

# **Springer Texts in Statistics**

## **Series Editors**

Genevera I. Allen, Department of Statistics, Houston, TX, USA

Rebecca Nugent, Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, USA

Richard D. De Veaux, Department of Mathematics and Statistics, Williams College, Williamstown, MA, USA

*Springer Texts in Statistics (STS)* includes advanced textbooks from 3rd- to 4th-year undergraduate courses to 1st- to 2nd-year graduate courses. Exercise sets should be included. The series editors are currently Genevera I. Allen, Richard D. De Veaux, and Rebecca Nugent. Stephen Fienberg, George Casella, and Ingram Olkin were editors of the series for many years.

More information about this series at <http://www.springer.com/series/417>

Bing Li · G. Jogesh Babu

# A Graduate Course on Statistical Inference

Bing Li  
Department of Statistics  
Penn State University  
University Park, PA, USA

G. Jogesh Babu  
Department of Statistics  
Penn State University  
University Park, PA, USA

ISSN 1431-875X

ISSN 2197-4136 (electronic)

Springer Texts in Statistics

ISBN 978-1-4939-9759-6

ISBN 978-1-4939-9761-9 (eBook)

<https://doi.org/10.1007/978-1-4939-9761-9>

© Springer Science+Business Media, LLC, part of Springer Nature 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Science+Business Media, LLC part of Springer Nature.

The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

To our parents:  
Jianmin Li and Liji Du  
Nagarathnam and Mallayya

and to our families:  
Yanling, Ann, and Terrence  
Sudha, Vinay, and Vijay

---

## Preface

It is our goal to write a compact, rigorous, self-contained, and accessible graduate textbook on statistical estimation and inference that reflects the current trends in statistical research.

The book contains three main themes: the finite-sample theory, the asymptotic theory, and Bayesian statistics. Chapters 2 through 4 are devoted to the finite-sample theory, which includes the classical theory of optimal estimation and hypothesis test, sufficiency, completeness, ancillarity, and exponential families. Chapters 5 to 6 are devoted to Bayesian statistics, covering prior and posterior distributions, Bayesian decision theory for estimation, hypothesis testing, and classification, empirical Bayes, shrinkage estimates. Chapters 8 through 11 are devoted to asymptotic theory, covering consistency and asymptotic normality of maximum likelihood estimation and estimating equations, the Le Cam-Hajek convolution theorem for regular estimates, and the asymptotic analysis of a wide variety of hypothesis testing procedures. Two chapters on preliminaries are included to make the book self-contained: Chapter 1 contains preliminaries for the finite-sample theory and Bayesian statistics; Chapter 7 for the asymptotic theory.

The topics and treatment of some material are different from a typical textbook on statistical inference, which we regard as a special feature of this book. For example, we devoted a chapter on estimating equations and used it as a unifying mechanism to cover some useful methodologies such as the generalized linear models, generalized estimation equations, quasi likelihood estimation, and conditional inference. We include a systematic exposition of the theory of regular estimates, from regularity, contiguity, the convolution theory, to asymptotic efficiency. This theory was then used in conjunction with the Local Asymptotic Normal (LAN) assumption to develop asymptotic local alternative distributions and the optimal properties for a wide variety of hypothesis testing procedures that can be written as quadratic forms in the limit.

One of the features of the book is the systematic use of a parsimonious set of assumptions and mathematical tools to streamline some recurring regularity conditions, and theoretical results that are fundamentally similar. This makes the development of the methodology more transparent and interconnected, and the book a coherent whole. For example, the conditions “differentiable under the integral sign (DUI)”, and “stochastic equicontinuity” are repeatedly used throughout many chapters of the book; the geometric projection and the multivariate Cauchy-Schwarz inequality are used to unify different types of optimal theories; the structures of asymptotic estimation and hypothesis testing echo their counterparts in the finite-sample theory.

This book can be used either as a one-semester or a two-semester textbook on statistical inference. For the two-semester courses, the first six chapters can be used for the first semester to cover finite-sample estimation and Bayesian statistics, and the last five for the second semester to cover asymptotic statistics. For a one-semester course, there are several pathways depending on the instructor’s emphasis. For example, one possibility is to use Chapters 1, 3, 4, 7, 10, 11 for an advanced course on hypothesis testing; another possibility is to use Chapters 1, 2, 5, part of 6, 7, 8, 9 as an advanced course on point estimation and Bayesian statistics.

