

# **Mathematical Statistics with Applications**



**ELSEVIER** *science &  
technology books*

• • Companion Web Site:

<http://www.elsevierdirect.com/companions/9780123748485>

*Mathematical Statistics with Applications*  
by Kandethody M. Ramachandran and Chris P. Tsokos

Resources for Professors:

- All figures from the book available as PowerPoint slides and as jpegs.
- Links to Web sites carefully chosen to supplement the content of the textbook.
- Online Student Solutions Manual is now available through separate purchase.
- Also available with purchase of *Mathematical Statistics with Applications*, password protected and activated upon registration, online Instructors' Solutions Manual.



TOOLS FOR **ALL** YOUR TEACHING NEEDS  
[textbooks.elsevier.com](http://textbooks.elsevier.com)



**ACADEMIC  
PRESS**

To adopt this book for course use, visit <http://textbooks.elsevier.com>

# Mathematical Statistics with Applications

**Kandethody M. Ramachandran**

*Department of Mathematics and Statistics  
University of South Florida  
Tampa, FL*

**Chris P. Tsokos**

*Department of Mathematics and Statistics  
University of South Florida  
Tampa, FL*



ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON  
NEW YORK • OXFORD • PARIS • SAN DIEGO  
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier



Elsevier Academic Press  
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA  
525 B Street, Suite 1900, San Diego, California 92101-4495, USA  
84 Theobald's Road, London WC1X 8RR, UK

This book is printed on acid-free paper. (∞)

Copyright © 2009, Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone: (+44) 1865 843830, fax: (+44) 1865 853333, E-mail: [permissions@elsevier.co.uk](mailto:permissions@elsevier.co.uk). You may also complete your request on-line via the Elsevier homepage (<http://elsevier.com>), by selecting "Customer Support" and then "Obtaining Permissions."

#### Library of Congress Cataloging-in-Publication Data

Ramachandran, K. M.

Mathematical statistics with applications / Kandethody M. Ramachandran, Chris P. Tsokos.  
p. cm.

ISBN 978-0-12-374848-5 (hardcover : alk. paper)

1. Mathematical statistics. 2. Mathematical statistics—Data processing. I. Tsokos, Chris P. II. Title.

QA276.R328 2009

519.5—dc22

2008044556

#### British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library.

ISBN 13: 978-0-12-374848-5

For all information on all Elsevier Academic Press publications  
visit our Web site at [www.elsevierdirect.com](http://www.elsevierdirect.com)

Printed in the United States of America  
09 10 9 8 7 6 5 4 3 2 1

Working together to grow  
libraries in developing countries

[www.elsevier.com](http://www.elsevier.com) | [www.bookaid.org](http://www.bookaid.org) | [www.sabre.org](http://www.sabre.org)

ELSEVIER

BOOK AID  
International

Sabre Foundation

*Dedicated to our families:*  
*Usha, Vikas, Vilas, and Varsha Ramachandran*  
*and*  
*Debbie, Matthew, Jonathan, and Maria Tsokos*



# Contents

Preface .....	xv
Acknowledgments .....	xix
About the Authors .....	xxi
Flow Chart .....	xxiii
<b>CHAPTER 1 Descriptive Statistics .....</b>	<b>1</b>
1.1 Introduction .....	2
1.1.1 Data Collection .....	3
1.2 Basic Concepts .....	3
1.2.1 Types of Data .....	5
1.3 Sampling Schemes .....	8
1.3.1 Errors in Sample Data .....	11
1.3.2 Sample Size .....	12
1.4 Graphical Representation of Data .....	13
1.5 Numerical Description of Data .....	26
1.5.1 Numerical Measures for Grouped Data .....	30
1.5.2 Box Plots .....	33
1.6 Computers and Statistics .....	39
1.7 Chapter Summary .....	40
1.8 Computer Examples .....	41
1.8.1 Minitab Examples .....	41
1.8.2 SPSS Examples .....	46
1.8.3 SAS Examples .....	47
Projects for Chapter 1 .....	51
<b>CHAPTER 2 Basic Concepts from Probability Theory .....</b>	<b>53</b>
2.1 Introduction .....	54
2.2 Random Events and Probability .....	55
2.3 Counting Techniques and Calculation of Probabilities .....	63
2.4 The Conditional Probability, Independence, and Bayes' Rule .....	71
2.5 Random Variables and Probability Distributions .....	83
2.6 Moments and Moment-Generating Functions .....	92
2.6.1 Skewness and Kurtosis .....	98
2.7 Chapter Summary .....	107
2.8 Computer Examples (Optional) .....	108
2.8.1 Minitab Computations .....	109
2.8.2 SPSS Examples .....	110
2.8.3 SAS Examples .....	110
Projects for Chapter 2 .....	112

