

Jun Shao Dongsheng Tu

The Jackknife and Bootstrap



Springer

Jun Shao
Department of Statistics
University of Wisconsin, Madison
1210 West Dayton Street
Madison, WI 53706-1685 USA

Dongsheng Tu
Institute of System Science
Academia Sinica
Beijing, 100080
People's Republic of China

With 4 figures.

Library of Congress Cataloging-in-Publication Data
Shao, Jun.

The jackknife and bootstrap / Jun Shao, Dongsheng Tu.

p. cm. — (Springer series in statistics)

Includes bibliographical references and index.

ISBN 978-1-4612-6903-8 ISBN 978-1-4612-0795-5 (eBook)

DOI 10.1007/978-1-4612-0795-5

1. Jackknife (Statistics). 2. Bootstrap (Statistics).
3. Resampling (Statistics). 4. Estimation theory. I. Tu,
Dongsheng. II. Title. III. Series.

QA276.6.S46 1995

519.5'44—dc20

95-15074

Printed on acid-free paper.

© 1995 Springer Science+Business Media New York

Originally published by Springer-Verlag New York, Inc in 1995

Softcover reprint of the hardcover 1st edition 1995

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher Springer Science+Business Media, LLC, 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 Hal Henglein; manufacturing supervised by Joe Quatela.
Photocomposed pages prepared from the authors' LaTeX file.

9 8 7 6 5 4 3 2 (Second corrected printing, 1996)

ISBN 978-1-4612-6903-8

SPIN 10544454

Contents

Preface	vii
Chapter 1. Introduction	1
1.1 Statistics and Their Sampling Distributions	1
1.2 The Traditional Approach	2
1.3 The Jackknife	4
1.4 The Bootstrap	9
1.5 Extensions to Complex Problems	17
1.6 Scope of Our Studies	19
Chapter 2. Theory for the Jackknife	23
2.1 Variance Estimation for Functions of Means	23
2.1.1 Consistency	24
2.1.2 Other properties	28
2.1.3 Discussions and examples	29
2.2 Variance Estimation for Functionals	32
2.2.1 Differentiability and consistency	33
2.2.2 Examples	37
2.2.3 Convergence rate	42
2.2.4 Other differential approaches	44
2.3 The Delete-d Jackknife	49
2.3.1 Variance estimation	50
2.3.2 Jackknife histograms	55
2.4 Other Applications	60
2.4.1 Bias estimation	61
2.4.2 Bias reduction	64

2.4.3	Miscellaneous results	68
2.5	Conclusions and Discussions	69
Chapter 3.	Theory for the Bootstrap	71
3.1	Techniques in Proving Consistency	72
3.1.1	Bootstrap distribution estimators	72
3.1.2	Mallows' distance	73
3.1.3	Berry-Esséen's inequality	74
3.1.4	Imitation	76
3.1.5	Linearization	78
3.1.6	Convergence in moments	79
3.2	Consistency: Some Major Results	80
3.2.1	Distribution estimators	80
3.2.2	Variance estimators	86
3.3	Accuracy and Asymptotic Comparisons	91
3.3.1	Convergence rate	91
3.3.2	Asymptotic minimaxity	97
3.3.3	Asymptotic mean squared error	99
3.3.4	Asymptotic relative error	102
3.3.5	Conclusions	104
3.4	Fixed Sample Performance	104
3.4.1	Moment estimators	105
3.4.2	Distribution estimators	108
3.4.3	Conclusions	112
3.5	Smoothed Bootstrap	113
3.5.1	Empirical evidences and examples	113
3.5.2	Sample quantiles	116
3.5.3	Remarks	117
3.6	Nonregular Cases	118
3.7	Conclusions and Discussions	127
Chapter 4.	Bootstrap Confidence Sets and Hypothesis Tests	129
4.1	Bootstrap Confidence Sets	129
4.1.1	The bootstrap-t	131
4.1.2	The bootstrap percentile	132

4.1.3	The bootstrap bias-corrected percentile	133
4.1.4	The bootstrap accelerated bias-corrected percentile	135
4.1.5	The hybrid bootstrap	140
4.2	Asymptotic Theory	141
4.2.1	Consistency	141
4.2.2	Accuracy	144
4.2.3	Other asymptotic comparisons	152
4.3	The Iterative Bootstrap and Other Methods	155
4.3.1	The iterative bootstrap	155
4.3.2	Bootstrap calibrating	160
4.3.3	The automatic percentile and variance stabilizing	161
4.3.4	Fixed width bootstrap confidence intervals	164
4.3.5	Likelihood based bootstrap confidence sets	165
4.4	Empirical Comparisons	166
4.4.1	The bootstrap-t, percentile, BC, and BC_a	166
4.4.2	The bootstrap and other asymptotic methods	170
4.4.3	The iterative bootstrap and bootstrap calibration	173
4.4.4	Summary	176
4.5	Bootstrap Hypothesis Tests	176
4.5.1	General description	177
4.5.2	Two-sided hypotheses with nuisance parameters	179
4.5.3	Bootstrap distance tests	182
4.5.4	Other results and discussions	184
4.6	Conclusions and Discussions	188
Chapter 5. Computational Methods		190
5.1	The Delete-1 Jackknife	190
5.1.1	The one-step jackknife	191
5.1.2	Grouping and random subsampling	195
5.2	The Delete-d Jackknife	197
5.2.1	Balanced subsampling	197
5.2.2	Random subsampling	198
5.3	Analytic Approaches for the Bootstrap	200
5.3.1	The delta method	201
5.3.2	Jackknife approximations	202

