

Bootstrap methods are computer-intensive methods of statistical analysis that use simulation to calculate standard errors, confidence intervals and significance tests. The methods are not only in general use by statisticians, but are also applied by quantitative researchers in the life sciences, medical sciences, social sciences, and business. The methods apply for any level of modelling, and so can be used for fully parametric, semiparametric, and completely nonparametric analysis. This book gives a broad and up-to-date coverage of bootstrap methods with numerous applied examples, together with the underlying general concepts and basic theory without emphasis on mathematical rigour, developed in a coherent way with the necessary theoretical basis.

Applications include stratified data; finite populations; censored and missing data; linear, nonlinear, and smooth regression models; time series and spatial confidence intervals; material on various diagnostic methods; and methods for efficient computation, including improved Monte Carlo simulation. Each chapter includes both practical and theoretical exercises. A valuable supplement is a website of purpose-written S-Plus programs, written by A. J. Canty, for implementing the methods described in the text.

The emphasis on practicalities, with detailed examples of how to apply the methods, will make it indispensable to users, teachers and students of statistical methods.

"A comprehensive and extremely readable overview of the current state of art in bootstrap methodology. Through the numerous exercises, practicals and examples the reader obtains a good understanding for the strength of bootstrap methods, the problems for which they work and how to avoid their pitfalls. I strongly recommend this book." B. Turlach, *Journal of Applied Statistics*

"Statisticians with little or no familiarity with the bootstrap will find [this book] to be a thorough introduction to its use in solving real-world problems. Practical advice . . . is given regularly throughout. . . . We recommend this book most highly. It made us stop and think regularly and contributed tremendously to our understanding of the bootstrap. It is an excellent book for professors, students, practitioners, and researchers alike." T. Loughin and C. Bilder, *JASA*

Cambridge Series in Statistical and Probabilistic Mathematics

Editorial Board:

R. Gill (Department of Mathematics, Utrecht University)

B. D. Ripley (Department of Statistics, University of Oxford)

S. Ross (Department of Industrial Engineering, University of California, Berkeley)

M. Stein (Department of Statistics, University of Chicago)

D. Williams (School of Mathematical Sciences, University of Bath)

CAMBRIDGE
UNIVERSITY PRESS

www.cambridge.org

ISBN 978-0-521-57471-6



9 780521 574716 >

Contents

<i>Preface</i>	ix
1 Introduction	1
2 The Basic Bootstraps	11
2.1 Introduction	11
2.2 Parametric Simulation	15
2.3 Nonparametric Simulation	22
2.4 Simple Confidence Intervals	27
2.5 Reducing Error	31
2.6 Statistical Issues	37
2.7 Nonparametric Approximations for Variance and Bias	45
2.8 Subsampling Methods	55
2.9 Bibliographic Notes	59
2.10 Problems	60
2.11 Practicals	66
3 Further Ideas	70
3.1 Introduction	70
3.2 Several Samples	71
3.3 Semiparametric Models	77
3.4 Smooth Estimates of F	79
3.5 Censoring	82
3.6 Missing Data	88
3.7 Finite Population Sampling	92
3.8 Hierarchical Data	100
3.9 Bootstrapping the Bootstrap	103

3.10	Bootstrap Diagnostics	113
3.11	Choice of Estimator from the Data	120
3.12	Bibliographic Notes	123
3.13	Problems	126
3.14	Practicals	131
4	Tests	136
4.1	Introduction	136
4.2	Resampling for Parametric Tests	140
4.3	Nonparametric Permutation Tests	156
4.4	Nonparametric Bootstrap Tests	161
4.5	Adjusted P-values	175
4.6	Estimating Properties of Tests	180
4.7	Bibliographic Notes	183
4.8	Problems	184
4.9	Practicals	187
5	Confidence Intervals	191
5.1	Introduction	191
5.2	Basic Confidence Limit Methods	193
5.3	Percentile Methods	202
5.4	Theoretical Comparison of Methods	211
5.5	Inversion of Significance Tests	220
5.6	Double Bootstrap Methods	223
5.7	Empirical Comparison of Bootstrap Methods	230
5.8	Multiparameter Methods	231
5.9	Conditional Confidence Regions	238
5.10	Prediction	243
5.11	Bibliographic Notes	246
5.12	Problems	247
5.13	Practicals	251
6	Linear Regression	256
6.1	Introduction	256
6.2	Least Squares Linear Regression	257
6.3	Multiple Linear Regression	273
6.4	Aggregate Prediction Error and Variable Selection	290
6.5	Robust Regression	307
6.6	Bibliographic Notes	315
6.7	Problems	316
6.8	Practicals	321

7	Further Topics in Regression	326
7.1	Introduction	326
7.2	Generalized Linear Models	327
7.3	Survival Data	346
7.4	Other Nonlinear Models	353
7.5	Misclassification Error	358
7.6	Nonparametric Regression	362
7.7	Bibliographic Notes	374
7.8	Problems	376
7.9	Practicals	378
8	Complex Dependence	385
8.1	Introduction	385
8.2	Time Series	385
8.3	Point Processes	415
8.4	Bibliographic Notes	426
8.5	Problems	428
8.6	Practicals	432
9	Improved Calculation	437
9.1	Introduction	437
9.2	Balanced Bootstraps	438
9.3	Control Methods	446
9.4	Importance Resampling	450
9.5	Saddlepoint Approximation	466
9.6	Bibliographic Notes	485
9.7	Problems	487
9.8	Practicals	494
10	Semiparametric Likelihood Inference	499
10.1	Likelihood	499
10.2	Multinomial-Based Likelihoods	500
10.3	Bootstrap Likelihood	507
10.4	Likelihood Based on Confidence Sets	509
10.5	Bayesian Bootstraps	512
10.6	Bibliographic Notes	514
10.7	Problems	516
10.8	Practicals	519

11 Computer Implementation	522
11.1 Introduction	522
11.2 Basic Bootstraps	525
11.3 Further Ideas	531
11.4 Tests	534
11.5 Confidence Intervals	536
11.6 Linear Regression	537
11.7 Further Topics in Regression	540
11.8 Time Series	543
11.9 Improved Simulation	545
11.10 Semiparametric Likelihoods	549
 Appendix A. Cumulant Calculations	 551
<i>Bibliography</i>	555
<i>Name Index</i>	568
<i>Example index</i>	572
<i>Subject index</i>	575