<b>CHAPTER 3</b>	<b>Additional Topics in Probability</b>	<b>113</b>
3.1	Introduction	114
3.2	Special Distribution Functions	114
3.2.1	The Binomial Probability Distribution	114
3.2.2	Poisson Probability Distribution	119
3.2.3	Uniform Probability Distribution	122
3.2.4	Normal Probability Distribution	125
3.2.5	Gamma Probability Distribution	131
3.3	Joint Probability Distributions	141
3.3.1	Covariance and Correlation	148
3.4	Functions of Random Variables	154
3.4.1	Method of Distribution Functions	154
3.4.2	The pdf of $Y = g(X)$ , Where $g$ Is Differentiable and Monotone Increasing or Decreasing	156
3.4.3	Probability Integral Transformation	157
3.4.4	Functions of Several Random Variables: Method of Distribution Functions	158
3.4.5	Transformation Method	159
3.5	Limit Theorems	163
3.6	Chapter Summary	173
3.7	Computer Examples (Optional)	175
3.7.1	Minitab Examples	175
3.7.2	SPSS Examples	177
3.7.3	SAS Examples	178
	Projects for Chapter 3	180
 <b>CHAPTER 4</b>	 <b>Sampling Distributions</b>	 <b>183</b>
4.1	Introduction	184
4.1.1	Finite Population	187
4.2	Sampling Distributions Associated with Normal Populations	191
4.2.1	Chi-Square Distribution	192
4.2.2	Student $t$ -Distribution	198
4.2.3	$F$ -Distribution	202
4.3	Order Statistics	207
4.4	Large Sample Approximations	212
4.4.1	The Normal Approximation to the Binomial Distribution	213
4.5	Chapter Summary	218
4.6	Computer Examples	219
4.6.1	Minitab Examples	219
4.6.2	SPSS Examples	219
4.6.3	SAS Examples	219
	Projects for Chapter 4	221



<b>CHAPTER 5</b>	<b>Point Estimation</b>	225
5.1	Introduction	226
5.2	The Method of Moments	227
5.3	The Method of Maximum Likelihood	235
5.4	Some Desirable Properties of Point Estimators	246
5.4.1	Unbiased Estimators	247
5.4.2	Sufficiency	252
5.5	Other Desirable Properties of a Point Estimator	266
5.5.1	Consistency	266
5.5.2	Efficiency	270
5.5.3	Minimal Sufficiency and Minimum-Variance Unbiased Estimation	277
5.6	Chapter Summary	282
5.7	Computer Examples	283
	Projects for Chapter 5	285
<b>CHAPTER 6</b>	<b>Interval Estimation</b>	291
6.1	Introduction	292
6.1.1	A Method of Finding the Confidence Interval: Pivotal Method	293
6.2	Large Sample Confidence Intervals: One Sample Case	300
6.2.1	Confidence Interval for Proportion, $p$	302
6.2.2	Margin of Error and Sample Size	303
6.3	Small Sample Confidence Intervals for $\mu$	310
6.4	A Confidence Interval for the Population Variance	315
6.5	Confidence Interval Concerning Two Population Parameters	321
6.6	Chapter Summary	330
6.7	Computer Examples	330
6.7.1	Minitab Examples	330
6.7.2	SPSS Examples	332
6.7.3	SAS Examples	333
	Projects for Chapter 6	334
<b>CHAPTER 7</b>	<b>Hypothesis Testing</b>	337
7.1	Introduction	338
7.1.1	Sample Size	346
7.2	The Neyman–Pearson Lemma	349
7.3	Likelihood Ratio Tests	355
7.4	Hypotheses for a Single Parameter	361
7.4.1	The $p$ -Value	361
7.4.2	Hypothesis Testing for a Single Parameter	363

