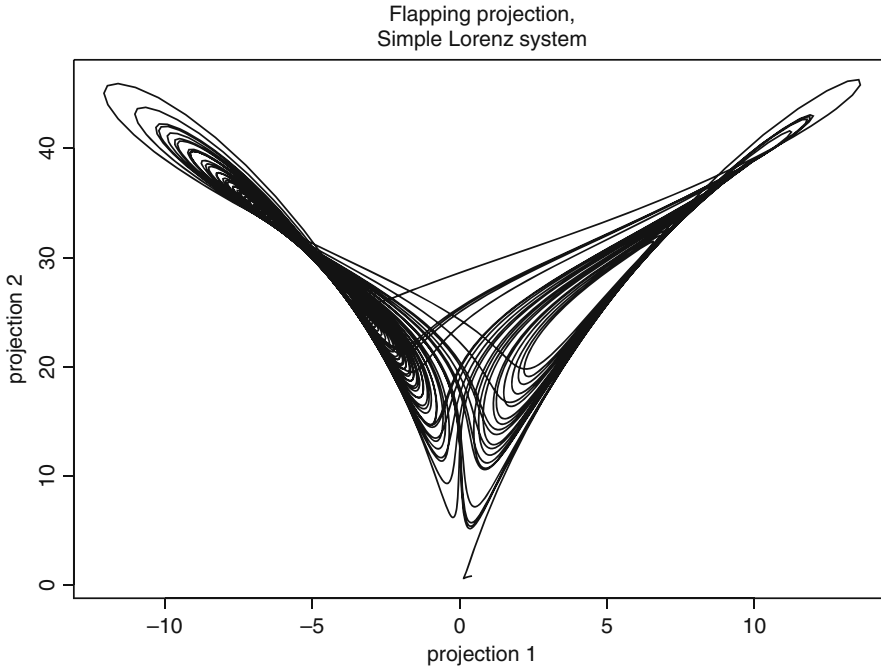


Figure 25.1. Revealing projection for Lorenz attractor

The Lorenz equations were derived from a system of equations studying convection in a layer of uniform depth when the temperature between the upper and lower surfaces was maintained constant with the lower surface having the higher temperature developed by Lord Raleigh [20]. Saltzmann [22], subsequently reduced Raleigh's system of partial differential equations to a system of ordinary differential equations via double Fourier expansion and a reduction approach proposed previously by Lorenz [14]. The solution of these equations result in only three dependent variables not having 0 as a fixed point. Subsequently Lorenz [15] reduced Saltzmann's set of equations to a set of 3 (given above) with the same long-term solutions.

The solutions to various forms of Equation (25.1) presented in this chapter were calculated by combining a symbolic operator exponentiation, such as described in Section 1–4 [6], with a recursive method of reducing the interval of the expansion and comparing with a procedure relaxing the estimated derivatives of the individual points, to see if the relaxation was self-consistent on the interval. The solution of Equation (25.1) with parameters (10,28,8/3) over the time interval of 0–500 with intervals of 0.1 is given in Figure 25.1. The view is orthogonal to the vector $(x, y, z)=(0.4,0.9,0.1)$.

25.1.2 The alternative for the simulation

The alternative model we consider for this study has the form

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -\sigma x_1 + \sigma x_2 \\ -x_1 x_3 + r x_1 - x_2 \\ x_1 x_2 |x_2|^{0.1} - b x_3 \end{pmatrix} \quad (25.2)$$

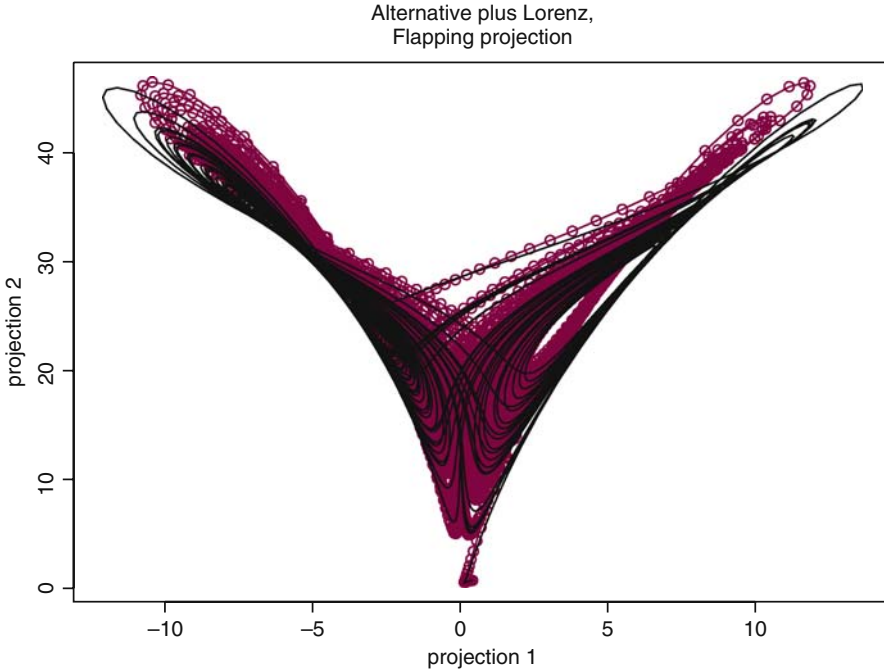


Figure 25.2. Null and alternative overlay in revealing projection

In Figure 25.2 we show both the null and the alternative model in the same projection used before. The alternative model is outlined in dots. The alternative attractor has stretched in directions orthogonal to the original attractor. Using imagination and a bit of literary allusion, Lorenz’s butterfly appears to be flapping its wings, hence the name used in this chapter for the projection. This picture, observed using the spin component of Splus (c), motivated the testing approach developed here.

25.2 The Test Statistic and Supporting Theory

A point X , on the trajectory in question, is tested by projecting it onto the computed points representing the set of null attractors, giving a point \widehat{X}_π , and then $X - \widehat{X}_\pi$ is projected onto the direction orthogonal to the manifold as estimated by the sample covariance matrix $\widehat{\Sigma}_\pi$ of nearest neighbors of \widehat{X}_π on the collection of attractors. The direction is estimated by the vector ξ corresponding to the smallest eigenvalue of $\widehat{\Sigma}_\pi$.

The estimate \widehat{X}_π has two sources of variation. One is the variation from the spin-up, Σ_π . Under a consistency assumption (discussed later) this covariance, appropriately normalized, converges to a constant with the increase in computed time on the spin-up and increase in number of points in the neighborhood of the nearest neighbor with the appropriate attention to relative growth rates. It also has a component from the projection of the measurement error in X (with covariance matrix Σ_μ) onto the manifold. If

we define Ξ as the matrix of eigenvectors of Σ_π and ξ as the eigenvector corresponding to the smallest eigenvalue of Σ_π , then the covariance of \widehat{X}_π has the form [24]

$$\Sigma_\pi + (\Xi \Xi^T - \xi \xi^T) \Sigma_\mu (\Xi \Xi^T - \xi \xi^T)^T. \quad (25.3)$$

For practical purposes, we will condition on X , in considering the test statistic so the second term in Eq. (25.3) vanishes. To form a test statistic, consider the terms which have the largest asymptotic order when we look at $(X - \widehat{X}_\pi) \widehat{\xi} \widehat{\xi}^T$. This can be rewritten as

$$\begin{aligned} (X - \widehat{X}_\pi) \widehat{\xi} \widehat{\xi}^T &= (X - X_\pi + X_\pi - \widehat{X}_\pi) (\xi \xi^T - \xi \xi^T + \widehat{\xi} \widehat{\xi}^T) \\ &= (X - X_\pi + op(1)) (\xi \xi^T + op(1)) \\ &\approx (X - X_\pi) (\xi \xi^T), \end{aligned} \quad (25.4)$$

where we note that terms that are $op(1)$ vanish in approximation [24], and that $X - X_\pi$ has an $Op(1)$ component from the measurement error. Now if this representation can be proved by invoking appropriate theorems for consistency of estimates in spin-ups of chaotic attractors, then an appropriate test statistic can have either of the two following forms, where the first includes contributions from the spin-up, and the second assumes a spin-up with computer sampling so large that the data approximates a continuous manifold. Note that when statistical noise is added in to the spin-up (more closely approximating conditions for weather for example), the spin-up contribution would have to be retained. This approximation should be good for statistical noise, small relative to the range of the attractor:

$$\begin{aligned} &(X - \widehat{X}_\pi)^T \widehat{\xi} \widehat{\xi}^T (\widehat{\Sigma}_\mu + \widehat{\Sigma}_\pi)^{-1} \widehat{\xi} \widehat{\xi}^T (X - \widehat{X}_\pi) \text{ or} \\ &(X - \widehat{X}_\pi)^T \widehat{\xi} \widehat{\xi}^T (\widehat{\Sigma}_\mu)^{-1} \widehat{\xi} \widehat{\xi}^T (X - \widehat{X}_\pi) \end{aligned} \quad (25.5)$$

In general the manifold would be m -dimensional in an n -dimensional space, where m has a fractional component in parts of the space, so typically $\xi \xi^T$ will be the $\lfloor (n - m) \rfloor$ smallest eigenvectors.

In the next section we will examine the testing procedure in a simulated example, including some discussion of initial attempts which did not work. Below we state the two main results. As stated, the results are theorems, but part of the statement of each is a conjecture of the continuity of the attractor [21] with respect to the parameters that requires more knowledge of dynamic systems than I have. As is consistent with the chaos literature [10] the simulation results are consistent with the theorems and so provides some support for making this assumption.

Define a function F to be Holder continuous function [1] if

$$|F(x) - F(y)| \leq C(F) [\rho(x, y)]^\beta, \quad (25.6)$$

where $\beta > 0$ is the holder exponent and ρ is the distance.

Let f be the time parameterized function corresponding to the system of equations generating the attractor. Then the global unstable manifold V_x^u is defined in [5] “provided that for every $t > 0$ there is x_{-t} such that $f^t(x_{-t}) = x$; the definition is

$$V_x^u = \left\{ y : \exists y_{-t} \text{ such that } f^t(y_{-t}) = y \text{ and } t \xrightarrow{\text{lim}} \infty \frac{1}{t} \log(\rho(x_{-t}, y_{-t})) < 0 \right\}'' , \quad (25.7)$$

which loosely translates as the set of y such that x and y were very close as we go into the infinite past.

An SRB measure μ is a measure on the attractor which is stable in the direction where the attractor is flowing. The precise definition of an SRB measure is beyond the scope of this chapter but is provided in [5]. However, one of its properties is that it is absolutely continuous with respect to Lebesgue measure on the V_x^u , the unstable manifolds.

Theorem 1. *Suppose that within the neighborhood of uncertainty of the parameters of the chaotic system, the system converges from a basin to a strange attractor, and:*

1. *On that strange attractor, averages of Holder continuous functions with 0 expectation over the attractor for an SRB measure μ converge in probability to 0 as T the elapsed time of the ‘spin-up’ increases to ∞ .*
2. *The flow is mixing on the strange attractor.*
3. *The strange attractor is continuous with respect to the parameters, both with respect to continuously defined geometric approximations of the attractor and with respect to properties one and two.*
4. *The measurement error has finite second moment.*
5. *Each calculated strange attractor starts at an initial value which is an early point in the measured trajectory. All calculations start at the same point.*
6. *The number of points in the neighborhood of X_π used to estimate $\Sigma_X \pi$ increases to infinity as T the spin-up time increases to infinity, at a rate slow enough that the neighborhood dimension goes to 0.*

Then the test statistic converges to a Null distribution depending only on the measurement error distribution and the geometry of the closest attractor(s) in the neighborhood as the time for each attractor spin-up tends to infinity.

