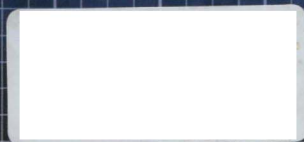


MATHEMATICS AND STATISTICS SERIES



Statistical Models and Methods for Reliability and Survival Analysis

Edited by

**Vincent Couallier, Léo Gerville-Réache
Catherine Huber-Carol, Nikolaos Limnios
and Mounir Mesbah**

ISTE

WILEY

Statistical Models and Methods for Reliability and Survival Analysis

Edited by

Vincent Couallier

Léo Gerville-Réache

Catherine Huber-Carol

Nikolaos Limnios

Mounir Mesbah

Series Editor

Nikolaos Limnios

ISTE

WILEY

First published 2014 in Great Britain and the United States by ISTE Ltd and John Wiley & Sons, Inc.

for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms and licenses issued by the CLA. Enquiries concerning reproduction outside these terms should be sent to the publishers at the undermentioned address:

ISTE Ltd
27-37 St George's Road
London SW19 4EU
UK

www.iste.co.uk

John Wiley & Sons, Inc.
111 River Street
Hoboken, NJ 07030
USA

www.wiley.com

© ISTE Ltd 2014

The rights of Vincent Couallier, Léo Gerville-Réache, Catherine Huber-Carol, Nikolaos Limnios and Mounir Mesbah to be identified as the authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

Library of Congress Control Number: 2013951426

British Library Cataloguing-in-Publication Data
A CIP record for this book is available from the British Library
ISBN: 978-1-84821-619-8



Printed and bound in Great Britain by CPI Group (UK) Ltd., Croydon, Surrey CR0 4YY

Statistical Models and Methods
for Reliability and Survival Analysis

Preface

Statistical Models and Methods for Reliability and Survival Analysis is a volume of contributions by specialists in statistical theory and their applications that provide up-to-date developments in methods used in survival analysis, statistical goodness of fit, stochastic processes for system reliability among others. Many of them are related to the work of Professor M. Nikulin in statistics which has spanned 30 years. The contributors accepted this challenging project to gather various contributions with a wide range of techniques and results, all of them on the topics of the past S2MRSA conference, 4–6th July 2012, dedicated to M. Nikulin for his 20th anniversary as a professor in the Bordeaux Segalen University (<http://www.sm.u-bordeaux2.fr/Stat.Bordeaux.2012>). The book is intended for researchers interested in statistical methodology and models useful in survival analysis, system reliability and statistical testing for censored and non-censored data.

Vincent COUALLIER
Léo GERVILLE-RÉACHE
Catherine HUBER-CAROL
Nikolaos LIMNIOS
Mounir MESBAH
October 2013

Biography of Mikhail Stepanovitch Nikouline

From his native city of St. Petersburg, whose name was Leningrad when he was born there on April 29, 1944, Mikhail Stepanovitch Nikulin (Nikouline by French passport) visited many foreign countries due to his interest in several different domains in his favorite disciplines: mathematics and statistics. It was Professor Login N. Bolshev who allowed him to be aware of his talent in mathematics and directed his thesis in the field of “Probability Theory and Mathematical Statistics”.

He obtained his Master’s degree in mathematics from the State University in St. Petersburg in 1966. In the meanwhile, he had met Valia, a young lady who had many admirers, and he was proud and happy to be the one she chose to be her husband. They married in early 1963 when they were both under 20. Valia was a scientist like her husband and they had a daughter named H  l  ne at the end of the same year.

Then they decided to learn French intensively in order to go to Brazzaville, capital of the “R  publique Populaire du Congo”, where they spent three years. There, both were professors of mathematics in the *Coll  ge d’Enseignement G  n  ral* from 1966 to 1969. Just after leaving Congo, their second child Alexis was born; both children were very promising. Then they went back to their native country. There, M.S. Nikouline became a PhD student at the Steklov Mathematical Institute in Moscow, Academy of Science of Russia, while Valia was doing research in mathematics at the Department of Applied Mathematics at the State University in St. Petersburg. Under the supervision of Professor L.N. Bolshev, for whom Nikouline always had a great admiration for both his mathematical skill and his personal human qualities, and whose early death at the age of 56 saddened Nikouline greatly, he obtained his PhD thesis entitled “A generalization of chi square tests”.

From 1974 to 2006, he was a member of the famous Ildar Ibragimov's Laboratory of Statistical Methods at the Mathematical Institute of V. Steklov (St. Petersburg).