7.5	Testing of Hypotheses for Two Samples .....	372
7.5.1	Independent Samples.....	373
7.5.2	Dependent Samples .....	382
7.6	Chi-Square Tests for Count Data .....	388
7.6.1	Testing the Parameters of Multinomial Distribution: Goodness-of-Fit Test .....	390
7.6.2	Contingency Table: Test for Independence .....	392
7.6.3	Testing to Identify the Probability Distribution: Goodness-of-Fit Chi-Square Test .....	395
7.7	Chapter Summary .....	399
7.8	Computer Examples .....	399
7.8.1	Minitab Examples .....	400
7.8.2	SPSS Examples .....	403
7.8.3	SAS Examples.....	405
	Projects for Chapter 7 .....	408

<b>CHAPTER 8</b>	<b>Linear Regression Models .....</b>	<b>411</b>
8.1	Introduction .....	412
8.2	The Simple Linear Regression Model .....	413
8.2.1	The Method of Least Squares.....	415
8.2.2	Derivation of $\hat{\beta}_0$ and $\hat{\beta}_1$ .....	416
8.2.3	Quality of the Regression .....	421
8.2.4	Properties of the Least-Squares Estimators for the Model $Y = \beta_0 + \beta_1 x + \varepsilon$ .....	422
8.2.5	Estimation of Error Variance $\sigma^2$ .....	425
8.3	Inferences on the Least Squares Estimators .....	428
8.3.1	Analysis of Variance (ANOVA) Approach to Regression .....	434
8.4	Predicting a Particular Value of $Y$ .....	437
8.5	Correlation Analysis .....	440
8.6	Matrix Notation for Linear Regression .....	445
8.6.1	ANOVA for Multiple Regression .....	449
8.7	Regression Diagnostics .....	451
8.8	Chapter Summary .....	454
8.9	Computer Examples .....	455
8.9.1	Minitab Examples .....	455
8.9.2	SPSS Examples .....	457
8.9.3	SAS Examples.....	458
	Projects for Chapter 8 .....	461

<b>CHAPTER 9</b>	<b>Design of Experiments .....</b>	<b>465</b>
9.1	Introduction .....	466
9.2	Concepts from Experimental Design .....	467
9.2.1	Basic Terminology .....	467

9.2.2	Fundamental Principles: Replication, Randomization, and Blocking .....	471
9.2.3	Some Specific Designs .....	474
9.3	Factorial Design .....	483
9.3.1	One-Factor-at-a-Time Design .....	483
9.3.2	Full Factorial Design .....	485
9.3.3	Fractional Factorial Design .....	486
9.4	Optimal Design .....	487
9.4.1	Choice of Optimal Sample Size .....	487
9.5	The Taguchi Methods .....	489
9.6	Chapter Summary .....	493
9.7	Computer Examples .....	494
9.7.1	Minitab Examples .....	494
9.7.2	SAS Examples .....	494
	Projects for Chapter 9 .....	497
<b>CHAPTER 10</b>	<b>Analysis of Variance .....</b>	<b>499</b>
10.1	Introduction .....	500
10.2	Analysis of Variance Method for Two Treatments (Optional) .....	501
10.3	Analysis of Variance for Completely Randomized Design .....	510
10.3.1	The $p$ -Value Approach .....	515
10.3.2	Testing the Assumptions for One-Way ANOVA .....	517
10.3.3	Model for One-Way ANOVA (Optional) .....	522
10.4	Two-Way Analysis of Variance, Randomized Complete Block Design .....	526
10.5	Multiple Comparisons .....	536
10.6	Chapter Summary .....	543
10.7	Computer Examples .....	543
10.7.1	Minitab Examples .....	543
10.7.2	SPSS Examples .....	546
10.7.3	SAS Examples .....	548
	Projects for Chapter 10 .....	554
<b>CHAPTER 11</b>	<b>Bayesian Estimation and Inference .....</b>	<b>559</b>
11.1	Introduction .....	560
11.2	Bayesian Point Estimation .....	562
11.2.1	Criteria for Finding the Bayesian Estimate .....	569
11.3	Bayesian Confidence Interval or Credible Intervals .....	579
11.4	Bayesian Hypothesis Testing .....	584
11.5	Bayesian Decision Theory .....	588
11.6	Chapter Summary .....	596
11.7	Computer Examples .....	596
	Projects for Chapter 11 .....	596

