Saddlepoint Approximations with Applications

Modern statistical methods use models that require the computation of probabilities from complicated distributions, which can lead to intractable computations. Saddlepoint approximations can be the answer. Written from the user's point of view, this book explains in clear, simple language how such approximate probability computations are made, taking readers from the very beginnings to current applications.

The book aims to make the subject accessible to the widest possible audience by using graduated levels of difficulty in which the core material is presented in chapters 1–6 at an elementary mathematical level. Readers are guided in applying the methods in various computations that will build their skills and deepen their understanding when later complemented with discussion of theoretical aspects. Chapters 7–9 address the p^* and r^* formulas of higher order asymptotic inference, developed through the Danish approach to the subject by Barndorff-Nielsen and others. These provide a readable summary of the literature and an overview of the subject beginning with the original work of Fisher. Later chapters address special topics where saddlepoint methods have had substantial impact through particular applications. These include applications in multivariate testing, applications to stochastic systems and applied probability, bootstrap implementation in the transform domain, and Bayesian computation and inference.

No previous background in the area is required as the book introduces the subject from the very beginning. Many data examples from real applications show the methods at work and demonstrate their practical value. Ideal for graduate students and researchers in statistics, biostatistics, electrical engineering, econometrics, applied mathematics, and other fields where statistical and probabilistic modeling are used, this is both an entry-level text and a valuable reference.

Ronald W. Butler is the C. F. Frensley Professor of Mathematical Sciences and Professor of Statistical Science at Southern Methodist University, Dallas, Texas.

CAMBRIDGE SERIES IN STATISTICAL AND PROBABILISTIC MATHEMATICS

This series of high-quality upper-division textbooks and expository monographs covers all aspects of stochastic applicable mathematics. The topics range from pure and applied statistics to probability theory, operations research, optimization, and mathematical programming. The books contain clear presentations of new developments in the field and also of the state of the art in classical methods. While emphasizing rigorous treatment of theoretical methods, the books also contain applications and discussions of new techniques made possible by advances in computational practice.

Already published

1. *Bootstrap Methods and Their Application*, by A. C. Davison and D. V. Hinkley
2. *Markov Chains*, by J. Norris
3. *Asymptotic Statistics*, by A. W. van der Vaart
4. *Wavelet Methods for Time Series Analysis*, by Donald B. Percival and Andrew T. Walden
5. *Bayesian Methods*, by Thomas Leonard and John S. J. Hsu
6. *Empirical Processes in M-Estimation*, by Sara van de Geer
7. *Numerical Methods of Statistics*, by John F. Monahan
8. *A User's Guide to Measure Theoretic Probability*, by David Pollard
9. *The Estimation and Tracking of Frequency*, by B. G. Quinn and E. J. Hannan
10. *Data Analysis and Graphics using R*, by John Maindonald and John Braun
11. *Statistical Models*, by A. C. Davison
12. *Semiparametric Regression*, by D. Ruppert, M. P. Wand, R. J. Carroll
13. *Exercises in Probability*, by Loic Chaumont and Marc Yor
14. *Statistical Analysis of Stochastic Processes in Time*, by J. K. Lindsey
15. *Measure Theory and Filtering*, by Lakhdar Aggoun and Robert Elliott
16. *Essentials of Statistical Inference*, by G. A. Young, R. L. Smith
17. *Elements of Distribution Theory*, by Thomas A. Severini
18. *Statistical Mechanics of Disordered Systems*, by Anton Bovier
19. *Random Graph Dynamics*, by Rick Durrett

Saddlepoint Approximations with Applications

Ronald W. Butler
Department of Statistical Science,
Southern Methodist University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521872508

First published 2007

Printed in the United Kingdom at the University Press, Cambridge

A catalog record for this publication is available from the British Library

ISBN 978-0-521-87250-8 hardback

Contents

Preface

Among the various tools that have been developed for use in statistics and probability over the years, perhaps the least understood and most remarkable tool is the saddlepoint approximation. It is remarkable because it usually provides probability approximations whose accuracy is much greater than the current supporting theory would suggest. It is least understood because of the difficulty of the subject itself and the difficulty of the research papers and books that have been written about it. Indeed this lack of accessibility has placed its understanding out of the reach of many researchers in both statistics and its related subjects.