The second condition is probably contained in the first, but I am not sure how to prove it at this point.

INDICATION OF PROOF. Assumption 5 of the theorem is required to ensure that the trajectory is not just caused by some simple perturbation of the system followed by a subsequent relaxation back to the attractor. The test is set up to be conservative in this case. Assumption 6 forces the conditional covariance of X_π to converge to a constant matrix assuming we can prove convergence.

To prove this we first note that strange attractors are compact. Further, Assumption 3 in the theorem implies that any attractor spun-up from the domain of uncertainty will be very close to any other attractor as long as the domain is sufficiently small. So if x is a vector on an attractor, then x and xx^T are bounded over the whole attractor by some number C_1 , and they can be seen by the absolute continuity of the SRB measure with respect to Lebesgue measure to be Holder continuous functions. If we then take a small neighborhood A and define $y = xI_{x \in A}$, then y and yy^T are also seen to be Holder continuous. Since they exist on a compact set, they have finite expectation, so Equation (25.6) is seen to hold. As T tends to ∞ the mixing implies the points calculated for each attractor become denser until the only points in the nearest neighbor estimate of $\Sigma_X \pi$ are in the closest attractor(s) to X from the trajectory, where multiple attractors

will be involved only if they overlap at that point in the set of the union of attractors. I assume without proof that this set of overlap will have measure 0 with respect to the SRB measure on each attractor. Thus conditional on X , ξ will depend only on $\Sigma_X \pi$. Conditionally on X I can estimate X_π , $\Sigma_X \pi$, and ξ with arbitrary accuracy by allowing T the spin-up time to go to infinity, and n the nearest neighbor neighborhood to increase, for example, as square root(T). The first test statistic takes into account the fact that in truth T will never be infinity, while the second is predicated on assuming T is for all essential purposes in a good neighborhood of infinity and ξ corresponds to a zero eigenvalue. \square

Theorem 2. *If the Lorenz attractor satisfies condition 3 of theorem 1, the measurement error of the trajectory satisfies condition 4, and the test statistic calculation satisfies conditions 5 and 6, then the test statistic converges to a Null distribution depending only on the measurement error distribution and the geometry of the closest attractor(s) in the neighborhood as the time for each attractor spin-up tends to infinity.*

INDICATION OF PROOF. A corollary of the main result of [11] is that Lorenz attractor obeys an almost sure invariance principle with the time parameter the variance. This implies condition 1 of theorem 1. The main result of [16] is that the Lorenz attractor is mixing. \square

Affraimovich et al. [1] refers to a paper [4] where a central limit theorem and mixing have been proven for the Lorenz process earlier however, I have not yet obtained an original.

In particular this corollary implies that if the measurement error is Gaussian, both statistics have an asymptotic Chi square distribution with 1 degree of freedom for the Lorenz attractor.

Theorem 3. *1. Suppose that the conditions 1–4 of theorem 1 hold, for the original attractor, and that an initial point in the original phase space is available from early in the experimental trajectory. Also suppose the reconstruction is done in a way so that the mapping from the original attractor to the reconstructed attractor is smooth and invertible (e.g., following Takens [27] and Sauer et al. [22]). Further suppose this map is continuous and invertible over the set spanned by the attractors in the range of uncertainty of the unknown parameters. Then the test statistic distributional result is similar to that in theorem 1.*

2. Suppose that $(E_{mu}(X - X_\pi))^T \xi > 0$ in the original phase space. Then as long as $(E_{mu}(X - X_\pi))^T \xi > 0$ in the reconstructed phase space, the test will retain some power in detecting the departure.

INDICATION OF PROOF. The smoothness (infinitely differentiable) and invertability of the map for the attractors imply that conditions 1 and 2 of theorem 1 carry through for each attractor. Condition 2 is required in order for any real separation between the trajectory and the set of attractors remain in the reconstructed space. I have not yet been able to obtain the original papers to see if for example Taken's theorem implies 2 above. \square

These theorems describe only how to test individual points in the trajectory. And all rely on a continuity conjecture about how attractors vary in a small region of the assumed parameters. From Figure 25.2 it is clear that there will be points where there is no separation between the attractors on some trajectories, so a way of simultaneously testing a large number of points on the trajectory is required. Fortunately the

false discovery rate (FDR) [3] approach is quite useful here. In this procedure, the p values are ordered and p_i is compared to $\frac{\alpha*i}{(n+1)}$, the interesting p values are the set $\left\{i : \forall j \leq i, p_j < \frac{\alpha*j}{n+1}\right\}$. The criterion that the FDR controls is $E\left(\frac{\text{Falserejections}}{\text{totalrejections}}\right)$, and in several situations, including some weak serial correlation situations, like those given by asymptotics for the y described in theorem 1, it provides a good upper bound [26].

25.3 Computer Experiments

Figure 25.3 shows both why only a small neighborhood of the parameters for the attractor can be considered at least for testing close alternatives and why the initial values of the attractor are confined to a point on the trajectory to be tested. The figure shows 17 attractors: the alternative and 16 constructed by first sampling 4 parameter sets, then within each parameter set sampling 4 starting points, essentially at random in the basin of attraction. Thus for this sort of test to be useful, there has to be significant scientific work shoring up plausible values of the attractor parameters.

Alternative plus 4 fully random Lorenz with 4 nested random starting points,
Flapping projection

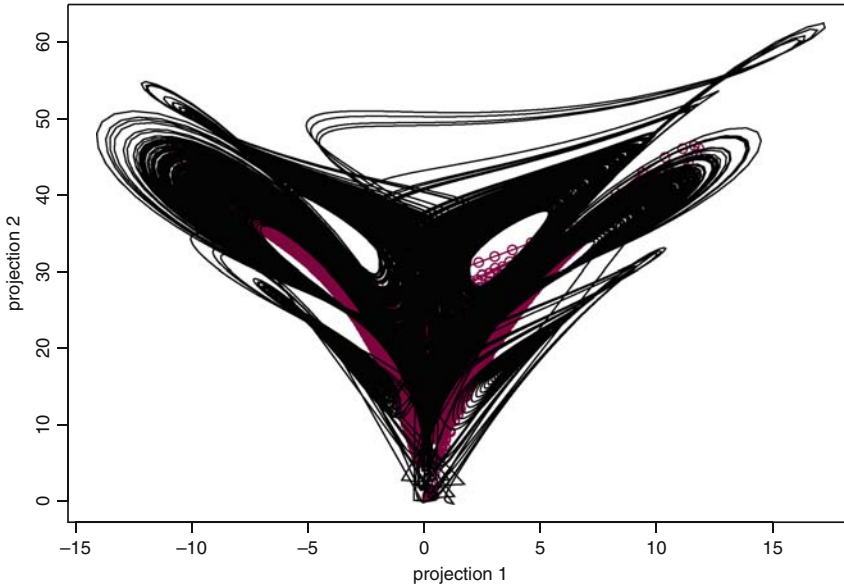


Figure 25.3. Large neighborhood of attractor in parameters for NULL plus alternative overlay

25.3.1 Description of the computer experiments

1. In the computer experiments shown below, the alternative, with parameter $(10, 28, 8/3)$, was spun out with 7,000 time steps (corresponding to a T of 700 in the dimensionless time of the equations).
2. Time step 1100 was selected as a starting point for the (10) spin-ups of the Null, and the parameter values of each spin-up was selected as $(10, 28, 8/3) \cdot \exp(z)$ where z was a Gaussian 3 vector with independent coordinates and standard error 0.01. The Null spin-ups were carried out for 4,000 time steps.
3. Tests were run using both test statistics on three reconstructions of the phase space of the attractor. The first, PS1, used the native phase space, (x, y, z) , the second, PS2, used $(x, x(2), z)$, and the third, PS3, used $(x, x(4), z)$, where $x(2)$ has a delay of two time steps and $x(4)$ has a delay of four time steps. When a delay is used, the time series was sampled at that rate to simplify sorting and ordering. So the second reconstruction had half as many time points, and the third $1/4$, while T still remain 700 and 400 between the alternative and null spin-ups, respectively.
4. The tests were compared to two different FDR criteria. In the first the p values were calculated by comparing the test statistic to a Chi square random variable with 1 degree of freedom, in the second p values were calculated by choosing each of the null spin-ups to create a trajectory, and developing statistics based on the remaining null spin-ups, and calculating p values based on the null statistics.
5. The trajectories tested on the first reconstructions ran from time steps 1251 to 1400 (trajectory 1) and 801 to 950 (trajectory 2). The trajectories on the second reconstruction ran from 626 to 700 (trajectory 3) and 401 to 475 (trajectory 4). The trajectories tested on the third reconstruction were from 313 to 350 (trajectory 5) and from 201 to 238 (trajectory 6). Note that trajectories 1, 3, and 5 cover the same region (region 1) of the attractors, as do trajectories 2, 4, and 6 (region 2).

In Figure 25.4 above we plot the results of applying statistic 1 and both Chi square 1 and empirical null distributions to the alternative and each of the 10 spun-up null distributions separately for trajectory 1 (in the original (x, y, z) phase space). The plot in the upper left-hand corner is the graphical display of the FDR test for the trajectory from the alternative using a Chi square 1 distribution. The plot in the upper right is similar showing each spun-up local null distribution compared to the FDR criterion using Chi square 1 p values. The FDR level was set at 0.05 for the upper plots. The plots in the lower left and right do the same comparisons, respectively, with the null distributions determined from the empirical null statistics. Because there were only 10 null distributions in the spin-ups, the FDR level had to be set at 0.1 in order to allow any rejections.

To compare power and specificity of the tests using each statistic, Null distribution estimation procedure, and phase space reconstruction procedure, every trajectory was tested using each test and null distribution comparison. To compare powers the FDR level was set at 0.1 for each.

