Lecture Notes in Statistics
Proceedings

Statistical Challenges in Modern Astronomy V

Springer

Lecture Notes in Statistics 209

Edited by
P. Bickel, P. Diggle, S. Fienberg, U. Gather,
I. Olkin, S. Zeger

For further volumes:
http://www.springer.com/series/694

Eric D. Feigelson • G. Jogesh Babu
Editors

Statistical Challenges in Modern Astronomy V

Springer

Editors
Eric D. Feigelson
Department of Astronomy
and Astrophysics
Pennsylvania State University
University Park, PA, USA

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

ISBN 978-1-4614-3519-8 ISBN 978-1-4614-3520-4 (eBook)
DOI 10.1007/978-1-4614-3520-4
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012939628

© Springer Science+Business Media New York 2013
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. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
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.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Twenty years ago, the first *Statistical Challenges in Modern Astronomy* (SCMA) conference was held at Penn State University. Serving as a gathering of two scholarly communities with common interests, SCMA meetings have been held every 5 years for cross-disciplinary discussions of methodological issues arising in astronomical research. These are the proceedings of the fifth SCMA conference held in June 2011. While some of the topics are the similar as those in the 1991 meeting, the level of sophistication and accomplishment has enormously increased. Astronomers and statisticians worldwide have developed collaborations to address some of the most challenging and important problems facing astronomy today. These involve data mining enormous datasets from widefield surveys obtained with major new telescope systems, fitting of cosmological and other astrophysical models to complex datasets, and studying the temporal behaviors of innumerable variable objects in the sky. Bayesian inference has gained considerable momentum in astrophysical model fitting. These advanced methods are gaining attention outside of the world of expert astrostatisticians, as the broad astronomical community realize that twenty-first century science goals can not be achieved with nineteenth and twentieth century statistical methods. At SCMA V, both young and experienced astrostatisticians presented work and engaged in discussions on how these problems can be best addressed.

The proceedings are divided into six sections; most invited talks are followed by invited commentaries by scholars in the other field. The volume begins with five talks on *Statistics in Cosmology* demonstrating significant recent accomplishments in this most-important field of astronomy and astrophysics. Modern accomplishments of modern quantitative cosmology rely heavily on sophisticated statistical analysis of large datasets. Topics reviewed include likelihood-free estimation of quasar luminosity functions (Schaefer and Freeman), estimation of galaxy photometric redshifts and quantification of voids in galaxy Large-Scale Structure (Wandelt), inference based on comparing data to cosmological simulations (Higdon), likelihood estimation of gravitational lensing of the cosmic microwave background (CMB) radiation (Anderes), and application of needlets to cosmic microwave background studies (Marinucci).

The second section provides a sampling of the growing applications of *Bayesian Analysis Across Astronomy*. Here we have both invited reviews by senior researchers, and a sampling of the many works by younger researchers. The reviews discuss Bayesian models constructed to model galaxy star formation histories (Weinberg), model selection within the consensus ΛCDM cosmological model family (Trotta), and measurement errors in astronomical regression and density estimation problems (Kelly). The shorter talks treat asteroseismology (Benomar), event detection in time series (Blocker and Protopapas), reverberation mapping in active galactic nuclei (Brewer), modeling of Poisson images (Guglielmetti et al.), treatment of instrument calibration errors (Kashyap et al.), modeling of Type Ia supernova data (Mandel), and faint source flux estimation (Switzer et al.). Advanced methods for hierarchical modeling and Monte Carlo Markov Chain computational techniques are discussed in many of these talks and associated commentaries.

The third section of the proceedings address the use of modern techniques techniques of *Data Mining and Astroinformatics* for the analysis of massive datasets emerging from many new observatories. Compressive sensing, an extension of wavelet analysis, is very promising for many problems (Starck). Diffusion maps can treat non-linear structures in high-dimensional datasets (Lee and Freeman). Nearest neighbor techniques are used for outlier detection in megadatasets (Borne and Vedachalam). Bayesian approaches can help cross-identification of sources between astronomical catalogs (Budavári). Likelihood-based data compression can assist parameter estimation in large datasets (Jimenez).

