

Wiley Series in Probability and Statistics

[Redacted]

Examples and Problems in Mathematical Statistics

Shelemyahu Zacks

WILEY

Examples and Problems in Mathematical Statistics

SHELEMYAHU ZACKS

Department of Mathematical Sciences
Binghamton University
Binghamton, NY



WILEY

Copyright © 2014 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Zacks, Shelemyahu, 1932- author.

Examples and problems in mathematical statistics / Shelemyahu Zacks.

pages cm

Summary: "This book presents examples that illustrate the theory of mathematical statistics and details how to apply the methods for solving problems" – Provided by publisher.

Includes bibliographical references and index.

ISBN 978-1-118-60550-9 (hardback)

1. Mathematical statistics—Problems, exercises, etc. I. Title.

QC32.Z265 2013

519.5—dc23

2013034492

Printed in the United States of America

ISBN: 9781118605509

10 9 8 7 6 5 4 3 2 1

Examples and Problems
in Mathematical Statistics

WILEY SERIES IN PROBABILITY AND STATISTICS

Established by WALTER A. SHEWHART and SAMUEL S. WILKS

Editors: *David J. Balding, Noel A. C. Cressie, Garrett M. Fitzmaurice,
Harvey Goldstein, Iain M. Johnstone, Geert Molenberghs, David W. Scott,
Adrian F. M. Smith, Ruey S. Tsay, Sanford Weisberg*

Editors Emeriti: *Vic Barnett, J. Stuart Hunter, Joseph B. Kadane, Jozef L. Teugels*

A complete list of the titles in this series appears at the end of this volume.

To my wife Hanna,
our sons Yuval and David,
and their families, with love.

Preface

I have been teaching probability and mathematical statistics to graduate students for close to 50 years. In my career I realized that the most difficult task for students is solving problems. Bright students can generally grasp the theory easier than apply it. In order to overcome this hurdle, I used to write examples of solutions to problems and hand it to my students. I often wrote examples for the students based on my published research. Over the years I have accumulated a large number of such examples and problems. This book is aimed at sharing these examples and problems with the population of students, researchers, and teachers.

The book consists of nine chapters. Each chapter has four parts. The first part contains a short presentation of the theory. This is required especially for establishing notation and to provide a quick overview of the important results and references. The second part consists of examples. The examples follow the theoretical presentation. The third part consists of problems for solution, arranged by the corresponding sections of the theory part. The fourth part presents solutions to some selected problems. The solutions are generally not as detailed as the examples, but as such these are examples of solutions. I tried to demonstrate how to apply known results in order to solve problems elegantly. All together there are in the book 167 examples and 431 problems.

The emphasis in the book is on statistical inference. The first chapter on probability is especially important for students who have not had a course on advanced probability. Chapter Two is on the theory of distribution functions. This is basic to all developments in the book, and from my experience, it is important for all students to master this calculus of distributions. The chapter covers multivariate distributions, especially the multivariate normal; conditional distributions; techniques of determining variances and covariances of sample moments; the theory of exponential families; Edgeworth expansions and saddle-point approximations; and more. Chapter Three covers the theory of sufficient statistics, completeness of families of distributions, and the information in samples. In particular, it presents the Fisher information, the Kullback–Leibler information, and the Hellinger distance. Chapter Four provides a strong foundation in the theory of testing statistical hypotheses. The Wald SPRT is

discussed there too. Chapter Five is focused on optimal point estimation of different kinds. Pitman estimators and equivariant estimators are also discussed. Chapter Six covers problems of efficient confidence intervals, in particular the problem of determining fixed-width confidence intervals by two-stage or sequential sampling. Chapter Seven covers techniques of large sample approximations, useful in estimation and testing. Chapter Eight is devoted to Bayesian analysis, including empirical Bayes theory. It highlights computational approximations by numerical analysis and simulations. Finally, Chapter Nine presents a few more advanced topics, such as minimaxity, admissibility, structural distributions, and the Stein-type estimators.

I would like to acknowledge with gratitude the contributions of my many ex-students, who toiled through these examples and problems and gave me their important feedback. In particular, I am very grateful and indebted to my colleagues, Professors A. Schick, Q. Yu, S. De, and A. Polunchenko, who carefully read parts of this book and provided important comments. Mrs. Marge Pratt skillfully typed several drafts of this book with patience and grace. To her I extend my heartfelt thanks. Finally, I would like to thank my wife Hanna for giving me the conditions and encouragement to do research and engage in scholarly writing.