The estimated false discovery rate is less than 0.02 in all of the calculations, even though the nominal (theoretical upper bound) on the FDR is 0.1 for this table. So the main concern is true discovery rate. Table 25.1 shows the true discovery rate for each region (1 and 2), statistic (1 and 2), method of null distribution construction (H0 measure), and phase space reconstruction (PS1, PS2, and PS3). Clearly the original phase space offers the most power, but also the second reconstruction, $(x, x(2), z)$, does

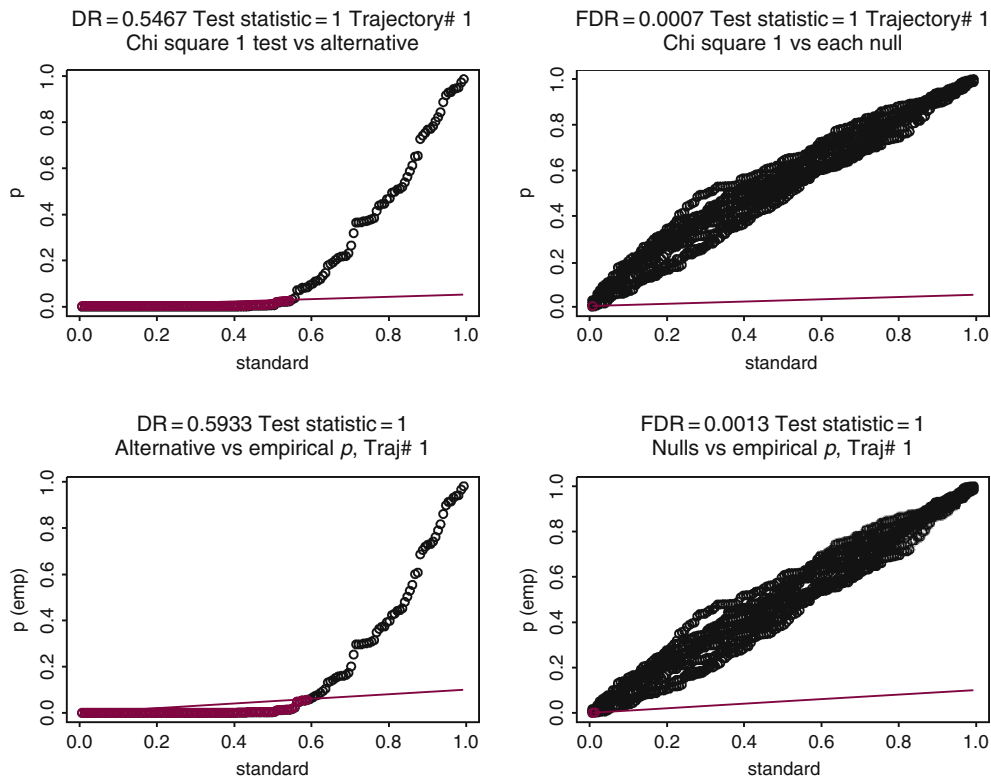


Figure 25.4. False discovery rate plots, alternative and null hypothesis, Chi square, and computational null distributions, points below line are rejections of null hypothesis

not do as well as the third, $(x, x(4), z)$. In fact in region 2, the discovery rate falls below the nominal bound on the false discovery rate. The two test statistics appear roughly equivalent. The methods of calculating the null are also close to one another, so the main difference is the region, of the attractor being tested, and the reconstruction method.

Table 25.1. Discovery rates for reconstruction procedures 1,2, and 3

Region -	Statistic -	H0 measure	PS1	PS2	PS3
Region 1	Statistic 1	Chi square	0.56	0.25	0.34
Region 1	Statistic 1	Empirical	0.59	0.33	0.37
Region 1	Statistic 2	Chi square	0.59	0.28	0.42
Region 1	Statistic 2	Empirical	0.58	0.33	0.40
Region 2	Statistic 1	Chi square	0.83	0.07	0.37
Region 2	Statistic 1	Empirical	0.85	0.11	0.29
Region 2	Statistic 2	Chi square	0.85	0.11	0.45
Region 2	Statistic 2	Empirical	0.84	0.08	0.29

25.4 Directions for Future Work

In addition to replacing each conjecture in the statement and proofs of the theorems with solid results, work needs to be done to extend this from working on toy examples to real problems. For this a number of points needs to be addressed:

1. Alternative definitions of neighborhoods need to be considered. For example, suppose small random perturbations are incorporated into the trajectories (instead of being added as measurement error). The resulting partly stochastic trajectories would tend to fall back onto the attractor as long as the perturbations are small enough, and enough time was allowed in between. In this case, combined with uncertainty in the underlying parameters, the space spanned by the ξ would have non-zero eigenvalues. Limits must also be placed on how small the statistical noise needs to be relative to the range of the attractor in order for the approximation used here to hold. The essential results should carry through, but the first statistic, using the null distribution calculated from the spun-up (now stochastic) attractors, should show more accuracy than the second in this situation. (Author's Note: Early simulations studying this problem indicate the second statistic is useless in this situation, and the simulated distribution is necessary for calculating p values correctly for the first statistic.)
2. The issue of good reconstructions for this kind of testing is an area that needs to be followed up. Perhaps an adaptation of the continuity metric of Pecora et al. [18] adapted both to points in the trajectory and in the attractor neighborhood of interest could be developed.
3. This test is only a partial goodness of fit test. The projection is onto the manifold, and the actual track between successive points is ignored. If this partial goodness of fit test agrees with the null hypothesis, likelihood estimation procedure could be developed by looking at one-step ahead procedures. Such an approach not only would allow the full statistical modeling machinery of likelihood estimation to be applied, but would allow a more accurate and precise goodness of fit test, making sure that not only does the data lie in the correct manifold, but it also has the correct track between successive data points.
4. This approach has the potential to test against higher dimensional attractors as long as the reconstruction space is larger than the manifold dimension. In addition the ξ provide an opportunity for meaningful dimension reduction, potentially identifying the subspaces of the phase space which are not being well modeled in given locations.

References

1. Afraimovich, V.S., Chernov, N.I., and Sateav, E.A., (1995), Statistical properties of 2-D generalized hyperbolic attractors, *Chaos an Interdisciplinary Journal of Nonlinear Science*, **5**, 238–252.

2. Barnett, W.A., Gallant, A.R., Hinich, M.J., Jungielges, J.A., Kaplan, D.T., and Jensen, M.J. (1997), A single-blind controlled competition among tests for nonlinearity and chaos, *Journal of Econometrics* **1**, 157–192.
3. Benjamini, Y. and Hochberg, Y. (1995), Controlling the false discovery rate, a practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society, Series B*, **57**, 289–300.
4. Bunimovich, L. (1983), Statistical properties of Lorenz attractors, In *Nonlinear Dynamics and Turbulence* (Ed., G.I. Barenblatt), pp. 71–92, Pitman, Boston.
5. Eckmann, J.P. and Ruelle, D. (1985), Ergodic theory of chaos and strange attractors, *Reviews of Modern Physics*, **57**, 617–656.
6. Edelen, D.G.B. (1985). *Applied Exterior Calculus*, Wiley Interscience, New York.
7. Gottwald, G.A. and Melbourne, I. (2005), Testing for chaos in deterministic systems with noise, *Physica D: Nonlinear Phenomena*, **1–2**, 100–111.
8. Guzzetti, S., Signorini, M.G., Cogliati, C., Mezzetti, S., Porta, A., Cerrutti, S., and Malliani, A. (1996), Non-linear dynamics and chaotic indices in heart rate variability of normal subjects and heart-transplanted patients, *Cardiovascular Research*, **31**, 441–446.
9. Hastings, A., Hom, C.L., Ellner, S., Turchin, P., and Godfray, H.C.J. (1993). Chaos in ecology: is mother nature a strange attractor, *Annual Review of Ecology and Systematics*, **24**, 1–33.
10. Henon, M. and Heiles, C. (1964), The applicability of the third integral of motion: some numerical experiments, *Astron. J.*, **69**, 73–79.
11. Holland, M. and Melbourne, I. (2007), Central limit theorems and invariance principles for Lorenz attractors, *Journal of the London Mathematical Society*, **76**, 345–364.
12. Kennel, M.B. and Isabelle, S. (1992), Method to distinguish possible chaos from colored noise and to determine embedding parameters, *Physical Review A*, **46**, 3111–3118.
13. McCaffrey, D.F., Ellner, S., Gallant, A.R. and Nychka, D.W. (1992), Estimating the Lyapunov exponent in chaotic system with nonparametric regression, *Journal of the American Statistical Association*, **97**, 682–695.
14. Lorenz, E.N. (1960), Maximal simplification of the dynamic equations, *Tellus*, **12**, 243–254.
15. Lorenz, E.N. (1963), Deterministic nonperiodic flow, *Journal of Atmospheric Science*, **20**, 130–141.
16. Luzzatto S., Melbourne, I., and Paccaut, F. (2005), The Lorenz attractor is mixing, *Communications in Mathematical Physics*, **260**, 393–401.
17. Nychka, D., Ellner, S., Gallant, R., and McCaffrey, D. (1992), Finding chaos in noisy systems, *Journal of the Royal Statistical Society, Series B*, **54**, 399–426.
18. Pecora, L.M., Moniz, L., Nichols, J., and Carrol, T.L. (2007), A unified approach to attractor reconstruction, *Chaos*, **17**, DOI:10.1063/1.2430294.
19. Poon, C.S. and Barahona, M. (2001), Titration of chaos with added noise, *Proceedings of the National Academy of Sciences*, **98**, 7107–7112.
20. Raleigh, L. (1916), On convective currents in a horizontal layer of fluid when the higher temperature is on the underside. *Philosophical Magazine* **32**, 529–546.
21. Ruelle, D. (1987), Diagnosis of dynamical systems with fluctuating parameters, *Proceedings of the Royal Society of London, A*, **413**, 5–8.
22. Sauer, T., Yorek, J., and Casdagli, M., (1991), Embedology, *Journal of Statistical Physics*, **65**, 579–616.
23. Saltzman, B. (1962), Finite amplitude free convection as an initial value problem, *Journal of Atmospheric Science*, **19**, 329–341.
24. Serfling, R. (1980), *Approximation Theorems in Mathematical Statistics*, Wiley Series in Probability and Statistics, New York.
25. Shimada, I. and Nagashima, T. (1979), A Numerical approach to ergodic problem of dissipative dynamical systems, *Progress of Theoretical Physics*, **61**, 1605–1616.
26. Storey, J.D. and Tibshirani, R. (2003), Statistical significance for genome wide studies, *Proceedings of the National Academy of Sciences*, **100**, 9440–9445.

27. Takens, F. (1981), Detecting strange attractors in turbulence, In *Dynamical Systems and Turbulence* (Eds., A. Dold and B. Eckmann), pp. 366–381, Springer Berlin/Heidelberg.
28. Tziperman, E., Stone, L., Cane, M.A., and Jarosh, H. (1994), El Nino chaos: overlapping of resonances between the seasonal cycle and the Pacific Ocean-atmosphere oscillator, *Science*, **264**, 72–74.

Probability Plotting with Independent Competing Risks

Francis G. Pascual¹ and Christopher Gast²

¹ Department of Statistics, Washington State University, Pullman, WA

² Axio Research Acquisition Co. LLC, Seattle, WA

Abstract: In this chapter, we discuss probability plotting for competing risk data in the presence of Type I or right censoring. The construction of probability plots is based on the linearization of the cumulative distribution function of the first-order statistic. We consider the case when risks are independent. We describe procedures for constructing pointwise confidence intervals based on large-sample properties of maximum likelihood estimators. The proposed method is compared with traditional probability plotting that assumes that causes of failure can be eliminated.

Keywords and phrases: Competing risks, maximum likelihood estimation, probability plotting, right censoring

26.1 Introduction

26.1.1 Competing risks

In engineering reliability, practitioners are often interested in the fraction of products that fail within a certain length of time due to either a known or an unknown single cause. Often, lifetime data collected to address this goal are right censored at possibly varying times. Basic nonparametric theory involves the use of Kaplan–Meier estimators, and computing pointwise confidence intervals about the estimated proportions failing or surviving. Parametric analyses fit statistical models to data, often by maximum likelihood (ML) methods.

There are situations, e.g., engineering and medical, when $k \geq 2$ failure modes or risks “compete” for the lifetime of an individual such that the subject dies when the first of these risks occurs. Thus, we are concerned with the first (or minimum)-order statistic $T_{(1)} = \min\{T_1, \dots, T_k\}$, where we assume T_i is the time to failure due to mode i for $i = 1, \dots, k$. We interpret subject life as the survival time of a k -part series system.

Reference [3] and Chapter 7 of [15] describe situations involving two or more competing risks. For instance, assemblies of ball bearings can fail because of either race or ball failures. Reference [10] describes how the failure of an electronic product can occur

either during production or as wearout during actual use. Reference [15] describes how motorett insulation can fail due to any of three specific causes. For the method we propose here, we assume that an autopsy can be performed in order to identify exactly what caused death or failure.