The book grew out of the lecture notes for two graduate-level courses that we have taught for more than two decades at the Pennsylvania State University. Over this period we have revamped the courses several times to adapt to the evolving trends, emphases, and demands in theoretical and methodological research. The authors are grateful to the Department of Statistics of the Pennsylvania State University for its constant support and the stimulating research and education environment it provides. The authors also gratefully acknowledge the support from the National Science Foundation grants.

State College  
April 2019

*Bing Li*  
*G. Jogesh Babu*

---

# Contents

<b>1</b>	<b>Probability and Random Variables</b> . . . . .	1
1.1	Sample space, events, and probability . . . . .	1
1.2	$\sigma$ -field and measure . . . . .	1
1.3	Measurable function and random variable . . . . .	3
1.4	Integral and its properties . . . . .	4
1.5	Some inequalities . . . . .	7
1.6	Logical statements modulo a measure . . . . .	7
1.7	Integration to the limit . . . . .	9
1.8	Differentiation under integral . . . . .	11
1.9	Change of variables . . . . .	12
1.10	The Radon-Nikodym Theorem . . . . .	13
1.11	Fubini's theorem . . . . .	15
1.12	Conditional probability . . . . .	16
1.13	Conditional expectation . . . . .	18
1.14	Conditioning on a random element . . . . .	21
1.15	Conditional distributions and densities . . . . .	22
1.16	Dynkin's $\pi$ - $\lambda$ theorem . . . . .	23
1.17	Derivatives and other notations . . . . .	24
	Problems . . . . .	25
	References . . . . .	29
<b>2</b>	<b>Classical Theory of Estimation</b> . . . . .	31
2.1	Families of probability measures . . . . .	31
2.1.1	Dominated and homogeneous families . . . . .	31
2.1.2	Parametric families . . . . .	34
2.1.3	Exponential families . . . . .	35
2.2	Sufficient, complete, and ancillary statistics . . . . .	37
2.3	Complete sufficient statistics for exponential family . . . . .	42
2.4	Unbiased estimator and Cramér-Rao lower bound . . . . .	43
2.5	Conditioning on complete and sufficient statistics . . . . .	49
2.6	Fisher consistency and two classical estimators . . . . .	53
	Problems . . . . .	55
	References . . . . .	59



<b>3</b>	<b>Testing Hypotheses for a Single Parameter</b> . . . . .	61
3.1	Basic concepts . . . . .	61
3.2	The Neyman-Pearson Lemma . . . . .	64
3.3	Uniformly Most Powerful test for one-sided hypotheses . . . . .	67
3.3.1	Definition and examples of UMP tests . . . . .	67
3.3.2	Monotone Likelihood Ratio . . . . .	70
3.3.3	The general form of UMP tests . . . . .	72
3.3.4	Properties of the one-sided UMP test . . . . .	74
3.4	Uniformly Most Powerful Unbiased test and two-sided hypotheses . . . . .	75
3.4.1	Uniformly Most Powerful Unbiased tests . . . . .	77
3.4.2	More properties of the exponential family . . . . .	78
3.4.3	Generalized Neyman-Pearson Lemma . . . . .	79
3.4.4	Quantile transformation and construction of two-sided tests . . . . .	80
3.4.5	UMP test for hypothesis III . . . . .	86
3.4.6	UMPU tests for hypotheses I and II . . . . .	88
	Problems . . . . .	93
	References . . . . .	97
<b>4</b>	<b>Testing Hypotheses in the Presence of Nuisance Parameters</b> . . . . .	99
4.1	Unbiased and Similar tests . . . . .	100
4.2	Sufficiency and completeness for a part of the parameter vector . . . . .	103
4.3	UMPU tests in the presence of nuisance parameters . . . . .	107
4.4	Invariant family and ancillarity . . . . .	113
4.5	Using Basu's theorem to construct UMPU test . . . . .	118
4.6	UMPU test for a linear function of $\theta$ . . . . .	121
4.7	UMPU test for nonregular family . . . . .	125
4.8	Confidence sets . . . . .	127
	Problems . . . . .	131
	References . . . . .	134
<b>5</b>	<b>Basic Ideas of Bayesian Methods</b> . . . . .	135
5.1	Prior, posterior, and likelihood . . . . .	135
5.2	Conditional independence and Bayesian sufficiency . . . . .	137
5.3	Conjugate families . . . . .	144
5.4	Two-parameter normal family . . . . .	147
5.5	Multivariate Normal likelihood . . . . .	151
5.6	Improper prior . . . . .	154
5.6.1	The motivation idea of improper prior . . . . .	154
5.6.2	Haar measures . . . . .	156
5.6.3	Jeffreys prior . . . . .	160
5.7	Statistical decision theory . . . . .	162
	Problems . . . . .	165
	References . . . . .	172