5.3.3	Saddle point approximations	203
5.3.4	Remarks	205
5.4	Simulation Approaches for the Bootstrap	206
5.4.1	The simple Monte Carlo method	207
5.4.2	Balanced bootstrap resampling	211
5.4.3	Centering after Monte Carlo	215
5.4.4	The linear bootstrap	219
5.4.5	Antithetic bootstrap resampling	221
5.4.6	Importance bootstrap resampling	223
5.4.7	The one-step bootstrap	228
5.5	Conclusions and Discussions	230
Chapter 6.	Applications to Sample Surveys	232
6.1	Sampling Designs and Estimates	232
6.2	Resampling Methods	238
6.2.1	The jackknife	238
6.2.2	The balanced repeated replication	241
6.2.3	Approximated BRR methods	244
6.2.4	The bootstrap	246
6.3	Comparisons by Simulation	251
6.4	Asymptotic Results	258
6.4.1	Assumptions	258
6.4.2	The jackknife and BRR for functions of averages	260
6.4.3	The RGBRR and RSBRR for functions of averages	264
6.4.4	The bootstrap for functions of averages	267
6.4.5	The BRR and bootstrap for sample quantiles	268
6.5	Resampling Under Imputation	270
6.5.1	Hot deck imputation	271
6.5.2	An adjusted jackknife	273
6.5.3	Multiple bootstrap hot deck imputation	277
6.5.4	Bootstrapping under imputation	278
6.6	Conclusions and Discussions	281
Chapter 7.	Applications to Linear Models	283
7.1	Linear Models and Regression Estimates	283
7.2	Variance and Bias Estimation	285

7.2.1	Weighted and unweighted jackknives	285
7.2.2	Three types of bootstraps	289
7.2.3	Robustness and efficiency	292
7.3	Inference and Prediction Using the Bootstrap	295
7.3.1	Confidence sets	295
7.3.2	Simultaneous confidence intervals	298
7.3.3	Hypothesis tests	301
7.3.4	Prediction	303
7.4	Model Selection	306
7.4.1	Cross-validation	307
7.4.2	The bootstrap	311
7.5	Asymptotic Theory	313
7.5.1	Variance estimators	313
7.5.2	Bias estimators	318
7.5.3	Bootstrap distribution estimators	320
7.5.4	Inference and prediction	324
7.5.5	Model selection	326
7.6	Conclusions and Discussions	329

Chapter 8. Applications to Nonlinear, Nonparametric, and Multivariate Models 331

8.1	Nonlinear Regression	331
8.1.1	Jackknife variance estimators	333
8.1.2	Bootstrap distributions and confidence sets	335
8.1.3	Cross-validation for model selection	337
8.2	Generalized Linear Models	338
8.2.1	Jackknife variance estimators	340
8.2.2	Bootstrap procedures	341
8.2.3	Model selection by bootstrapping	343
8.3	Cox's Regression Models	345
8.3.1	Jackknife variance estimators	346
8.3.2	Bootstrap procedures	349
8.4	Kernel Density Estimation	350
8.4.1	Bandwidth selection by cross-validation	351
8.4.2	Bandwidth selection by bootstrapping	353
8.4.3	Bootstrap confidence sets	356

8.5	Nonparametric Regression	360
8.5.1	Kernel estimates for fixed design	360
8.5.2	Kernel estimates for random regressor	364
8.5.3	Nearest neighbor estimates	366
8.5.4	Smoothing splines	370
8.6	Multivariate Analysis	373
8.6.1	Analysis of covariance matrix	373
8.6.2	Multivariate linear models	376
8.6.3	Discriminant analysis	379
8.6.4	Factor analysis and clustering	382
8.7	Conclusions and Discussions	384

Chapter 9. Applications to Time Series and Other Dependent Data 386

9.1	m-Dependent Data	387
9.2	Markov Chains	392
9.3	Autoregressive Time Series	394
9.3.1	Bootstrapping residuals	395
9.3.2	Model selection	397
9.4	Other Time Series	400
9.4.1	ARMA(p, q) models	401
9.4.2	Linear regression with time series errors	403
9.4.3	Dynamical linear regression	406
9.5	Stationary Processes	407
9.5.1	Moving block and circular block	407
9.5.2	Consistency of the bootstrap	410
9.5.3	Accuracy of the bootstrap	411
9.5.4	Remarks	413
9.6	Conclusions and Discussions	414

Chapter 10. Bayesian Bootstrap and Random Weighting 416

10.1	Bayesian Bootstrap	416
10.1.1	Bayesian bootstrap with a noninformative prior	417
10.1.2	Bayesian bootstrap using prior information	420
10.1.3	The weighted likelihood bootstrap	422
10.1.4	Some remarks	424

10.2 Random Weighting	425
10.2.1 Motivation	425
10.2.2 Consistency	427
10.2.3 Asymptotic accuracy	429
10.3 Random Weighting for Functionals and Linear Models . . .	434
10.3.1 Statistical functionals	434
10.3.2 Linear models	437
10.4 Empirical Results for Random Weighting	440
10.5 Conclusions and Discussions	445
Appendix A. Asymptotic Results	447
A.1 Modes of Convergence	447
A.2 Convergence of Transformations	448
A.3 $O(\cdot)$, $o(\cdot)$, and Stochastic $O(\cdot)$, $o(\cdot)$	448
A.4 The Borel-Cantelli Lemma	449
A.5 The Law of Large Numbers	449
A.6 The Law of the Iterated Logarithm	450
A.7 Uniform Integrability	450
A.8 The Central Limit Theorem	451
A.9 The Berry-Esséen Theorem	451
A.10 Edgeworth Expansions	452
A.11 Cornish-Fisher Expansions	454
Appendix B. Notation	455
References	457
Author Index	493
Subject Index	499