The fourth section considers challenges arising in astronomical *Image and Time Series Analysis*. Techniques of mathematical morphology are applied to classifying sunspots (Stenning et al.). Realistic images are simulated using knowledge of celestial populations and telescope characteristics (Connolly et al.). Structure recognition algorithms are discussed for three-dimensional astronomical datacubes (Rosolowsky). The problem of locating faint transient sources in multiepoch image datasets is addressed by controlling the False Discovery Rate (Clements et al.). Wavelets are a valuable tool for modeling irregularly spaced time series (Mondal and Percival).

The fifth section provides perspectives on *The Future of Astrostatistics*. The field is gaining a presence in international organizations (Hilbe). The public domain **R** statistical computing environment is a very promising new software environment to implement existing and develop new statistical analyses for astronomical research (Tierney). A Panel Discussion discusses various aspects of astrostatistical practice and research for the coming decade (van Dyk, Feigelson, Loredo, Scargle). The final section of the proceedings gives brief presentations of the contributed posters. Many fascinating problems and sophisticated statistical methods are described.

The work of many individuals and organizations contributed to the success of the SCMA V conference. The invited speakers and cross-disciplinary commentators were the central pillar of the conference, and we are grateful for their presentations and manuscripts. Staff in the Departments of Statistics and Astronomy and Astrophysics provided administrative support. Funding support for the conference was provided by the two departments, Penn State's Eberly College of Science, and the National Science Foundation through grant AST-1113001. Finally, we are appreciative of our families' support during the many phases of this conference organization.

Pennsylvania State University, PA, USA Eric D. Feigelson
Pennsylvania State University, PA, USA G. Jogesh Babu

Contents

Part I Statistics in Cosmology

Part II Bayesian Analysis Across Astronomy

Part III Data Mining and Astroinformatics

Part IV Image and Time Series Analysis

Part V The Future of Astrostatistics

Part VI Contributed Papers

Contributors

Mark Allen Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA, USA

Steven W. Allen Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA, USA

Ethan Anderes Statistics Department, University of California, Davis, CA, USA

Keith Arnaud NASA Goddard Space Flight Center, Greenbelt, MD, USA

Zsolt Bagoly Eötvös University, Budapest, Hungary

Paul D. Baines Department of Statistics, University of California, Davis, CA, USA

Lajos G. Balázs Konkoly Observatory, Budapest, Hungary

Nicholas M. Ball National Research Council Herzberg Institute of Astrophysics, Victoria, BC, Canada

Iréne Balmés Laboratoire Univers et Théories (LUTh), UMR 8102 CNRS, Observatoire de Paris, Université Paris Diderot, Meudon, France

A.J. Banday Institut de Recherche in Astrophysique et Planétologie, Université de Toulouse, Toulouse, France

David Banks Department of Statistical Science, Duke University, Durham, NC, USA

Richard Baraniuk Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA

D. Bard SLAC National Accelerator Laboratory, Menlo Park, CA, USA

Graham Barnes CoRA, NorthWest Research Associates, Boulder, CO, USA

Shantanu Basu Department of Physics and Astronomy, London, ON, Canada

Guillaume Belanger European Space Agency, Villanueva de la Canada, Spain

Othman Benomar School of Physics, University of Sydney, Sydney, Australia

Sharmodeep Bhattacharyya Department of Statistics, University of California, Berkeley, CA, USA

P. Bielewicz Centre d'Étude Spatiale des Rayonnements, Toulouse Cedex 4, France

Alexander W. Blocker Department of Statistics, Harvard University, Cambridge, MA, USA

Joshua S. Bloom Department of Astronomy, University of California, Berkeley, CA, USA

John Bochanski Department of Astronomy and Astrophysics, Penn State University, University Park, PA, USA

Hans Böhringer Max-Planck-Institut für extraterrestrische Physik, Garching, Germany

Kirk D. Borne Department of Computational and Data Sciences, George Mason University, Fairfax, VA, USA

Brendon J. Brewer Department of Physics, University of California, Santa Barbara, CA, USA

Tamara Broderick Department of Statistics, University of California, Berkeley, CA, USA

Tamás Budavári Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD, USA

Nathaniel R. Butler School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA

Yonk-Ik Byun Department of Astronomy, Yonsei University, Seoul, Korea

Seo-Won Chang Department of Astronomy, Yonsei University, Seoul, Korea

David Chernoff Cornell University, Ithaca, NY, USA

Jessica Cisewski Department of Statistics and Operations Research, University of North Carolina, Carrboro, NC, USA

Nicolle Clements Department of Statistics, Fox School of Business and Management, Temple University, Philadelphia, PA