<b>6</b>	<b>Bayesian Inference</b> .....	173
	6.1 Estimation .....	173
	6.2 Bayes rule and unbiasedness .....	178
	6.3 Error assessment of estimators .....	179
	6.4 Credible sets .....	180
	6.5 Hypothesis test .....	182
	6.6 Classification .....	187
	6.7 Stein's phenomenon .....	190
	6.8 Empirical Bayes .....	193
	Problems .....	195
	References .....	200
<b>7</b>	<b>Asymptotic tools and projections</b> .....	203
	7.1 Laws of Large Numbers .....	203
	7.2 Convergence in distribution .....	206
	7.3 Argument via subsequences .....	212
	7.4 Argument via simple functions .....	214
	7.5 The Central Limit Theorems .....	215
	7.6 The $\delta$ -method .....	218
	7.7 Mann-Wald notation for order of magnitude .....	220
	7.8 Hilbert spaces .....	223
	7.9 Multivariate Cauchy-Schwarz inequality .....	228
	7.10 Projections .....	230
	Problems .....	233
	References .....	236
<b>8</b>	<b>Asymptotic theory for Maximum Likelihood Estimation</b> ..	237
	8.1 Maximum Likelihood Estimation .....	237
	8.2 Cramér's approach to consistency .....	240
	8.3 Almost everywhere uniform convergence .....	244
	8.4 Wald's approach to consistency .....	247
	8.5 Asymptotic normality .....	251
	Problems .....	255
	References .....	259
<b>9</b>	<b>Estimating equations</b> .....	261
	9.1 Optimal Estimating Equations .....	261
	9.2 Quasi likelihood estimation .....	266
	9.3 Generalized Estimating Equations .....	268
	9.4 Other optimal estimating equations .....	271
	9.5 Asymptotic properties .....	273
	9.6 One-step Newton-Raphson estimate .....	275
	9.7 Asymptotic linear form .....	278
	9.8 Efficient score for parameter of interest .....	279
	Problems .....	287
	References .....	292

<b>10 Convolution Theorem and Asymptotic Efficiency</b> . . . . .	295
10.1 Contiguity . . . . .	296
10.2 Le Cam's first lemma . . . . .	296
10.3 Le Cam's third lemma . . . . .	300
10.4 Local asymptotic Normality . . . . .	302
10.5 The convolution theorem . . . . .	305
10.6 Asymptotically efficient estimates . . . . .	310
10.7 Augmented LAN . . . . .	312
10.8 Le Cam's third lemma under ALAN . . . . .	315
10.9 Superefficiency . . . . .	317
Problems . . . . .	320
References . . . . .	326
<b>11 Asymptotic Hypothesis Test</b> . . . . .	329
11.1 Quadratic Form test . . . . .	329
11.2 Wilks's likelihood ratio test . . . . .	331
11.3 Wald's, Rao's, and Neyman's tests . . . . .	335
11.3.1 Wald's test . . . . .	335
11.3.2 Rao's test . . . . .	336
11.3.3 Neyman's $C(\alpha)$ test . . . . .	336
11.4 Asymptotically efficient test . . . . .	338
11.5 Pitman efficiency . . . . .	342
11.6 Hypothesis specified by an arbitrary constraint . . . . .	343
11.6.1 Asymptotic analysis of constrained MLE . . . . .	344
11.6.2 Likelihood ratio test for general hypotheses . . . . .	349
11.6.3 Wald's test and Rao's test for general hypotheses . . . . .	350
11.6.4 Neyman's $C(\alpha)$ test for general hypotheses . . . . .	352
11.6.5 Asymptotic efficiency . . . . .	353
11.7 QF tests for estimating equations . . . . .	356
11.7.1 Wald's, Rao's, and Neyman's tests for estimating equations . . . . .	356
11.7.2 QF tests for canonical estimating equations . . . . .	362
11.7.3 Wilks's test for conservative estimating equations . . . . .	364
Problems . . . . .	365
References . . . . .	373
<b>Index</b> . . . . .	375