<b>CHAPTER 12 Nonparametric Tests</b>	599
12.1 Introduction	600
12.2 Nonparametric Confidence Interval	601
12.3 Nonparametric Hypothesis Tests for One Sample	606
12.3.1 The Sign Test	607
12.3.2 Wilcoxon Signed Rank Test	611
12.3.3 Dependent Samples: Paired Comparison Tests	617
12.4 Nonparametric Hypothesis Tests for Two Independent Samples	620
12.4.1 Median Test	620
12.4.2 The Wilcoxon Rank Sum Test	625
12.5 Nonparametric Hypothesis Tests for $k \geq 2$ Samples	630
12.5.1 The Kruskal–Wallis Test	631
12.5.2 The Friedman Test	634
12.6 Chapter Summary	640
12.7 Computer Examples	642
12.7.1 Minitab Examples	642
12.7.2 SPSS Examples	646
12.7.3 SAS Examples	648
Projects for Chapter 12	652
<b>CHAPTER 13 Empirical Methods</b>	657
13.1 Introduction	658
13.2 The Jackknife Method	658
13.3 An Introduction to Bootstrap Methods	663
13.3.1 Bootstrap Confidence Intervals	667
13.4 The Expectation Maximization Algorithm	669
13.5 Introduction to Markov Chain Monte Carlo	681
13.5.1 Metropolis Algorithm	685
13.5.2 The Metropolis–Hastings Algorithm	688
13.5.3 Gibbs Algorithm	692
13.5.4 MCMC Issues	695
13.6 Chapter Summary	697
13.7 Computer Examples	698
13.7.1 SAS Examples	699
Projects for Chapter 13	699
<b>CHAPTER 14 Some Issues in Statistical Applications: An Overview</b>	701
14.1 Introduction	702
14.2 Graphical Methods	702
14.3 Outliers	708
14.4 Checking Assumptions	713
14.4.1 Checking the Assumption of Normality	714
14.4.2 Data Transformation	716

14.4.3	Test for Equality of Variances .....	719
14.4.4	Test of Independence.....	724
14.5	Modeling Issues .....	727
14.5.1	A Simple Model for Univariate Data .....	727
14.5.2	Modeling Bivariate Data .....	730
14.6	Parametric versus Nonparametric Analysis .....	733
14.7	Tying It All Together .....	735
14.8	Conclusion .....	746
<b>Appendices</b>	.....	<b>747</b>
A.I	Set Theory .....	747
A.II	Review of Markov Chains .....	751
A.III	Common Probability Distributions .....	757
A.IV	Probability Tables .....	759
<b>References</b>	.....	<b>799</b>
<b>Index</b>	.....	<b>803</b>



# Preface

This textbook is of an interdisciplinary nature and is designed for a two- or one-semester course in probability and statistics, with basic calculus as a prerequisite. The book is primarily written to give a sound theoretical introduction to statistics while emphasizing applications. If teaching statistics is the main purpose of a two-semester course in probability and statistics, this textbook covers all the probability concepts necessary for the theoretical development of statistics in two chapters, and goes on to cover all major aspects of statistical theory in two semesters, instead of only a portion of statistical concepts. What is more, using the optional section on computer examples at the end of each chapter, the student can also simultaneously learn to utilize statistical software packages for data analysis. It is our aim, without sacrificing any rigor, to encourage students to apply the theoretical concepts they have learned. There are many examples and exercises concerning diverse application areas that will show the pertinence of statistical methodology to solving real-world problems. The examples with statistical software and projects at the end of the chapters will provide good perspective on the usefulness of statistical methods. To introduce the students to modern and increasingly popular statistical methods, we have introduced separate chapters on Bayesian analysis and empirical methods.