From 1988 to 1992, M.S. Nikouline was a docent in the Department of Statistics and Probability at the State University in St. Petersburg. In 1992, he was hired as an Associate Professor at the University Victor Segalen-Bordeaux 2, France, where he spent his career serving in various positions: Professor in 1996, Full Professor in 2006 and Distinguished Professor in 2008 until 2012 when he became Emeritus. During the same time, from 1996 to 2001, he was the head of the Department "SCIMS- Science et Modélisation", and, from 1999 to 2006, he was the head of the Laboratory "Equipe d'Accueil 2961: Mathematical Statistics and its Applications".

He published a large number of papers (more than 250), alone for most of them, or jointly, with several colleagues and friends from the USSR, Europe and Canada, such as V. Solev, V. Bagdonavicius, V.G. Voinov, N.N. Lyashenko, K. Dhzapridze, P. Greenwood, L. Gerville-Réache, J. Kruopis and many others.

His first field of interest, following the theme of his PhD thesis, was goodness-of-fit tests and best unbiased estimators for parametric models. Between 1973 and 1992, he published more than 30 papers on this subject.

From 1993 to 2013, his subjects of interest became more diversified. We can see from the references of Nikulin's papers that although he was interested in parametric models, goodness of fit for them and unbiased estimators, he also tackled semi-parametric and non-parametric models that are more general. In particular, he was interested in human survival data analysis and he was one of the few people who made the link with reliability whose traditional field is industry that seemed for a long time to be very much apart from the medical field. From a mathematical point of view, many advances in any of these two domains could be provided by the other domain. With this perspective in mind, M.S. Nikouline organized with three French colleagues M. Mesbah of the University of Vannes (Britanny) and then University of Paris VI, N. Limnios of the University of Technology of Compiègne and C. Huber-Carol of the University of Paris Descartes, a European Seminar devoted to the "Survival Analysis and Reliability Theory".

During his career in Bordeaux, Nikouline has directed 11 PhD theses and most of his students have obtained good positions either in universities or in industry.

Moreover, Nikouline was the author and editor of a number of volumes on mathematical statistics. He has also written several articles for the *Encyclopaedia* and was part of several editorial boards and scientific societies such as the International Statistical Institute (ISI) and the Mathematical Society of St. Petersburg. Throughout his long career, he has met a large number of colleagues and as he is a friendly and generous person, many of them became his friends and were always very faithful to him.

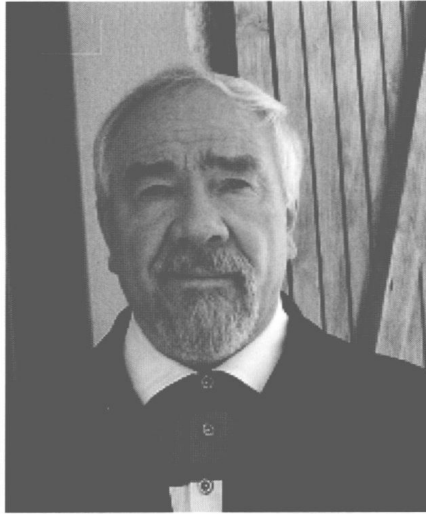


Table of Contents

Preface	xv
Biography of Mikhail Stepanovitch Nikouline	xvii
Vincent COUALLIER, Léo GERVILLE-RÉACHE, Catherine HUBER-CAROL, Nikolaos LIMNIOS and Mounir MESBAH	
PART 1. STATISTICAL MODELS AND METHODS	1
Chapter 1. Unidimensionality, Agreement and Concordance Probability	3
Zhezhen JIN and Mounir MESBAH	
1.1. Introduction	3
1.2. From reliability to unidimensionality: CAC and curve	4
1.2.1. Classical unidimensional models for measurement	4
1.2.2. Reliability of an instrument: CAC	6
1.2.3. Unidimensionality of an instrument: BRC	9
1.3. Agreement between binary outcomes: the kappa coefficient	10
1.3.1. The kappa model	10
1.3.2. The kappa coefficient	10
1.3.3. Estimation of the kappa coefficient	10
1.4. Concordance probability	11
1.4.1. Relationship with Kendall's τ measure	12
1.4.2. Relationship with Somer's D measure	12
1.4.3. Relationship with ROC curve	13
1.5. Estimation and inference	14
1.6. Measure of agreement	14
1.7. Extension to survival data	15
1.7.1. Harrell's c -index	15
1.7.2. Measure of discriminatory power	16