Modeling competing risks has been discussed in the literature. Such models are interchangeably called competing risk, multi-risk, compound or series-system models. For example, exponential distributions were used by [2] to model two competing risks and, more generally, by [6] to model three or more risks. Independent Weibull distributions were used by [7] to model failure times of two risks with product lifetime defined to be the first occurrence of these two risks.

Herman and Pattel [6] discussed maximum likelihood methods for fitting a two-risk model to data from an experiment terminated at one specific point in time. Assuming independent risks, they derived expressions for the likelihood and asymptotic distribution of ML estimators. They demonstrated their results with the exponential and Weibull distributions. Moeschberger and David [14] and David and Moeschberger [5] derived likelihood expressions for right-censored data when competing risks are distributed lognormal or Weibull. They also presented the asymptotic properties of ML estimators.

26.1.2 Probability plotting

Probability plotting is a technique for assessing how well a distribution describes a data set. One such type of plot is the quantile–quantile (Q–Q) plot wherein the quantiles of the data are plotted against the quantiles of a candidate distribution. This works very well with location-scale distributions so that the candidate can be replaced with the standardized distribution for which the location parameter is 0 and the scale parameter is 1. That is, a parametric estimate of the candidate is not necessary for plotting. Chapter 6 of [11] presents probability (Q–Q) plots with emphasis on location-scale distributions. They demonstrate how to choose plotting axes by linearizing the cumulative distribution function (cdf).

In the Q–Q plot, the sample quantiles can be replaced with the Kaplan–Meier estimate of the cdf, and the quantiles of the candidate distribution with their corresponding cdf values. This plot is what [4] refers to as the P–P plot. Transformations of the P–P plot coordinates may also be performed to stabilize the variances of the plotting points particularly at the lower and upper regions. See, for example, the stabilized probability (S–P) plot discussed in [13] that uses the arcsin transformation.

In general, linearity of the points in a probability plot suggests an adequate fit of the distribution to the data. More sophisticated interpretations of plots may involve deriving so-called simulated envelopes and performing goodness-of-fit tests. On page 66 of [4], the authors provide pertinent references. In Chapter 6 of [11], the authors provide guidelines for interpreting probability plots. In particular, they suggest including simultaneous confidence bands around the estimated cdf as an aid in assessing fit.

Reference [8] proposed a probability plot for fitting an independent Weibull competing risk model to data. Product failure time was defined to be when the first of several risks occurred. The method was demonstrated on complete (i.e., no censoring) data without information on actual cause of failure. Beforehand, the number of active risks is unknown and is determined, along with model parameters, from graphical properties of the probability plot.

26.1.3 Outline

In this chapter, we propose a probability plotting method for right-censored lifetime data under competing risks based on the cdf of $T_{(1)}$. We consider situations when the risks are statistically independent. Section 26.2 presents conventional notation and model assumptions regarding competing risks. In Section 26.3, we review nonparametric and parametric estimation methods as it relates to competing risks. We introduce our proposed method of creating probability plots based on the distribution of $T_{(1)}$ in Section 26.4. We apply the method to actual survival/reliability and simulated data sets in Section 26.5.

26.2 Notation and Model Assumptions

In our examples below, we consider data from both engineering and medical applications which use different, but practically equivalent, terminologies. Below, “subject” pertains to a product (engineering reliability) or a patient (clinical trials). We shall use “lifetime” to denote time to failure or death, and interchangeably use the terms “failure” and “death” and the terms “experiment” and “clinical trial.”

From this point on, we suppose that there are k independent competing risks indexed by $i = 1, \dots, k$. Let T_i be the time to failure of a subject due to risk i for $i = 1, \dots, k$. Let $F_i(t_i)$ and $f_i(t_i)$ be the cdf and probability density function (pdf) of T_i , respectively.

26.2.1 Distributions of individual risks

Below, we consider the cases when risk failure times are either Weibull or lognormal. These are log-location-scale distributions commonly used in survival analysis. It is known that the (natural) logarithm of Weibull is smallest extreme value (SEV) and the logarithm of lognormal is normal. Let μ_i and σ_i denote the location and scale parameters, respectively, of either the SEV or the normal distribution. Assign $z_i = [\log(t) - \mu_i]/\sigma_i$. If risk i is Weibull, then its cdf and pdf can be written as

$$\begin{aligned} F_i(t) &= \Phi_i(z) = 1 - \exp[-\exp(z_i)], \\ f_i(t) &= \frac{1}{\sigma_i t} \exp[z_i - \exp(z_i)], \end{aligned} \quad (26.1)$$

respectively, for $t > 0$. If risk i is lognormal, then its cdf and pdf are given by

$$\begin{aligned} F_i(t) &= \Phi_i(z) = \int_{-\infty}^{z_i} \frac{1}{\sqrt{2\pi}} \exp(-u^2/2) du, \\ f_i(t) &= \frac{1}{\sigma_i t \sqrt{2\pi}} \exp(-z_i^2/2), \end{aligned} \quad (26.2)$$

respectively, for $t > 0$.

26.2.2 Distribution of subject lifetime

The subject lifetime T is denoted by the time when the first of the k competing risks occurs, i.e., $T = \min\{T_1, \dots, T_k\}$. One or all of T_1, \dots, T_k can be latent or unobserved. For example, if $T = T_1$, then the cause of death is risk 1, and T_2, \dots, T_k are right censored at T_1 . If an experiment is terminated before any risk could occur, then all of T_1, \dots, T_k are right censored at the termination time.

For independent risks, time to failure T has cdf

$$F(t) = 1 - \prod_{i=1}^k [1 - \Phi_i(z_i)],$$

where Φ_i is given by either (26.1) or (26.2). Furthermore, we can write $F(t)$ in the form

$$F(t) = 1 - [1 - \Phi_i(z_i)] \times \prod_{\substack{j=1, \dots, k \\ j \neq i}} \{1 - \Phi_j[z_i \sigma_i / \sigma_j + (\mu_i - \mu_j) / \sigma_j]\} \quad (26.3)$$

for $i = 1, \dots, k$.

26.2.3 The likelihood function

Because of censoring, maximum likelihood (ML) methods are often used to estimate model parameters. Below, we give the general form of the likelihood function for competing risks with right censoring as derived in [5].

Assume that when a unit fails, the practitioner is able to perform an “autopsy” to determine the exact cause of failure. Let n be the sample size, and for $j = 1, \dots, n$, let t_j be the j th realization of T and c_j denote the cause of death ($c_j = 1, \dots, k$) or right censoring ($c_j = k + 1$). Then, lifetime data consist of ordered pairs, namely, $\{(t_1, c_1), (t_2, c_2), \dots, (t_n, c_n)\}$. Let m_i denote the total number of failures due to risk i for $i = 1, \dots, k$. Let m_{k+1} be the total number of right-censored observations. Let t_{i1}, \dots, t_{im_i} be mode i failure times, and let $t_{k+1,1}, \dots, t_{k+1,m_{k+1}}$ be the right-censored observations. From (2.9) of [5], the likelihood of the data when risks are independent is given by

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^k \prod_{j=1}^{m_i} \left\{ f_i(t_{ij}) \times \prod_{\substack{l=1, \dots, k \\ l \neq i}} [1 - F_l(t_{ij})] \right\} \times \prod_{j=1}^{m_{k+1}} \prod_{i=1}^k [1 - F_i(t_{k+1,j})], \quad (26.4)$$

where $\boldsymbol{\theta}$ is the vector of parameters to be estimated from data. The value $\hat{\boldsymbol{\theta}}$ that maximizes the likelihood is the ML estimate of $\boldsymbol{\theta}$. By invariance, the ML estimate of a function of $\boldsymbol{\theta}$ (e.g., quantile, probability, pdf) is obtained by substituting $\hat{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}$.

Let $\boldsymbol{\theta}_i = (\mu_i, \sigma_i)$ for $i = 1, \dots, k$. Let $L_i(\boldsymbol{\theta}_i)$ be the likelihood when risk i is the only cause of failure, and failures due to the other risk are treated as right-censored data. A close inspection of (26.4) reveals that $L(\boldsymbol{\theta}) = L_1(\boldsymbol{\theta}_1) \times \dots \times L_k(\boldsymbol{\theta}_k)$ and it follows that if $\hat{\boldsymbol{\theta}}_i$ maximizes $L_i(\boldsymbol{\theta}_i)$, then the ML estimate based on $L(\boldsymbol{\theta})$ is $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\theta}}_1, \dots, \hat{\boldsymbol{\theta}}_k)$.

26.3 The Kaplan–Meier Estimator And Probability Plotting

26.3.1 Kaplan–Meier estimator

If nonparametric estimates for the proportion failing are desired, the most commonly used approach is the Kaplan–Meier (KM) estimator introduced in [9]. Suppose that t_1, \dots, t_m are the m unique failure or censoring times of the data sorted in increasing order. If n is the sample size, then $m \leq n$ with equality holding when there are no ties. Assign $t_0 = 0$. Let d_i denote the number of failures at time t_i , and let r_i denote the number of units censored at time t_i . Set $d_0 = 0$ and $r_0 = m$. Then the number of subjects alive at time t_i is

$$n_i = n - \sum_{j=0}^{i-1} d_j - \sum_{j=0}^{i-1} r_j, \quad i = 1, \dots, m.$$

The KM estimate for the survival probability $S(t_i) = 1 - F(t_i)$ is

$$\widehat{S}(t_i) = \prod_{j=1}^i \left[1 - \frac{d_j}{n_j} \right], \quad i = 1, \dots, m,$$

and the estimate for $F(t_i)$ is

$$\widehat{F}(t_i) = 1 - \widehat{S}(t_i) = 1 - \prod_{j=1}^i \left[1 - \frac{d_j}{n_j} \right], \quad i = 1, \dots, m.$$

For $t_i \leq t < t_{i+1}$, $\widehat{F}(t) = \widehat{F}(t_i)$.

An estimate of the standard error of $\widehat{F}(t_i)$ is computed by Greenwood's Formula as

$$\widehat{s}_i = \widehat{S}(t_i) \sqrt{\sum_{j=1}^i \frac{d_j}{n_j(n_j - d_j)}} \quad (26.5)$$

For simpler notation below, we omit the argument t_i or t in the cdf.

Confidence intervals (CIs) for the cdf F can be computed using a normal approximation of the distribution of \widehat{F} . This method's coverage probability can be improved, particularly in small samples, by utilizing the logit transformation $\text{logit}(p) = \log[p/(1-p)]$. Thus, an approximate nonparametric $100(1-\alpha)\%$ CI for F is given by

$$(E, \bar{F}) = \left[\frac{\widehat{F}}{\widehat{F} + (1 - \widehat{F})w_i}, \frac{\widehat{F}}{[\widehat{F} + (1 - \widehat{F})]/w_i} \right], \quad (26.6)$$

where \widehat{F} is the KM estimate (or any other estimate) of F , $w_i = \exp\{z_{(1-\alpha/2)} \times \widehat{s}_i / [\widehat{F}(1 - \widehat{F})]\}$, and $z_{(1-\alpha/2)}$ is the $1 - \alpha/2$ quantile of the standard normal distribution.

26.3.2 Linearizing the Cdf under one risk

In this section, we review a parametric approach to probability plotting for one cause of failure. This involves linearizing the cdf F that has a known functional form. This will allow us to plot the data along with a straight line that represents a cdf from the candidate distribution on the same set of axes. The fit of the distribution to the data is assessed by how much the data follow a straight line.