SHELEMYAHU ZACKS

List of Random Variables

$B(n, p)$	Binomial, with parameters n and p
$E(\mu)$	Exponential with parameter μ
$EV(\lambda, \alpha)$	Extreme value with parameters λ and α
$F(v_1, v_2)$	Central F with parameters v_1 and v_2
$F(n_1, n_2; \lambda)$	Noncentral F with parameters v_1, v_2, λ
$G(\lambda, p)$	Gamma with parameters λ and p
$H(M, N, n)$	Hyper-geometric with parameters M, N, n
$N(\mu, V)$	Multinormal with mean vector μ and covariance matrix V
$N(\mu, \sigma)$	Normal with mean μ and σ
$NB(\psi, v)$	Negative-binomial with parameters ψ , and v
$P(\lambda)$	Poisson with parameter λ
$R(a, b)$	Rectangular (uniform) with parameters a and b
$t[n; \lambda]$	Noncentral Student's t with parameters n and λ
$t[n; \xi, V]$	Multivariate t with parameters n, ξ and V
$t[n]$	Student's t with n degrees of freedom
$W(\lambda, \alpha)$	Weibul with parameters λ and α
$\beta(p, q)$	Beta with parameters p and q
$\chi^2[n, \lambda]$	Noncentral chi-squared with parameters n and λ
$\chi^2[n]$	Chi-squared with n degrees of freedom

List of Abbreviations

a.s.	Almost surely
ANOVA	Analysis of variance
c.d.f.	Cumulative distribution function
$\text{cov}(x, y)$	Covariance of X and Y
CI	Confidence interval
CLT	Central limit theorem
CP	Coverage probability
CR	Cramer Rao regularity conditions
$E\{X Y\}$	Conditional expected value of X , given Y
$E\{X\}$	Expected value of X
FIM	Fisher information matrix
i.i.d.	Independent identically distributed
LBUE	Linear best unbiased estimate
LCL	Lower confidence limit
m.g.f.	Moment generating function
m.s.s.	Minimal sufficient statistics
MEE	Moments equations estimator
MLE	Maximum likelihood estimator
MLR	Monotone likelihood ratio
MP	Most powerful
MSE	Mean squared error
MVU	Minimum variance unbiased
OC	Operating characteristic
p.d.f.	Probability density function
p.g.f.	Probability generating function
$P\{E A\}$	Conditional probability of E , given A
$P\{E\}$	Probability of E
PTE	Pre-test estimator
r.v.	Random variable
RHS	Right-hand side
s.v.	Stopping variable

SE	Standard error
SLLN	Strong law of large numbers
SPRT	Sequential probability ratio test
$\text{tr}\{A\}$	trace of the matrix A
UCL	Upper control limit
UMP	Uniformly most powerful
UMPI	Uniformly most powerful invariant
UMPU	Uniformly most powerful unbiased
UMVU	Uniformly minimum variance unbiased
$V\{X Y\}$	Conditional variance of X , given Y
$V\{X\}$	Variance of X
w.r.t.	With respect to
WLLN	Weak law of large numbers

Contents

Preface	xv
List of Random Variables	xvii
List of Abbreviations	xix
1 Basic Probability Theory	1
PART I: THEORY, 1	
1.1 Operations on Sets, 1	
1.2 Algebra and σ -Fields, 2	
1.3 Probability Spaces, 4	
1.4 Conditional Probabilities and Independence, 6	
1.5 Random Variables and Their Distributions, 8	
1.6 The Lebesgue and Stieltjes Integrals, 12	
1.6.1 General Definition of Expected Value: The Lebesgue Integral, 12	
1.6.2 The Stieltjes–Riemann Integral, 17	
1.6.3 Mixtures of Discrete and Absolutely Continuous Distributions, 19	
1.6.4 Quantiles of Distributions, 19	
1.6.5 Transformations, 20	
1.7 Joint Distributions, Conditional Distributions and Independence, 21	
1.7.1 Joint Distributions, 21	
1.7.2 Conditional Expectations: General Definition, 23	
1.7.3 Independence, 26	
1.8 Moments and Related Functionals, 26	
1.9 Modes of Convergence, 35	
1.10 Weak Convergence, 39	
1.11 Laws of Large Numbers, 41	

1.11.1	The Weak Law of Large Numbers (WLLN),	41
1.11.2	The Strong Law of Large Numbers (SLLN),	42
1.12	Central Limit Theorem,	44
1.13	Miscellaneous Results,	47
1.13.1	Law of the Iterated Logarithm,	48
1.13.2	Uniform Integrability,	48
1.13.3	Inequalities,	52
1.13.4	The Delta Method,	53
1.13.5	The Symbols o_p and O_p ,	55
1.13.6	The Empirical Distribution and Sample Quantiles,	55
PART II: EXAMPLES,		56
PART III: PROBLEMS,		73
PART IV: SOLUTIONS TO SELECTED PROBLEMS,		93