平岸 氏

1.8. Discussion	17
1.9. Bibliography	18
Chapter 2. A Universal Goodness-of-Fit Test Based on Regression Techniques	21
Florence GEORGE and Sneha GULATI	
2.1. Introduction	21
2.2. The Brain and Shapiro procedure for the exponential distribution	22
2.3. Applications of the Brain and Shapiro test	24
2.4. Small sample null distribution of the test statistic for specific distributions	25
2.5. Power studies	28
2.6. Some real examples	28
2.7. Conclusions	31
2.8. Acknowledgment	32
2.9. Bibliography	32
Chapter 3. Entropy-type Goodness-of-Fit Tests for Heavy-Tailed Distributions	33
Andreas MAKRIDES, Alex KARAGRIGORIOU and Filia VONTA	
3.1. Introduction	33
3.2. The entropy test for heavy-tailed distributions	35
3.2.1. Development and asymptotic theory	35
3.2.2. Discussion	39
3.3. Simulation study	40
3.4. Conclusions	42
3.5. Bibliography	42
Chapter 4. Penalized Likelihood Methodology and Frailty Models	45
Emmanouil ANDROULAKIS, Christos KOUKOUVINOS and Filia VONTA	
4.1. Introduction	45
4.2. Penalized likelihood in frailty models for clustered data	48
4.2.1. Gamma distributed frailty	52
4.2.2. Inverse Gaussian distributed frailty	52
4.2.3. Uniform distributed frailty	54
4.3. Simulation results	55
4.4. Concluding remarks	57
4.5. Bibliography	57

Chapter 5. Interactive Investigation of Statistical Regularities in Testing Composite Hypotheses of Goodness of Fit	61
Boris LEMESHKO, Stanislav LEMESHKO and Andrey ROGOZHNIKOV	
5.1. Introduction	61
5.2. Distributions of the test statistics in the case of testing composite hypotheses	63
5.3. Testing composite hypotheses in “real-time”	68
5.4. Conclusions	73
5.5. Acknowledgment	73
5.6. Bibliography	73
Chapter 6. Modeling of Categorical Data	77
Henning LÄUTER	
6.1. Introduction	77
6.2. Continuous conditional distributions	78
6.2.1. Conditional normal distribution	78
6.2.1.1. Estimation of parameters	78
6.2.2. More general continuous conditional distributions	81
6.2.2.1. Conditional distribution	82
6.2.2.2. Normal copula	83
6.3. Discrete conditional distributions	84
6.3.1. Parametric conditional distributions	84
6.3.2. Estimation of parameters	86
6.4. Goodness of fit	86
6.4.1. Distribution of \hat{X}^2	87
6.5. Modeling of categorical data	88
6.5.1. Contingency tables	89
6.5.1.1. General tables	89
6.5.1.2. Further examples	93
6.6. Bibliography	93
Chapter 7. Within the Sample Comparison of Prediction Performance of Models and Submodels: Application to Alzheimer’s Disease	95
Catherine HUBER-CAROL, Shulamith T. GROSS and Annick ALPÉROVITCH	
7.1. Introduction	95
7.2. Framework	96
7.2.1. General description of the data set and the models to be compared	96
7.2.2. Definition of the performance prediction criteria: IDI and BRI	96
7.3. Estimation of IDI and BRI	97
7.3.1. General estimating equations for IDI and BRI	98
7.3.2. Estimation of IDI and BRI in the logistic case	98
7.3.2.1. Asymptotics of $\widehat{IDI}_{2/1}$ for logistic predictors	99

7.3.2.2. Asymptotics of $B\widehat{RI}_{2/1}$ for logistic predictors	100
7.4. Simulation studies	102
7.4.1. First simulation	102
7.4.2. Second simulation: Gu and Pepe's example	104
7.5. The three city study of Alzheimer's disease	106
7.6. Conclusion	108
7.7. Bibliography	109
Chapter 8. Durbin–Knott Components and Transformations of the Cramér-von Mises Test	111
Gennady MARTYNOV	
8.1. Introduction	111
8.2. Weighted Cramér-von Mises statistic	111
8.3. Examples of the Cramér-von Mises statistics	113
8.3.1. Classical Cramér-von Mises statistic	113
8.3.2. Anderson–Darling statistic	113
8.3.3. Cramér-von Mises statistic with the power weight function	114
8.4. Weighted parametric Cramér-von Mises statistic	114
8.4.1. Covariance functions of weighted parametric empirical process	114
8.4.2. Eigenvalues and eigenfunctions for weighted parametric Cramér- von Mises statistic	116
8.5. Transformations of the Cramér-von Mises statistic	117
8.5.1. Preliminary notes	117
8.5.2. Replacement of eigenvalues	118
8.5.3. Transformed statistics	119
8.6. Bibliography	122
Chapter 9. Conditional Inference in Parametric Models	125
Michel BRONIATOWSKI and Virgile CARON	
9.1. Introduction and context	125
9.2. The approximate conditional density of the sample	127
9.2.1. Approximation of conditional densities	127
9.2.2. The proxy of the conditional density of the sample	129
9.2.3. Comments on implementation	131
9.3. Sufficient statistics and approximated conditional density	131
9.3.1. Keeping sufficiency under the proxy density	131
9.3.2. Rao–Blackwellization	132
9.4. Exponential models with nuisance parameters	135
9.4.1. Conditional inference in exponential families	135
9.4.2. Application of conditional sampling to MC tests	137
9.4.2.1. Context	137
9.4.2.2. Bimodal likelihood: testing the mean of a normal distribution in dimension 2	139