Let $p = F(t_p)$, that is, t_p is the p quantile of a candidate cdf F . In linearizing the cdf, the most common approach is to find a transformation $g(t_p)$ so that there is a linear relationship between $g(t_p)$ and functions of the model parameters. Let β_0 and β_1 be functions of only the distribution parameters, and $h(p)$ be a function, ideally, of only p . We must obtain a linear relationship of the form

$$h(p) = \beta_0 + \beta_1 g(t_p).$$

The horizontal axis $g(t_p)$ and vertical axis $h(p)$ are labeled using values of t_p and p , respectively. The line on this plot with slope β_1 and intercept β_0 represents the cdf $F(t)$ and is called the cdf line.

Sort the data $t_1 \leq \dots \leq t_n$, and then plot the points $(g(t_i), h(\hat{p}_i))$ for $i = 1, \dots, n$ on the $g(t_p) - h(p)$ axes where $\hat{p}_i = \hat{F}(t_i)$ is an estimate of $F(t_i)$. Choices for \hat{p}_i are $(i + 0.5)/n$, the KM estimate of $F(t_i)$, and the ML estimate of $F(t_i)$. If the points are reasonably linear, the underlying distribution is said to describe the data adequately. Otherwise, the candidate distribution is said to exhibit a lack of fit to the data.

The above procedure works well with (log) location-scale distributions such as the normal, lognormal, SEV, and Weibull distributions where the slope is $1/\sigma$ and the intercept is either μ or e^μ . See Table 6.2 of [11]. If $\hat{\mu}$ and $\hat{\sigma}$ are known parameter estimates, the estimated cdf line can also be drawn using the appropriate slope and intercept. CIs for $F(t)$ are computed as in (26.5)–(26.6) and included in the plot.

26.3.3 Probability plotting and competing risks

Meeker and Escobar [11] discuss strategies for probability plotting with $k = 2$ competing risks. We review them below.

1. Analyze the mode i failures only while treating failures due to other failures as right-censored data with respect to mode i . The results estimate the failure-time distribution if all other modes $j \neq i$ could be completely eliminated.
2. As in the method we propose here, this strategy considers a series system with independent parts for which the cdf of time to failure is $F(t) = 1 - [1 - F_1(t)] \times [1 - F_2(t)]$. ML estimates for $F_1(t)$ are obtained by treating mode 2 failures as right-censored in addition to the m_3 original right-censored data. An analogous ML estimation is done for $F_2(t)$.
3. Analyze the data ignoring the failure modes, i.e., treat failure modes as one type of failure. Sometimes this is adequate within the range of the data, but can be seriously incorrect when extrapolating.

Strategies 1 and 3 ignore the information on the specific cause of failure. Strategy 1 is appropriate only under the hypothesis that one or the other failure mode can be eliminated. Strategy 3 is useful only when the mode of failure is not of interest or the distributions of failure modes are very similar.

The ML estimates of parameters under strategies 1 and 2 are exactly the same as those under the likelihood given in (26.4). For example, this can be verified for strategy 2 by noting that (26.4) is the product of the likelihoods corresponding to modes $1, \dots, k$. If mode failure times come from the same family of distributions, the k probability plots for strategy 1 can be combined in one plot. On this plot the ML estimate of $F(t)$ generally plots as a curve instead of a straight line.

Below, we propose a method for probability plotting that uses all available failure-cause information without additional assumptions unlike strategies 1 and 3. Also, the ML estimate of $F(t)$ plots as a straight line on the axes that we suggest. We provide expressions for large-sample CIs for $F(t)$. As in the one-risk case, the fit of the model is judged by how closely the failure data follow the cdf line.

26.4 Proposed Method

Before we present our proposed probability plotting for k competing risks, we review the notation which we shall use below. $t_1 \leq \dots \leq t_n$ represent the data set disregarding the causes of failure. $t_{i1} \leq \dots \leq t_{im_i}$ are the failures due to risk i for $i = 1, \dots, k$. $t_{k+1,1} \leq \dots \leq t_{k+1,m_{k+1}}$ are the right-censored observations. Generically, t refers to failure time.

We linearize F as follows. Recall that $z_i = [\log(t) - \mu_i]/\sigma_i$ for $i = 1, \dots, k$. Let $G_i(z_i)$ be the right-hand side of (26.3) so that $F(t) = G_i(z_i)$ for $i = 1, \dots, k$. By solving for $\log(t)$ we obtain the relationships

$$G_i^{-1}[F(t)] = -\frac{\mu_i}{\sigma_i} + \frac{1}{\sigma_i} \log(t), \quad i = 1, \dots, k \quad (26.7)$$

which are all linear in $\log(t)$ (horizontal axis) and G_i^{-1} (vertical axis). Observe that G_i^{-1} is uniquely determined by the underlying distributions, and the values of σ_i/σ_j and $(\mu_i - \mu_j)/\sigma_j$ for $i, j = 1, \dots, k$ and $i \neq j$. Thus, to initially establish the vertical axis, we obtain estimates for these quantities based on the likelihood in (26.4).

Our approach results in k probability plots, one for each failure mode. The cdf line has k representations according to (26.7). We obtain an estimate $\hat{F}(t_{ij})$ of $F(t_{ij})$ using Kaplan–Meier. If the underlying distribution is appropriate, for each of $i = 1, \dots, k$, the plot of $\log(t_{ij})$ versus $G_i^{-1}[\hat{F}(t_{ij})]$ for $j = 1, \dots, m_i$ should be linear and the correlation between these coordinates should be high. If risks are independently distributed Weibull with common shapes, i.e., $\sigma_1 = \dots = \sigma_k$, one can obtain a closed form expression for G_i^{-1} based on the exponential cdf. Otherwise, G_i^{-1} is numerically computed using a root-finding algorithm.

An estimate of the variance of $\hat{F}(t_i)$ is given by

$$\hat{s}_i^2 = \left(\frac{\partial F}{\partial \theta} \right)' \hat{\Sigma}_{\theta} \left(\frac{\partial F}{\partial \theta} \right) \bigg|_{\theta = \hat{\theta}, t = t_i}, \quad i = 1, \dots, n \quad (26.8)$$

where $\partial F/\partial \theta$ is the vector of first-order partial derivatives of $F(t)$ with respect to θ , and $\hat{\Sigma}_{\theta}$ is the variance–covariance matrix (the inverse negative Hessian) which can be calculated numerically or obtained from standard output of ML estimation software.

In our computations, we reparameterize by replacing σ_i with $\lambda_i = \log(\sigma_i)$ so that the parameter space and the optimization of the likelihood are unconstrained. Thus, the vector of parameters is $\boldsymbol{\theta} = (\mu_1, \lambda_1, \dots, \mu_k, \lambda_k)$. The partial derivatives of F with respect to $\boldsymbol{\theta}$ are given by

$$\frac{\partial F(t)}{\partial \mu_i} = -\frac{\phi_i(z_i)}{\sigma_i} \times \prod_{\substack{j=1, \dots, k \\ j \neq i}} [1 - \Phi_j(z_j)]$$

and

$$\frac{\partial F(t)}{\partial \lambda_i} = \sigma_i z_i \frac{\partial F(t)}{\partial \mu_i}$$

for $i = 1, \dots, k$. Observe that

$$\frac{\partial F(t)}{\partial \sigma_i} = \frac{1}{\sigma_i} \frac{\partial F(t)}{\partial \lambda_i} = z_i \frac{\partial F(t)}{\partial \mu_i}.$$

We use the non-linear minimization function *nlm* of the R Project for Statistical Computing (<http://www.r-project.org>) to compute ML estimates and the Hessian.

We compute CIs for F using the logit-based approach in (26.6) where the standard error \widehat{s}_i is given by (26.8). Thus, for each risk i , we have a probability plot with

- points $\left(\log(t_{ij}), G_i^{-1} \left[\widehat{F}_{KM}(t_{ij}) \right] \right)$ for $j = 1, \dots, m_i$ where \widehat{F}_{KM} denotes the Kaplan–Meier estimate;
- cdf line $\left(\log(t), G_i^{-1} \left[\widehat{F}_{ML}(t) \right] \right)$ where \widehat{F}_{ML} denotes the ML estimate; and
- pointwise CI curves $\left(\log(t), G_i^{-1} \left[\underline{F}(t) \right] \right)$ and $\left(\log(t), G_i^{-1} \left[\overline{F}(t) \right] \right)$.

where the horizontal axis is in log scale. Points that follow a linear trend and fall within the CI curves suggest an adequate fit. The slope and intercept of the estimated cdf line correspond to estimates of $1/\sigma_i$ and $-\mu_i/\sigma_i$, respectively.

26.5 Applications

In this section, we apply the proposed probability plot to several data sets. We consider competing risk data from medicine and engineering that are found in the literature. We demonstrate probability plotting for the cases when risks are independent lognormal or Weibull. In all our examples, subjects failed due to $k = 2$ risks.

26.5.1 Breast cancer study

The first data set from [1], also found in [5], consists of paired observations (T =survival time in months, cause of death) for 121 patients treated for breast cancer. There were 78 deaths due to cancer (mode 1), 18 deaths due to other causes (mode 2), and 25 survivors (right-censored observations). On average, cancer deaths occurred about 16 months earlier than deaths due to other causes. We remove two observations with survival time $t = 0.3$, because initial probability plots suggest that these are outliers, possibly “infant mortality” data, and, hence, possibly not representative of the rest of

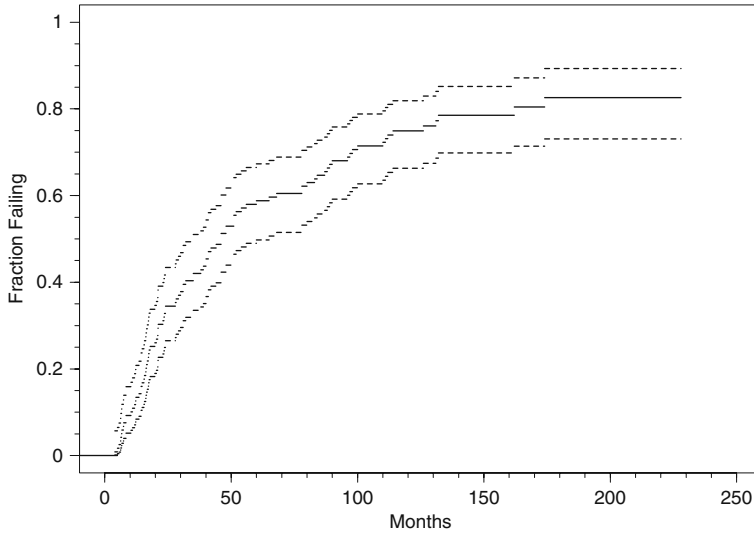


Figure 26.1. Plot of the KM estimate of the cumulative distribution function F with pointwise confidence intervals for the cancer data set

the data. Figure 26.1 plots the KM estimate of the cdf F with pointwise CIs based on (26.5) and (26.6). This plot was generated using the SPLIDA software developed by [12].