2 Statistical Distributions

106

PART I: THEORY,		106
2.1	Introductory Remarks,	106
2.2	Families of Discrete Distributions,	106
2.2.1	Binomial Distributions,	106
2.2.2	Hypergeometric Distributions,	107
2.2.3	Poisson Distributions,	108
2.2.4	Geometric, Pascal, and Negative Binomial Distributions,	108
2.3	Some Families of Continuous Distributions,	109
2.3.1	Rectangular Distributions,	109
2.3.2	Beta Distributions,	111
2.3.3	Gamma Distributions,	111
2.3.4	Weibull and Extreme Value Distributions,	112
2.3.5	Normal Distributions,	113
2.3.6	Normal Approximations,	114
2.4	Transformations,	118
2.4.1	One-to-One Transformations of Several Variables,	118
2.4.2	Distribution of Sums,	118
2.4.3	Distribution of Ratios,	118
2.5	Variances and Covariances of Sample Moments,	120
2.6	Discrete Multivariate Distributions,	122
2.6.1	The Multinomial Distribution,	122
2.6.2	Multivariate Negative Binomial,	123
2.6.3	Multivariate Hypergeometric Distributions,	124

2.7	Multinormal Distributions, 125
2.7.1	Basic Theory, 125
2.7.2	Distribution of Subvectors and Distributions of Linear Forms, 127
2.7.3	Independence of Linear Forms, 129
2.8	Distributions of Symmetric Quadratic Forms of Normal Variables, 130
2.9	Independence of Linear and Quadratic Forms of Normal Variables, 132
2.10	The Order Statistics, 133
2.11	t -Distributions, 135
2.12	F -Distributions, 138
2.13	The Distribution of the Sample Correlation, 142
2.14	Exponential Type Families, 144
2.15	Approximating the Distribution of the Sample Mean: Edgeworth and Saddlepoint Approximations, 146
2.15.1	Edgeworth Expansion, 147
2.15.2	Saddlepoint Approximation, 149
	PART II: EXAMPLES, 150
	PART III: PROBLEMS, 167
	PART IV: SOLUTIONS TO SELECTED PROBLEMS, 181

3 Sufficient Statistics and the Information in Samples

191

	PART I: THEORY, 191
3.1	Introduction, 191
3.2	Definition and Characterization of Sufficient Statistics, 192
3.2.1	Introductory Discussion, 192
3.2.2	Theoretical Formulation, 194
3.3	Likelihood Functions and Minimal Sufficient Statistics, 200
3.4	Sufficient Statistics and Exponential Type Families, 202
3.5	Sufficiency and Completeness, 203
3.6	Sufficiency and Ancillarity, 205
3.7	Information Functions and Sufficiency, 206
3.7.1	The Fisher Information, 206
3.7.2	The Kullback–Leibler Information, 210
3.8	The Fisher Information Matrix, 212
3.9	Sensitivity to Changes in Parameters, 214
3.9.1	The Hellinger Distance, 214
	PART II: EXAMPLES, 216
	PART III: PROBLEMS, 230
	PART IV: SOLUTIONS TO SELECTED PROBLEMS, 236

4	Testing Statistical Hypotheses	246
	PART I: THEORY, 246	
4.1	The General Framework, 246	
4.2	The Neyman–Pearson Fundamental Lemma, 248	
4.3	Testing One-Sided Composite Hypotheses in MLR Models, 251	
4.4	Testing Two-Sided Hypotheses in One-Parameter Exponential Families, 254	
4.5	Testing Composite Hypotheses with Nuisance Parameters—Unbiased Tests, 256	
4.6	Likelihood Ratio Tests, 260	
4.6.1	Testing in Normal Regression Theory, 261	
4.6.2	Comparison of Normal Means: The Analysis of Variance, 265	
4.7	The Analysis of Contingency Tables, 271	
4.7.1	The Structure of Multi-Way Contingency Tables and the Statistical Model, 271	
4.7.2	Testing the Significance of Association, 271	
4.7.3	The Analysis of 2×2 Tables, 273	
4.7.4	Likelihood Ratio Tests for Categorical Data, 274	
4.8	Sequential Testing of Hypotheses, 275	
4.8.1	The Wald Sequential Probability Ratio Test, 276	
	PART II: EXAMPLES, 283	
	PART III: PROBLEMS, 298	
	PART IV: SOLUTIONS TO SELECTED PROBLEMS, 307	
5	Statistical Estimation	321
	PART I: THEORY, 321	
5.1	General Discussion, 321	
5.2	Unbiased Estimators, 322	
5.2.1	General Definition and Example, 322	
5.2.2	Minimum Variance Unbiased Estimators, 322	
5.2.3	The Cramér–Rao Lower Bound for the One-Parameter Case, 323	
5.2.4	Extension of the Cramér–Rao Inequality to Multiparameter Cases, 326	
5.2.5	General Inequalities of the Cramér–Rao Type, 327	
5.3	The Efficiency of Unbiased Estimators in Regular Cases, 328	
5.4	Best Linear Unbiased and Least-Squares Estimators, 331	
5.4.1	BLUEs of the Mean, 331	
5.4.2	Least-Squares and BLUEs in Linear Models, 332	
5.4.3	Best Linear Combinations of Order Statistics, 334	