9.4.3. Estimation through conditional likelihood	140
9.5. Bibliography	142
Chapter 10. On Testing Stochastic Dominance by Exceedance, Precedence and Other Distribution-Free Tests, with Applications	145
Paul DEHEUVELS	
10.1. Introduction	145
10.2. Results	148
10.2.1. The experimental data set	148
10.2.2. An application of the Wilcoxon–Mann–Whitney statistics	149
10.2.3. One-sided Kolmogorov–Smirnov tests	150
10.2.4. Precedence and Exceedance Tests.	152
10.3. Negative binomial limit laws	155
10.4. Conclusion	159
10.5. Bibliography	159
Chapter 11. Asymptotically Parameter-Free Tests for Ergodic Diffusion Processes	161
Yury A. KUTOYANTS and Li ZHOU 遍历打散过程的 近似无参检验	
11.1. Introduction	161
11.2. Ergodic diffusion process and some limits	165
11.3. Shift parameter	168
11.4. Shift and scale parameters	172
11.5. Bibliography	175
Chapter 12. A Comparison of Homogeneity Tests for Different Alternative Hypotheses	177
Sergey POSTOVALOV and Petr PHILONENKO 齐性检验	
12.1. Homogeneity tests	178
12.1.1. Tests for data without censoring	179
12.1.2. Tests for data with censoring	180
12.2. Alternative hypotheses	184
12.3. Power simulation	185
12.3.1. Power of tests without censoring	187
12.3.2. Power of tests with censoring	189
12.3.2.1. How does the distribution of censoring time affect the power of the test?	189
12.3.2.2. How does the censoring rate affect the power of the test?	191
12.4. Statistical inference	191
12.5. Acknowledgment	192
12.6. Bibliography	193

Chapter 13. Some Asymptotic Results for Exchangeably Weighted Bootstraps of the Empirical Estimator of a Semi-Markov Kernel with Applications 195
 Salim BOUZEBDA and Nikolaos LIMNIOS

13.1. Introduction 195
 13.2. Semi-Markov setting 197
 13.3. Main results 201
 13.4. Bootstrap for a multidimensional empirical estimator of a continuous-time semi-Markov kernel 205
 13.5. Confidence intervals 208
 13.6. Bibliography 210

Chapter 14. On Chi-Squared Goodness-of-Fit Test for Normality 213
 Mikhail NIKULIN, Léo GERVILLE-RÉACHE and Xuan Quang TRAN

14.1. Chi-squared test for normality 213
 14.2. Simulation study 221
 14.3. Bibliography 226

PART 2. STATISTICAL MODELS AND METHODS IN SURVIVAL ANALYSIS 229

Chapter 15. Estimation/Imputation Strategies for Missing Data in Survival Analysis 231
 Elodie BRUNEL, Fabienne COMTE and Agathe GUILLOUX

15.1. Introduction 231
 15.2. Model and strategies 233
 15.2.1. Model assumptions 233
 15.2.2. Strategy involving knowledge of ζ 234
 15.2.3. Strategy involving knowledge of π 235
 15.2.4. Estimation of ζ or π : logit or non-parametric regression 236
 15.2.5. Computing the hazard estimators 236
 15.2.6. Theoretical results 239
 15.3. Imputation-based strategy 241
 15.4. Numerical comparison 242
 15.5. Proofs 244
 15.6. Bibliography 251

Chapter 16. Non-Parametric Estimation of Linear Functionals of a Multivariate Distribution Under Multivariate Censoring with Applications 253
 Olivier LOPEZ and Philippe SAINT-PIERRE

16.1. Introduction 253