Springer Series in Statistics

Advisors:

J. Berger, S. Fienberg, J. Gani,
K. Krickeberg, I. Olkin, B. Singer

Springer Series in Statistics

- Anderson*: Continuous-Time Markov Chains: An Applications-Oriented Approach.
- Andrews/Herzberg*: Data: A Collection of Problems from Many Fields for the Student and Research Worker.
- Anscombe*: Computing in Statistical Science through APL.
- Berger*: Statistical Decision Theory and Bayesian Analysis, 2nd edition.
- Bolfarine/Zacks*: Prediction Theory for Finite Populations.
- Brémaud*: Point Processes and Queues: Martingale Dynamics.
- Brockwell/Davis*: Time Series: Theory and Methods, 2nd edition.
- Daley/Vere-Jones*: An Introduction to the Theory of Point Processes.
- Dzhaparidze*: Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series.
- Farrell*: Multivariate Calculation.
- Fienberg/Hoaglin/Kruskal/Tanur (Eds.)*: A Statistical Model: Frederick Mosteller's Contributions to Statistics, Science, and Public Policy.
- Goodman/Kruskal*: Measures of Association for Cross Classifications.
- Grandell*: Aspects of Risk Theory.
- Hall*: The Bootstrap and Edgeworth Expansion.
- Härdle*: Smoothing Techniques: With Implementation in S.
- Hartigan*: Bayes Theory.
- Heyer*: Theory of Statistical Experiments.
- Jolliffe*: Principal Component Analysis.
- Kotz/Johnson (Eds.)*: Breakthroughs in Statistics Volume I.
- Kotz/Johnson (Eds.)*: Breakthroughs in Statistics Volume II.
- Kres*: Statistical Tables for Multivariate Analysis.
- Leadbetter/Lindgren/Rootzén*: Extremes and Related Properties of Random Sequences and Processes.
- Le Cam*: Asymptotic Methods in Statistical Decision Theory.
- Le Cam/Yang*: Asymptotics in Statistics: Some Basic Concepts.
- Manoukian*: Modern Concepts and Theorems of Mathematical Statistics.
- Miller, Jr.*: Simultaneous Statistical Inference, 2nd edition.
- Mosteller/Wallace*: Applied Bayesian and Classical Inference: The Case of *The Federalist Papers*.
- Pollard*: Convergence of Stochastic Processes.
- Pratt/Gibbons*: Concepts of Nonparametric Theory.
- Read/Cressie*: Goodness-of-Fit Statistics for Discrete Multivariate Data.
- Reiss*: Approximate Distributions of Order Statistics: With Applications to Nonparametric Statistics.
- Ross*: Nonlinear Estimation.
- Sachs*: Applied Statistics: A Handbook of Techniques, 2nd edition.
- Salsburg*: The Use of Restricted Significance Tests in Clinical Trials.
- Särndal/Swensson/Wretman*: Model Assisted Survey Sampling.
- Seneta*: Non-Negative Matrices and Markov Chains.
- Shedler*: Regeneration and Networks of Queues.

(continued after index)

Heleno Bolfarine Shelemyahu Zacks

Prediction Theory for Finite Populations

Springer-Verlag

New York Berlin Heidelberg London Paris
Tokyo Hong Kong Barcelona Budapest

Heleno Bolfarine
Department of Statistics
Institute of Mathematics and Statistics
University of Sao Paulo
Sao Paulo
Brazil

Shelemyahu Zacks
Department of Mathematical Sciences
State University of New York
at Binghamton
Binghamton, NY 13902-6000
USA

Mathematics Subject Classification (1991): 62D05

Library of Congress Cataloging-in-Publication Data
Bolfarine, Heleno.

Prediction theory for finite populations/Heleno Bolfarine,
Shelemyahu Zacks.

p. cm. — (Springer series in statistics)

Includes bibliographical references and index.

ISBN-13:978-1-4612-7713-2

1. Prediction theory. I. Zacks, Shelemyahu.

II. Title. III. Series.

QA279.2.B65 1992

519.5'4—dc20

91-45891

Printed on acid-free paper.

© 1992 Springer-Verlag New York Inc.

Softcover reprint of the hardcover 1st edition 1992

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, 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 of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

Production managed by Natalie Johnson; manufacturing supervised by Robert Paella.
Photocomposed copy prepared from the authors' $\mathcal{A}_{\mathcal{M}}\mathcal{S}$ - $\mathcal{T}_{\mathcal{E}}\mathcal{X}$ files.

9 8 7 6 5 4 3 2 1

ISBN-13:978-1-4612-7713-2

e-ISBN-13:978-1-4612-2904-9

DOI: 10.1007/978-1-4612-2904-9

To our families

Preface

A large number of papers on predicting finite population quantities have appeared in the literature during the last two decades. This book is designed to present current theory in a systematic and consistent manner. An overview of the material in the book is presented in the Synopsis, which guides the reader to the various topics treated. Theoretical examples, which illustrate concepts and provide formulae of predictors for special cases of common interest, are given in each chapter. In addition, a large number of problems for solution accompany each chapter. Many of these problems are based on material found in research papers. Thus, the book can be used in a graduate course and can well prepare students for research in the field. The first draft was written while H. Bolfarine spent his sabbatical year at the Center for Statistics at the State University of New York at Binghamton, New York. We started by translating a short monograph (in Portuguese) written by Rodrigues and Bolfarine (1984). Soon we realized that much more work was required. This monograph turned out to be completely different from the short Portuguese version. We would like to acknowledge the advice and suggestions offered generously by Professor J. Rodrigues. H. Bolfarine would like to acknowledge the financial support of Conselho Nacional de Pesquisas (CNPq-Brasil) and the University of São Paulo, which made the writing of this book possible. Special thanks are due to Mrs. Marge Pratt for her excellent $\text{T}_{\text{E}}\text{X}$ typing of the book. We would like to thank also the State University of New York at Binghamton for providing facilities and our families for their understanding and support.