Figure 26.2 is a lognormal probability plot of survival times of cancer deaths while treating other deaths as right censored. This is the first strategy described in Section 26.3.3. The corresponding Weibull probability plot for survival times of Other deaths is in Figure 26.3. The two broken lines in these plots represent the ML esti-

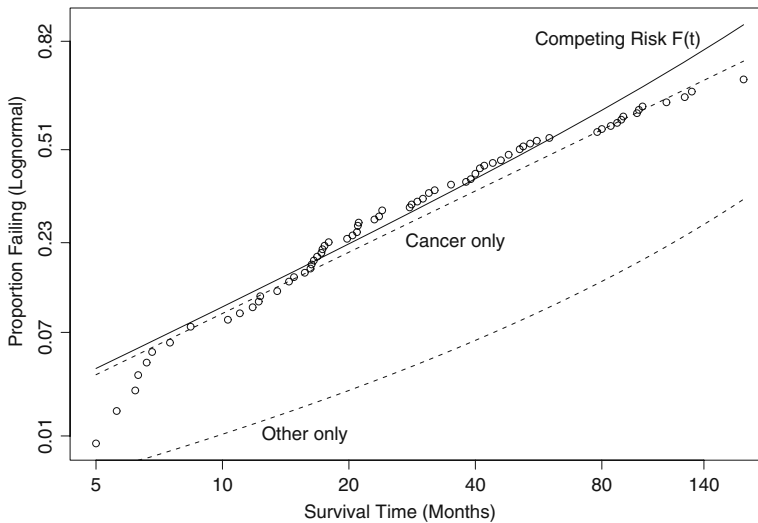


Figure 26.2. Lognormal probability plot of survival times for deaths due to cancer with estimates of the cdf assuming that cancer is the only risk, Other is the only risk, and cancer and Other are independent lognormal and Weibull, respectively

mates of the cdfs when only one risk is assumed to act on the subjects. The solid line represents the ML estimate of $F(t)$. Note that the latter plots as a curve instead of a straight line on these plots.

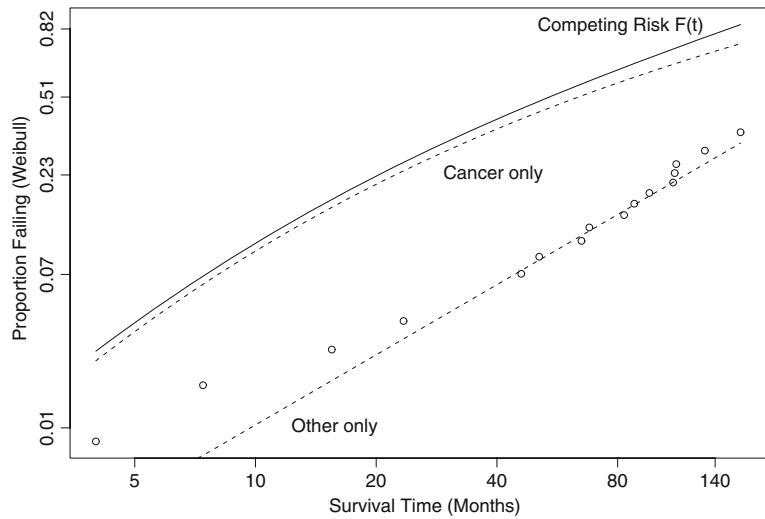


Figure 26.3. Weibull probability plot of survival times for deaths due to Other with estimates of the cdf assuming that cancer is the only risk, Other is the only risk, and cancer and Other are independent lognormal and Weibull, respectively

We consider all four combinations of lognormal or Weibull for modeling the times to death due to cancer or other causes. Table 26.1 provides the maximum likelihood values achieved from fitting these models. There are three sets of loglikelihood values in the table. The first, in column 3, is the maximum loglikelihood when cancer is the only risk, and deaths from other causes are treated as right censored. Column 4 gives the maximum likelihood when other causes are treated as the only risk, and cancer deaths are treated as right censored. As mentioned earlier, the sum of these loglikelihoods is exactly the maximum loglikelihood under the independent risks model with likelihood given in (26.4). See column 5.

Table 26.1. Maximum likelihood results for the cancer data set

Distributions		Maximum loglikelihood			Correlations		
Cancer	Other	Cancer	Other	Total	Cancer	Other	
Lognormal	Lognormal	-168.57	-54.76	-223.33	0.9906	0.9903	
Lognormal	Weibull	-168.57	-53.97	-222.54	0.9892	0.9887	
Weibull	Lognormal	-177.06	-54.76	-231.81	0.9645	0.9593	
Weibull	Weibull	-177.06	-53.97	-231.03	0.9626	0.9583	

Based on total likelihood, the lognormal–cancer and Weibull–other-causes combination provides the best fit. This is followed closely by the lognormal–lognormal combination. These models differ mainly in describing deaths due to other causes. The others provide about equally worse fits to the data.

The proposed competing risk probability plots for the two best fits are in Figures 26.4 and 26.5. The plots include 95% pointwise confidence intervals for the cdf of survival. It is not clear from the plots which of the lognormal–lognormal and lognormal–Weibull models provide a significantly better fit. Both sets of plots indicate an adequate description of survival time. The plotted points are concentrated around the estimated cdf line, with nearly all failure times contained within the 95% confidence intervals. We compute the linear correlations between the plotting coordinates $\log(t)$ and $G_i^{-1}[\hat{F}(t)]$ for all models, and these are given in the last two columns of Table 26.1. The correlations for the lognormal–lognormal model, followed by those for the lognormal–Weibull model, are the highest among the fitted models.

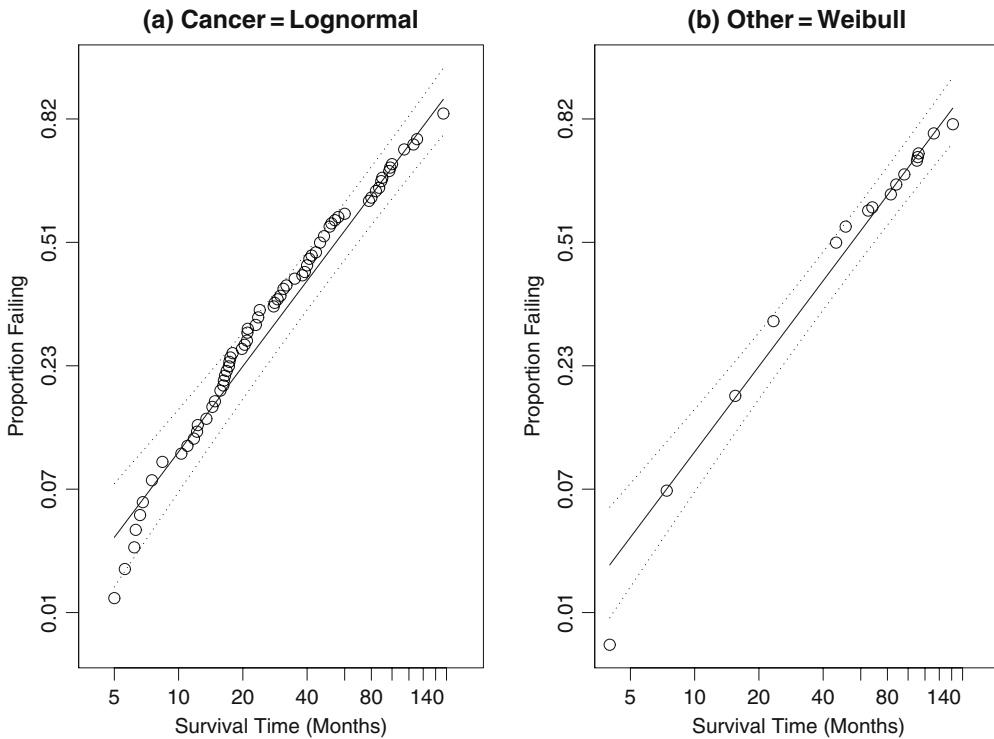


Figure 26.4. Probability plots for the cancer data set with pointwise confidence intervals for failure-time cdf F assuming independent lognormal cancer deaths and Weibull other deaths. Plot (a) is for deaths due to cancer. Plot (b) is for deaths due to other causes

26.5.2 Shock absorber failure data

We also apply the proposed method to a reliability data set on shock absorbers found in [16] and [11]. This data set describes the failure times of vehicle shock absorbers, recorded in number of kilometers of use, with failure modes M1 and M2. Out of 38 observations, 7 were M1 failures, 4 were M2 failures, and 27 were right censored at various times. On average, M1 failures occurred about 1,200 km later than M2 failures. Figure 26.6 plots the KM estimate of the cdf F .

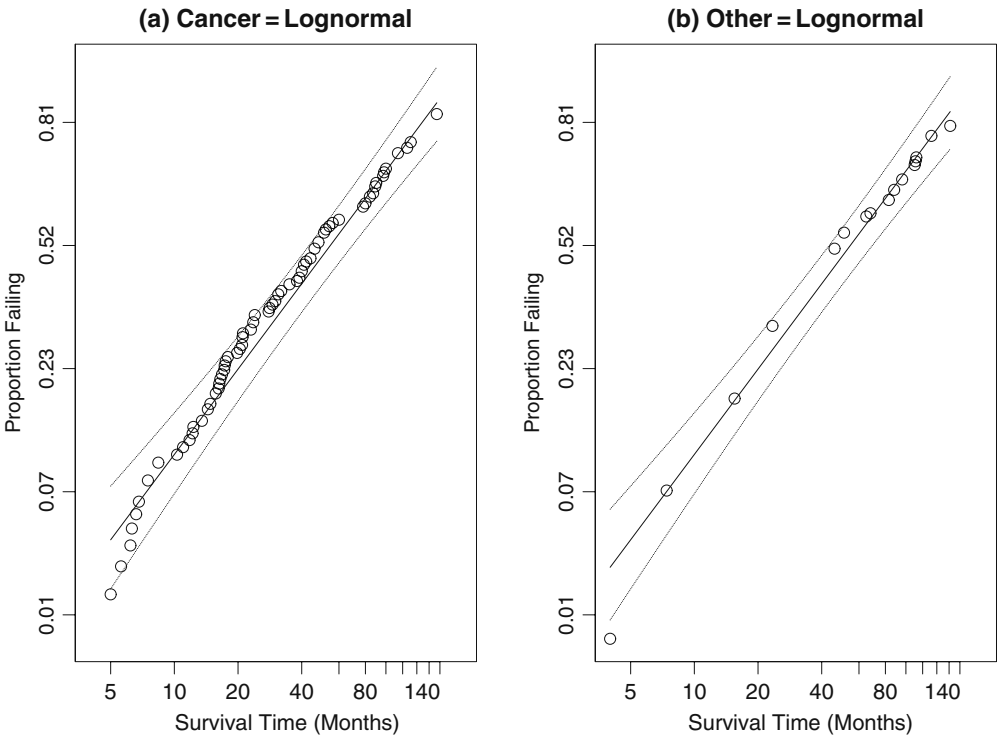


Figure 26.5. Probability plots for the cancer data set with pointwise confidence intervals for failure-time cdf F assuming independent lognormal risks. Plot (a) is for deaths due to cancer. Plot (b) is for deaths due to other causes

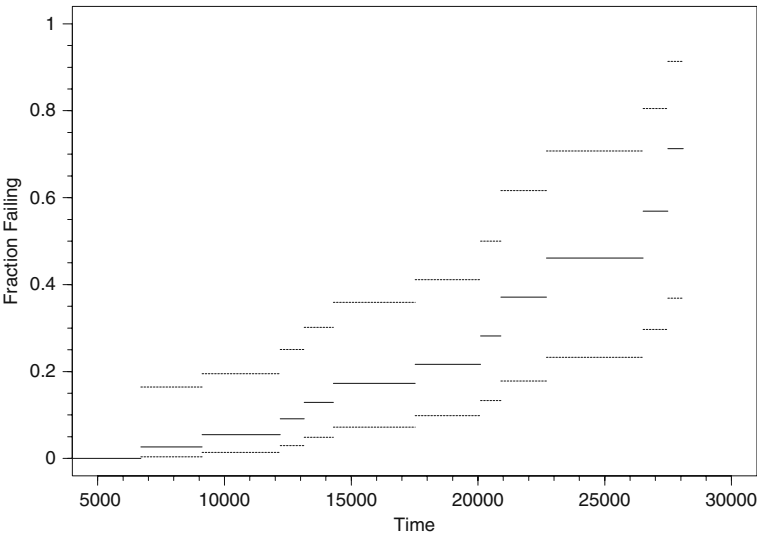


Figure 26.6. Plot of the KM estimate of the cumulative distribution function F with pointwise confidence intervals for the shock absorber data set

Table 26.2. Maximum likelihood results for the shock absorber data set

Distributions		Maximum loglikelihood			Correlations	
M1	M2	M1	M2	Total	M1	M2
Lognormal	Lognormal	-14.31	-10.98	-25.29	0.9756	0.9903
Lognormal	Weibull	-14.31	-11.18	-25.49	0.9833	0.9919
Weibull	Lognormal	-13.50	-10.98	-24.48	0.9893	0.9934
Weibull	Weibull	-13.50	-11.18	-24.68	0.9930	0.9944

Table 26.2 provides the maximum likelihood values and the competing risk probability plot correlations achieved from fitting models to the data set. All four total loglikelihood values are similar, but the best one is from Weibull–lognormal followed by Weibull–Weibull. These two model M2 failures differently, but their M2 loglikelihood values (column 4) are similar. Weibull–Weibull provides correlations for M1 and M2 failures that are higher than those under Weibull–lognormal. Competing risk probability plots for these models are given in Figures 26.7 and 26.8, respectively. These plots and correlations suggest that either model provides an adequate description of shock absorber failure times.