One of the main aims of this book is to prepare advanced undergraduates and beginning graduate students in the theory of statistics with emphasis on interdisciplinary applications. The audience for this course is regular full-time students from mathematics, statistics, engineering, physical sciences, business, social sciences, materials science, and so forth. Also, this textbook is suitable for people who work in industry and in education as a reference book on introductory statistics for a good theoretical foundation with clear indication of how to use statistical methods. Traditionally, one of the main prerequisites for this course is a semester of the introduction to probability theory. A working knowledge of elementary (descriptive) statistics is also a must. In schools where there is no statistics major, imposing such a background, in addition to calculus sequence, is very difficult. Most of the present books available on this subject contain full one-semester material for probability and then, based on those results, continue on to the topics in statistics. Also, some of these books include in their subject matter only the theory of statistics, whereas others take the cookbook approach of covering the mechanics. Thus, even with two full semesters of work, many basic and important concepts in statistics are never covered. This book has been written to remedy this problem. We fuse together both concepts in order for students to gain knowledge of the theory and at the same time develop the expertise to use their knowledge in real-world situations.

Although statistics is a very applied subject, there is no denying that it is also a very abstract subject. The purpose of this book is to present the subject matter in such a way that anyone with exposure to basic calculus can study statistics without spending two semesters of background preparation. To prepare students, we present an optional review of the elementary (descriptive) statistics in Chapter 1. All the probability material required to learn statistics is covered in two chapters. Students with a probability background can either review or skip the first three chapters. It is also our belief that any statistics course is not complete without exposure to computational techniques. At

the end of each chapter, we give some examples of how to use Minitab, SPSS, and SAS to statistically analyze data. Also, at the end of each chapter, there are projects that will enhance the knowledge and understanding of the materials covered in that chapter. In the chapter on the empirical methods, we present some of the modern computational and simulation techniques, such as bootstrap, jackknife, and Markov chain Monte Carlo methods. The last chapter summarizes some of the steps necessary to apply the material covered in the book to real-world problems. The first eight chapters have been class tested as a one-semester course for more than 3 years with five different professors teaching. The audience was junior- and senior-level undergraduate students from many disciplines who had had two semesters of calculus, most of them with no probability or statistics background. The feedback from the students and instructors was very positive. Recommendations from the instructors and students were very useful in improving the style and content of the book.

## **AIM AND OBJECTIVE OF THE TEXTBOOK**

This textbook provides a calculus-based coverage of statistics and introduces students to methods of theoretical statistics and their applications. It assumes no prior knowledge of statistics or probability theory, but does require calculus. Most books at this level are written with elaborate coverage of probability. This requires teaching one semester of probability and then continuing with one or two semesters of statistics. This creates a particular problem for non-statistics majors from various disciplines who want to obtain a sound background in mathematical statistics and applications. It is our aim to introduce basic concepts of statistics with sound theoretical explanations. Because statistics is basically an interdisciplinary applied subject, we offer many applied examples and relevant exercises from different areas. Knowledge of using computers for data analysis is desirable. We present examples of solving statistical problems using Minitab, SPSS, and SAS.

## **FEATURES**

- During years of teaching, we observed that many students who do well in mathematics courses find it difficult to understand the concept of statistics. To remedy this, we present most of the material covered in the textbook with well-defined step-by-step procedures to solve real problems. This clearly helps the students to approach problem solving in statistics more logically.
- The usefulness of each statistical method introduced is illustrated by several relevant examples.
- At the end of each section, we provide ample exercises that are a good mix of theory and applications.
- In each chapter, we give various projects for students to work on. These projects are designed in such a way that students will start thinking about how to apply the results they learned in the chapter as well as other issues they will need to know for practical situations.
- At the end of the chapters, we include an optional section on computer methods with Minitab, SPSS, and SAS examples with clear and simple commands that the student can use to analyze



data. This will help students to learn how to utilize the standard methods they have learned in the chapter to study real data.