The primary aim of this book is to provide an accessible account of the theory and application of saddlepoint methods that can be understood by the widest possible audience. To do this, the book has been written at graduated levels of difficulty with the first six chapters forming the easiest part and the core of the subject. These chapters use little mathematics beyond the difficulty of advanced calculus (no complex variables) and should provide relatively easy reading to first year graduate students in statistics, engineering, and other quantitative fields. These chapters would also be accessible to senior-level undergraduate mathematics and advanced engineering majors. With the accessibility issue in mind, the first six chapters have been purposefully written to address the issue and should assure that the widest audience is able to read and learn the subject.

The presentation throughout the book takes the point of view of users of saddlepoint approximations; theoretical aspects of the methods are also covered but are given less emphasis than they are in journal articles. This is why, for example, on page 3 of chapter 1 the basic saddlepoint density approximation is introduced without a lot of fuss and used in many computations well before the reader understands what the formulas actually mean. In this way, readers gain practical experience that deepens their understanding when later complemented with the discussion of theoretical aspects. With users in mind, a wide range of practical applications has been included.

Chapters 7–9 address the p^* and r^* formulas of higher order asymptotic inference that have been developed out of the Danish approach to the subject by Barndorff-Nielsen and others. It is unavoidable that the difficulty level must increase somewhat by the very nature of the topics covered here. However the account given for p^* and r^* is considerably simpler and more transparent than much of the literature and these chapters should still be quite readable to the better first year graduate students in statistics. The chapter shows the evolution of the ideas starting from Fisher's original formulation of the p^* formula, through the work of Fraser, and Efron and Hinkley and ending with r^* and related approximations due to Skovgaard, and Fraser and Reid.

Chapters 10–16 address important special topics where saddlepoint methods have had substantial impact in particular applications or subjects. For example, chapter 11 provides a comprehensive survey of the use of saddlepoint approximations in multivariate testing. The majority of commonly used multivariate tests may now be implemented by using the more accurate saddlepoint methods instead of the usual statistical package procedures. All of these results are organized in one place with a common notation.

Applications to stochastic systems and applied probability are presented in chapter 13. The emphasis here is on approximation to first passage times in stochastic systems because such computations underlie the subject of system reliability. The subject is also basic to transfer function computation and inversion, which are encountered with many electrical and control systems in the engineering sciences. This chapter should appeal directly to electrical engineers, who increasingly are embracing saddlepoint methods out of the need to invert Laplace transforms that often arise in the mathematical models of the engineering sciences.

Saddlepoint methods are also useful in avoiding much of the simulation requisite when implementing another modern statistical tool, the bootstrap. Chapter 14 shows how the bootstrap may be implemented in the transform domain thus forgoing the usual resampling. The emphasis is on multistate survival models that are used to represent degenerative medical conditions in biostatistics. Chapter 15 shows how Bayesian computation and inference may also benefit from using saddlepoint approximations particularly in settings for which the likelihood function may not be tractable when using the standard methods.

The audience for the book includes graduate students and faculty in statistics, biostatistics, probability, engineering sciences, applied mathematics, and other subjects wherein more sophisticated methods of statistical and probabilistic modeling are used. The book is an entry level text from which readers may learn the subject for the first time. The book does not attempt to cover the most advanced aspects of the subject as one would find in Jensen (1995), Field and Ronchetti (1990), and Kolassa (1994) but rather provides much of the background needed to begin to understand these more advanced presentations. These more advanced monographs also require a relatively strong background in complex analysis, something that the majority of statisticians and biostatisticians lack. While complex analysis cannot be avoided in an advanced treatment of the subject, the use of such methods assures a rather narrow audience and this is contrary to the aims of this book.

Acknowledgements

During the years needed to write this book, my own understanding of saddlepoint methods and their potential uses has deepened and broadened. With this I have gained both an appreciation of the efforts by fellow researchers and a deep respect and admiration for the original work by the late Henry Daniels who led the way for us all. My hope is that readers from very diverse backgrounds will be able to acquire similar appreciation while putting forth far less effort.

I would like to thank all those who have helped in some manner to make this book possible. In particular, James Booth, Douglas Bronson, Robert Paige, Marc Paolella, and Andrew Wood provided invaluable help for which I am very grateful. Chapters 7–10 were written on a sabbatical leave at the University of Sydney and I thank John Robinson for an enjoyable and productive visit. Financial aid for this work was provided by the National Science Foundation.