Heleno Bolfarine
Shelemyahu Zacks

Binghamton, NY
October, 1991

Contents

Preface	vii
Synopsis	1
1. Basic Ideas and Principles	5
1.1. The Fixed Finite Population Model.....	6
1.2. The Superpopulation Model.....	8
1.2.1. The Regression Model.....	9
1.3. Predictors of Population Quantities.....	12
1.4. The Model-Based Design-Based Approach.....	17
1.5. Exercises.....	18
2. Optimal Predictors of Population Quantities	23
2.1. Best Linear Unbiased Predictors.....	23
2.2. Best Unbiased Predictors.....	30
2.3. Equivariant Predictors.....	42
2.3.1. A General Formulation.....	43
2.3.2. Location Equivariant Predictors of T Under Model SM1.....	44
2.3.3. Location Equivariant Predictors of T Under the Regression Model.....	48
2.3.4. Scale Equivariant Predictors of T	49
2.3.5. Location-Scale Equivariant Predictors of T	52
2.3.6. Location-Scale Equivariant Predictors of S_y^2	54
2.4. Stein-Type Shrinkage Predictors.....	60
2.5. Exercises.....	62
3. Bayes and Minimax Predictors	66
3.1. The Multivariate Normal Model.....	66

3.1.1. Bayes Predictors of T	70
3.1.2. Bayes Predictors of β_N	72
3.1.3. Bayes Predictors of S_y^2	73
3.2. Bayes Linear Predictors	76
3.3. Minimax and Admissible Predictors	79
3.4. Dynamic Bayesian Prediction	86
3.4.1. The Multinormal Dynamic Model	87
3.4.1.1. Dynamic Prediction of T_t	87
3.4.1.2. Dynamic Prediction of S_{ty}^2	89
3.5. Empirical Bayes Predictors	91
3.6. Exercises	95
4. Maximum-Likelihood Predictors	99
4.1. Predictive Likelihoods	99
4.1.1. Estimative Predictive Likelihoods	99
4.1.2. Profile Predictive Likelihoods	102
4.1.3. The Lauritzen-Hinkley Predictive Likelihoods	103
4.1.4. The Royall Predictive Likelihoods	104
4.2. Maximum Likelihood Predictors of T Under the Normal Superpopulation Model	105
4.2.1. Estimative Likelihood Predictors	105
4.2.2. Profile Likelihood Predictors	106
4.2.3. The LH Likelihood Predictors	107
4.2.4. The Royall Maximum-Likelihood Predictors	108
4.3. Maximum-Likelihood Predictors of the Population Variance S_y^2 Under the Normal Regression Model	113
4.3.1. Estimative Likelihood Predictors	113
4.3.2. Profile Likelihood Predictors	114
4.3.3. LH Likelihood Predictors	116
4.4. Exercises	116
5. Classical and Bayesian Prediction Intervals	119
5.1. Confidence Prediction Intervals	119
5.2. Tolerance Prediction Intervals for T	122
5.3. Bayesian Prediction Intervals	124
5.4. Exercises	127
6. The Effects of Model Misspecification, Conditions For Robustness, and Bayesian Modeling	128
6.1. Robust Linear Prediction of T	129
6.2. Estimation of the Prediction Variance	133
6.3. Simulation Estimates of the ψ^* MSE of $\hat{T}_R$	135
6.4. Bayesian Robustness	137
6.5. Bayesian Modeling	140
6.5.1. The Framework	140

6.5.2. The Normal Linear Model.....	141
6.6. Exercises.....	145
7. Models with Measurement Errors.....	150
7.1. The Location and Simple Regression Models.....	151
7.1.1. Model SM1.....	151
7.1.2. Simple Regression Model.....	153
7.1.3. Regression Type Predictors.....	154
7.2. Bayesian Models with Measurement Errors.....	159
7.2.1. Model SM1.....	159
7.2.2. Model SM6.....	161
7.2.3. Simple Regression Model.....	163
7.2.4. Orthogonal Transformations.....	165
7.3. Exercises.....	167
8. Asymptotic Properties in Finite Populations.....	169
8.1. Predictors of T	169
8.2. The Asymptotic Distribution of $\hat{\beta}_{s_k}$	172
8.3. The Linear Regression Model with Measurement Errors.....	176
8.4. Exercises.....	180
9. Design Characteristics of Predictors.....	182
9.1. The QR Class of Predictors.....	182
9.2. ADU Predictors.....	184
9.3. Optimal ADU Predictors.....	187
9.4. Exercises.....	191
Glossary of Predictors.....	193
Bibliography.....	196
Author Index.....	203
Subject Index.....	205