- We introduce many of the modern statistical computational and simulation concepts, such as the jackknife and bootstrap methods, the EM algorithms, and the Markov chain Monte Carlo methods such as the Metropolis algorithm, the Metropolis–Hastings algorithm, and the Gibbs sampler. The Metropolis algorithm was mentioned in *Computing in Science and Engineering* as being among the top 10 algorithms having the “greatest influence on the development and practice of science and engineering in the 20th century.”
- We have introduced the increasingly popular concept of Bayesian statistics and decision theory with applications.
- A separate chapter on design of experiments, including a discussion on the Taguchi approach, is included.
- The coverage of the book spans most of the important concepts in statistics. Learning the material along with computational examples will prepare students to understand and utilize software procedures to perform statistical analysis.
- Every chapter contains discussion on how to apply the concepts and what the issues are related to applying the theory.
- A student’s solution manual, instructor’s manual, and data disk are provided.
- In the last chapter, we discuss some issues in applications to clearly demonstrate in a unified way how to check for many assumptions in data analysis and what steps one needs to follow to avoid possible pitfalls in applying the methods explained in the rest of this textbook.



# Acknowledgments

We express our sincere appreciation to our late colleague, co-worker, and dear friend, Professor A. N. V. Rao, for his helpful suggestions and ideas for the initial version of the subject textbook. In addition, we thank Bong-jin Choi and Yong Xu for their kind assistance in the preparation of the manuscript. Finally, we acknowledge our students at the University of South Florida for their useful comments and suggestions during the class testing of our book. To all of them, we are very thankful.

**K. M. Ramachandran**  
**Chris P. Tsokos**  
*Tampa, Florida*



# About the Authors

**Kandethody M. Ramachandran** is Professor of Mathematics and Statistics at the University of South Florida. He received his B.S. and M.S. degrees in Mathematics from the Calicut University, India. Later, he worked as a researcher at the Tata Institute of Fundamental Research, Bangalore center, at its Applied Mathematics Division. Dr. Ramachandran got his Ph.D. in Applied Mathematics from Brown University.

His research interests are concentrated in the areas of applied probability and statistics. His research publications span a variety of areas such as control of heavy traffic queues, stochastic delay equations and control problems, stochastic differential games and applications, reinforcement learning methods applied to game theory and other areas, software reliability problems, applications of statistical methods to microarray data analysis, and mathematical finance.

Professor Ramachandran is extensively involved in activities to improve statistics and mathematics education. He is a recipient of the Teaching Incentive Program award at the University of South Florida. He is a member of the MEME Collaborative, which is a partnership among mathematics education, mathematics, and engineering faculty to address issues related to mathematics and mathematics education. He was also involved in the calculus reform efforts at the University of South Florida.

**Chris P. Tsokos** is Distinguished University Professor of Mathematics and Statistics at the University of South Florida. Dr. Tsokos received his B.S. in Engineering Sciences/Mathematics, his M.A. in Mathematics from the University of Rhode Island, and his Ph.D. in Statistics and Probability from the University of Connecticut. Professor Tsokos has also served on the faculties at Virginia Polytechnic Institute and State University and the University of Rhode Island.

Dr. Tsokos's research has extended into a variety of areas, including stochastic systems, statistical models, reliability analysis, ecological systems, operations research, time series, Bayesian analysis, and mathematical and statistical modeling of global warming, among others. He is the author of more than 250 research publications in these areas.

Professor Tsokos is the author of several research monographs and books, including *Random Integral Equations with Applications to Life Sciences and Engineering*, *Probability Distribution: An Introduction to Probability Theory with Applications*, *Mainstreams of Finite Mathematics with Applications*, *Probability with the Essential Analysis*, and *Applied Probability Bayesian Statistical Methods with Applications to Reliability*, among others.

Dr. Tsokos is the recipient of many distinguished awards and honors, including Fellow of the American Statistical Association, USF Distinguished Scholar Award, Sigma Xi Outstanding Research Award, USF Outstanding Undergraduate Teaching Award, USF Professional Excellence Award, URI Alumni Excellence Award in Science and Technology, Pi Mu Epsilon, and election to the International Statistical Institute, among others.



# Flow Chart

This flow chart gives some options on how to use the book in a one-semester or two-semester course. For a two-semester course, we recommend coverage of the complete textbook. However, Chapters 1, 9, and 14 are optional for both one- and two-semester courses and can be given as reading exercises. For a one-semester course, we suggest the following options: A, B, C, D.