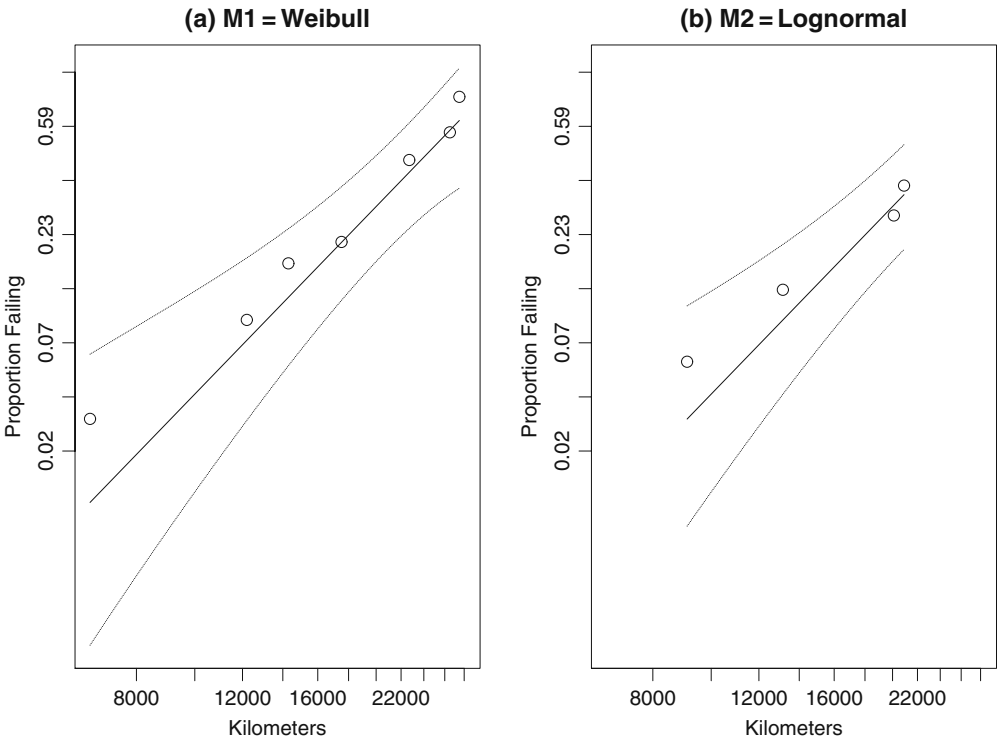


Figure 26.7. Probability plots for the shock absorber data set with pointwise confidence intervals for failure-time cdf F assuming independent Weibull M1 failures and lognormal M2 failures. Plot (a) is for failures due to mode M1. Plot (b) is failures due to mode M2

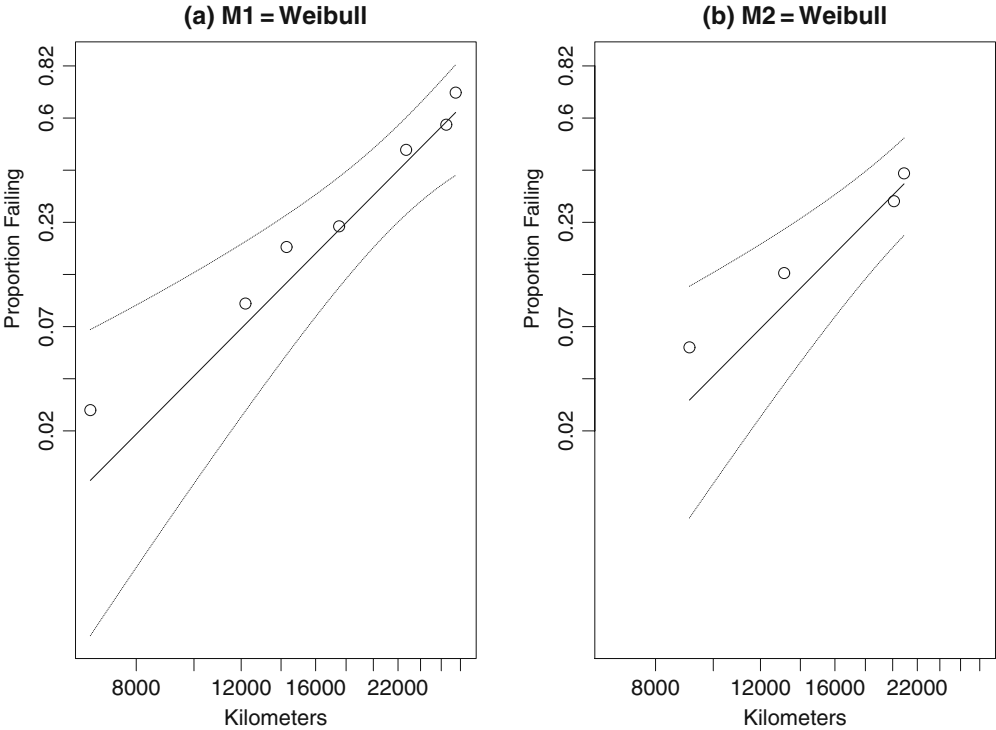


Figure 26.8. Probability plots for the shock absorber data set with pointwise confidence intervals for failure-time cdf F assuming independent Weibull failure times. Plot (a) is for failures due to mode M1. Plot (b) is failures due to mode M2

26.5.3 Simulated data set

We also apply the proposed probability plot on a simulated data set. We assume that risks 1 and 2 are independent and distributed, respectively, lognormal with $\mu_1 = 6$ and $\sigma_1 = 1$ and Weibull with $\mu_2 = 5$ and $\sigma_2 = 0.7$. We generate $n = 100$ data points and censor at time $\exp(6)$. There are $m_1 = 12$ failures due to risk 1, $m_2 = 86$ failures due to risk 2, and $m_3 = 2$ right-censored observations. The maximum loglikelihoods and probability plot correlations from fitting the four combinations of lognormal and Weibull are given in Table 26.3. Probability plots for the lognormal–

Table 26.3. Maximum likelihood results for a simulated data set assuming that risk 1 and risk 2 are independent and distributed, respectively, lognormal with $\mu_1 = 6$ and $\sigma_1 = 1$ and Weibull with $\mu_2 = 5$ and $\sigma_2 = 0.7$

Distributions		Maximum loglikelihood		Correlations	
Risk 1	Risk 2	Risk 1	Risk 2	Total	Risk 1 Risk 2
Lognormal	Lognormal	−33.62	−115.71	−149.33	0.9955 0.9890
Lognormal	Weibull	−33.62	−116.37	−149.99	0.9911 0.9941
Weibull	Lognormal	−35.39	−115.71	−151.10	0.9954 0.9895
Weibull	Weibull	−35.39	−116.37	−151.76	0.9902 0.9936

lognormal and lognormal–Weibull combinations are given in Figures 26.9 and 26.10. Based on the loglikelihood values, the best models are lognormal–lognormal followed closely by lognormal–Weibull. The correlation values do not suggest a better fitting model between the two. However, the risk 2 probability plot for lognormal–lognormal (Figure 26.9) indicates a lack of fit in the lower tail, while that for the true model lognormal–Weibull (Figure 26.10) shows a better fit.

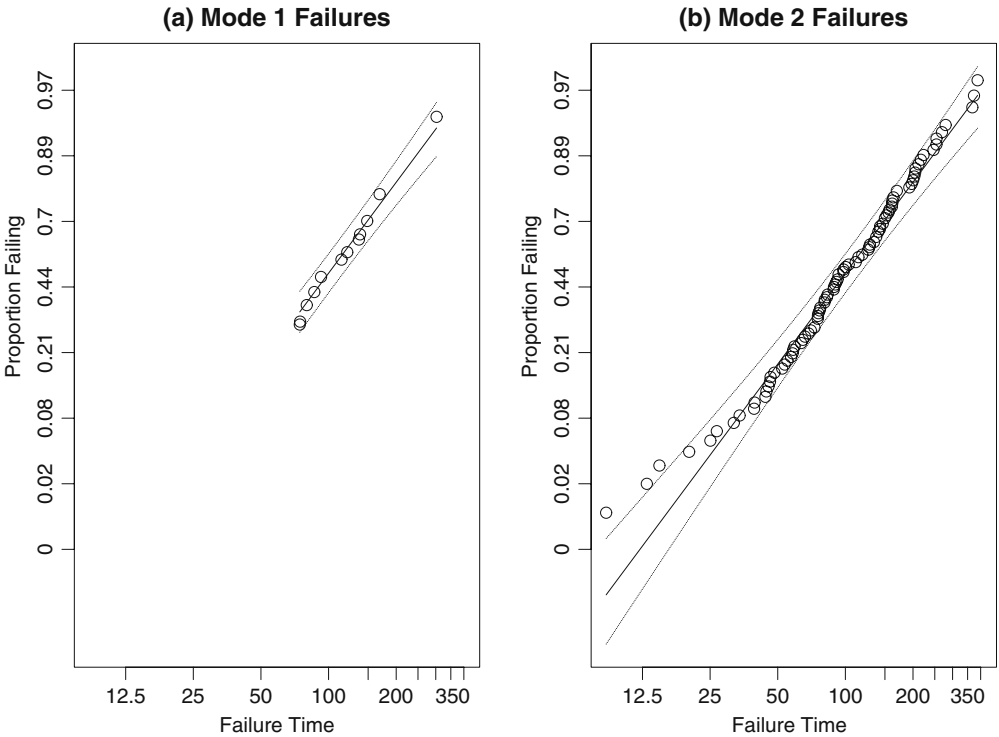


Figure 26.9. Probability plots for the simulated data set (lognormal risk 1, Weibull risk 2) with pointwise confidence intervals for failure-time cdf F assuming independent lognormal risks. Plot (a) is for failures due to risk 1. Plot (b) is failures due to risk 2

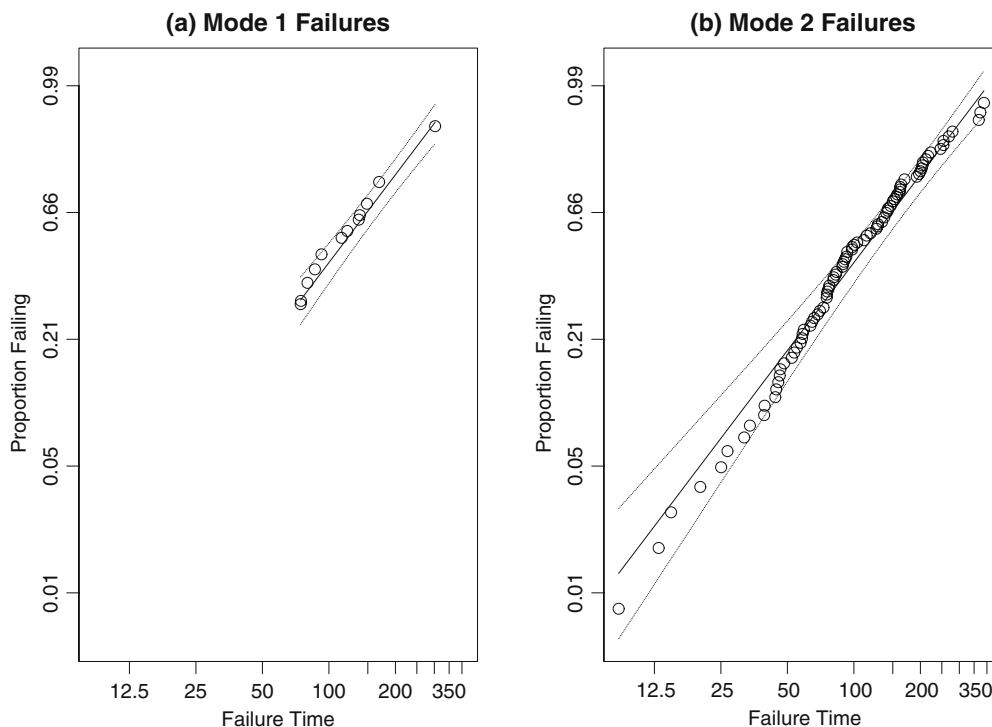


Figure 26.10. Probability plots for the simulated data set (lognormal risk 1, Weibull risk 2) with pointwise confidence intervals for failure-time cdf F assuming independent lognormal risk 1 failures and Weibull risk 2 failures. Plot (a) is for failures due to risk 1. Plot (b) is failures due to risk 2

26.6 Conclusions

The past approach for probability plotting competing risk data is to focus on a specific risk, treat failures due to other risks as right censored, and construct the usual single-distribution probability plot. Under this strategy, there is loss of information by ignoring the causes of failure, and the cdf $F(t)$ of the first time to failure does not usually plot as a straight line on the axes. In this chapter, we propose an alternative plotting method so that the cdf $F(t)$ is a straight line on the axes. We also provide expressions for computing large-sample pointwise confidence intervals for $F(t)$. Primary advantages of our technique include (a) simpler interpretation and assessment of model fit to data and (b) use of all information available from the data (the cause of failure, in particular) for constructing CIs for the cdf $F(t)$.

The proposed method here can be used to assess how well combinations of independent (log) location-scale distributions describe competing risk behavior. This can also be used as a graphical diagnostic in the analysis of competing risk data from accelerated life testing (ALT). A plot is constructed at each level of the accelerating variable to assess fit at individual conditions, and, taken together, these plots could give an

overall sense of model adequacy. The competing risk ALT plans discussed in [17–19], and [20] require an assumed model. Thus, if data are available at the planning stage, then the proposed method here can be applied to determine candidate models.

For future research, it would be worthwhile to extend the methods to the case of dependent risks. For risks which are lognormally distributed, the multivariate normal distribution provides a straightforward way to describe risk dependencies. We expect identifiability issues with estimating the correlation coefficients when the practitioner is unable to obtain times to failure for two or more risks from each subject. This is certainly true in medical studies. One possible way to address this issue is to assume with justification that the correlation coefficients are known. The time-to-failure cdf in this case can also be written in the form (26.7). There are, however, certain engineering applications where risk failure times can be observed for two or more risks. In Chapter 7 of [15], the author describes how Class-H and Class-B insulations for motorettes can fail due to phase, ground, or turn failures. Because these failures occur at separate parts of the product, the occurrence of one does not necessarily preclude those of the other two. The author presents data sets with as much as three risk failure times reported per motorette. The dependent Weibull case is not as straightforward due to the lack of a unifying approach to model dependent Weibull variables. We expect that probability plotting procedures will vary with the model.

References

1. Boag, J.W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of Royal Statistical Society, Series B*, **11**, 15–53.
2. Cox, D.R. (1959). The analysis of exponentially distributed lifetimes with two types of failures. *Journal of the Royal Statistical Society, Series B*, **21**, 411–421.
3. Craiu, R.V. and Lee, T.C.M. (2005). Model selection for the competing-risks model with and without masking. *Technometrics*, **47**, 457–467.
4. Crowder, M.J., Kimber, A.C., Smith, R.L., and Sweeting, T.J. (1991). *Statistical Analysis of Reliability Data*. Chapman and Hall, New York.
5. David, H.A. and Moeschberger, M.L. (1978). *The Theory of Competing Risks*. Macmillan Publishing Co., Inc., New York.
6. Herman, R.J. and Pattel, R.K.N. (1971). Maximum likelihood estimation for multi-risk model. *Technometrics*, **13**, 385–396.
7. Ishioka, T. and Nonaka, Y. (1991). Maximum likelihood estimation of Weibull parameters for two independent competing risks. *IEEE Transactions on Reliability*, **40**, 71–74.
8. Jiang, R. and Murthy, D.N.P. (2003). Study of n -fold Weibull competing risk model. *Mathematical and Computer Modelling*, **38**, 1259–1273.
9. Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.
10. Kim, C.M. and Bai, D.S. (2002). Analyses of accelerated life test data under two failure modes. *International Journal of Reliability, Quality and Safety Engineering*, **9**, 111–126.
11. Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*. John Wiley & Sons, New York.
12. Meeker, W.Q. and Escobar, L.A. (2008). Splida Splus life data analysis. <http://www.public.iastate.edu/~splida>
13. Michael, J.R. (1983). The stabilized probability plot. *Biometrika*, **70**, 11–17.

14. Moeschberger, M.L. and David, H.A. (1971). Life tests under competing causes of failure and the theory of competing risks. *Biometrics*, **27**, 909–933.
15. Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*. John Wiley & Sons, New York.
16. O'Connor, P.D.T. (1995). *Practical Reliability Engineering*. John Wiley & Sons, New York.
17. Pascual, F.G. (2007). Accelerated life test planning with independent Weibull competing risks with known shape parameter. *IEEE Transactions on Reliability*, **56**, 85–93.
18. Pascual, F.G. (2008). Errata to accelerated life test planning with independent Weibull competing risks with known shape parameter. *IEEE Transactions on Reliability*, **57**, 531–532.
19. Pascual, F.G. (2008). Accelerated life test planning with independent Weibull competing risks. *IEEE Transactions on Reliability*, **57**, 435–444.
20. Pascual, F.G. (2009). Accelerated life test planning with independent lognormal competing risks. Under review.

Index

- accelerated life testing, 235
- accelerated model, 323
- accelerated testing, 3, 311, 355
- AFT model, 293
- asymptotic properties, 77

- ball bearings, 253
- Bayesian approach, 157
- bayesian methods, 3
- bootstrap method, 355

- censored data, 27, 371
- censoring, 323
- competing risks, 397
- compound Poisson process, 195
- confidence limits, 27
- confident interval, 293
- contamination, 323
- convergence, 59
- counting processes, 339
- covariate process, 209
- covariates, 275
- Cox model, 323
- crack growth, 241
- cumulative damage, 241
- cumulative exposure model, 355
- cumulative shock, 59

- defect failure, 27
- defect growth, 27
- defect initiation, 27
- defined failure, 27
- degradation, 105, 117, 181, 241, 275
- degradation data, 3
- degradation data analysis, 157
- degradation model, 27, 147
- degradation models, 235
- degradation process, 127, 209
- degradation processes, 17
- degradation-threshold-shock model, 209
- demonstration of system reliability, 311
- deterioration, 241
- deterioration process, 195
- deterministic chaos, 383
- doubly stochastic Poisson process, 77

- estimation of density, 371
- exact distribution, 59
- exponential and Weibull distributions, 311
- extended hazard regression model, 253
- extreme shock, 59
- extreme value theory, 241

- failure rate, 241
- failure time, 127, 209, 293
- failures, 275
- false discover rate, 383
- fatigue, 241
- first passage times, 59
- fuzzy information, 235
- fuzzy numbers, 235
- fuzzy probability distributions, 235

- gamma process, 275
- gamma process, Lévy Process, 209
- goodness of fit test, 383

- hazard function, 195, 355
- hazard potential, 241
- hazard rate, 181
- Hellinger consistency, 371
- Hellinger distance, 323

- imperfect repairs, 339

- increasing failure rate, 17
- increasing failure rate average, 17
- increasing threshold function, 195
- influence curve, 323
- initiation time distribution, 27
- insurance, 181
- intensity, 147
- intershock time, 59
- Kaplan-Meier, 253, 323
- lagged effects, 355
- LEYP model, 339
- life distributions, 17
- life-time analysis, 235
- life-times, 235
- likelihood function, 127
- limit distribution, 59
- locally independent increments process, 105
- lognormal distribution, 39
- maximum-likelihood estimation, 77
- maintenance and replacement policies, 17
- marked point process, 77, 209
- marker processes, 241
- Markovian shock model, 59
- maximum likelihood, 3, 293
- maximum likelihood estimation, 39, 127, 209, 339, 397
- maximum likelihood fitting, 27
- mixed shock, 59
- moment estimation, 77
- moment estimator, 209
- moments, 59
- multiple failure modes, 3
- natural and traumatic failures, 147
- neighborhoods, 323
- non-monotone, 181
- nonparametric estimation, 3, 253, 293
- optimal Markov time, 181
- parameter estimation, 127, 293
- parametric, 147
- parametric estimation, 253
- parametric models, 77
- path models, 275
- piecewise deterministic Markov process, 105
- Poisson approximation, 105
- Poisson process, 117
- position-dependent marking, 77
- probability plotting, 397
- process observation at discrete time points, 127
- product design, 3
- Prohorov distance, 323
- prophylaxis, 181
- random resource, 117
- recurrence data, 3
- recurrent events, 339
- redundant system, 293, 294
- regression, 253
- regression model, 27
- reliability, 105, 157, 241, 253, 293
- reliability function, 195
- renewal theory, 59
- repairable item, 209
- reparable systems, 339
- residuals, 27
- right censoring, 397
- risk of error, 311
- robustness, 323
- Sedyakin's model, 293
- semi parametric estimation, 253
- semi-Markov process, 105
- semi-parametric and non-parametric estimation methods, 147
- semimartingale, 105
- semiparametric estimation, 209
- shock, 59
- shock model, 77
- shock process, 275
- shocks, 117
- short trajectory, 383
- shot noise process, 117
- singular perturbation, 105
- statistical software, 3
- step-stress model, 355
- stochastic processes, 241
- stopped random walk, 59
- survival function, 147
- time dependent covariates, 339
- time scale, 209
- traumatic event, 209
- traumatic failure, 275
- varying critical load, 59
- warm stand-by unit, 293
- warranty data, 3
- weak convergence, 105
- weakest link principle, 241
- wear, 275
- Weibull distribution, 39, 293
- Wiener process, 127, 209, 241