Andrew Connolly Department of Astronomy, University of Washington, Seattle, WA, USA

Alanna Connors Eureka Scientific, Oakland CA, USA

Thomas M. Crawford The Kavli Institute for Cosmological Physics, The University of Chicago, Chicago, IL, USA

Istvan Csabai Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD, USA

Nathan Deg Department of Astronomy, Queens University, Kingston, ON, Canada

Stewart DeSoto Department of Physics, Wheaton College, Wheaton, IL, USA

Volker Dose Max-Planck-Institut für Plasmaphysik, Garching, Germany

David Donoho Statistics Department, Stanford University, Stanford, CA, USA

Jeremy Drake Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

Garcia Etchegaray Estefania, Department of Statistics, Baker Hall, Carnegie Mellon University, Pittsburgh, PA, USA

Benjamin Farr Dearborn Observatory, Northwestern University, Evanston, IL, USA

Krzysztof Findeisen Department of Astronomy, California Institute of Technology, Pasadena, CA, USA

Rainer Fischer Max-Planck-Institut für Plasmaphysik, Garching, Germany

Didier Fraix-Burnet Université Jules Fournier, Grenoble Cedex 9, France

Peter E. Freeman Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, USA

Christopher Genovese Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, USA

Lluís Gil Tavernes Blanques, University of València, València, Spain

Facundo Gomez Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA

Carlo Graziani Department of Astronomy, University of Chicago, Chicago, IL, USA

Philip C. Gregory Physics and Astronomy, University of British Columbia, Vancouver, BC, Canada

Dirk Grupe Department of Astronomy and Asrophysics, Penn State University, University Park, PA, USA

Fabrizia Guglielmetti Max Planck Institut für Exterterrestrischephysik, Giessenbachstr, Garching bei Muenchen, Germany

Wenge Guo Department of Mathematical Sciences, New Jersey Institute of Technology, University Heights, Newark, NJ

Jaegyoon Hahm Supercomputing Center, Korea Institute of Science and Technology Information, Daejeon, Korea

Kenji Hamaguchi NASA Goddard Space Flight Center, Greenbelt, MD, USA

Jiangang Hao Fermi National Laboratory, Batavia, IL, USA

Salman Habib Argonne National Laboratory, Argonne, IL, USA

Alan Heavens SUPA, Institute for Astronomy, Royal Observatory of Edinburgh, Edinburgh, UK

Katrin Heitmann Argonne National Laboratory, Argonne, IL 60439, USA

Martin A. Hendry School of Physics and Astronomy, Kelvin Building, University of Glasgow, Glasgow, UK

David Higdon Los Alamos National Laboratory, Los Alamos, NM, USA

Joseph M. Hilbe School of Mathematics and Statistical Sciences, Arizona State University, Tempe, AZ, USA

Darren Homrighausen Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, USA

István Horváth Department of Physics, Bolyai Military University, Budapest, Hungary

Talvikki Hovatta Department of Physics, Purdue University, West Lafayette, IN, USA

R. Izbicki Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, USA

Fadili M. Jalal Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen, École Nationale Superieure d'Ingéntiers, Caen, France

John Berian James Niels Bohr Institute, Copenhagen, Denmark

Jens Jasche Argelander-Institut für Astronomie, Bonn, Germany

Garrett Jernigan Space Sciences Laboratory, University California, Berkeley, CA, USA

Raul Jimenez Instituto de Ciencias del Cosmos, Facultad de Fisica, Universidad de Barcelona, Barcelona, Spain

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

Vinay L. Kashyap Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

Neal Katz Department of Astronomy, University of Massachusetts, Amherst, MA, USA

David Keitel Max-Planck-Institut für Gravitationsphysik, Hannover, Germany

Brandon C. Kelly Department of Physics, University of California, Santa Barbara CA, USA

Michael Kuhn Department of Astronomy and Astrophysics, Penn State University, University Park, PA, USA

Aleksandra Kurek Astronomical Observatory, Jagiellonian University, Krakow, Poland

Guilhem Lavaux Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, Canada

Earl Lawrence Los Alamos National Laboratory, Los Alamos, NM, USA

Ann B. Lee Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, USA

Thomas C.M. Lee Department of Statistics, University of Statistics, Davis, CA, USA

Adrienne Leonard IRFU Service d'Astrophysique, Centre d'Études Atomiques, Gif sur Yvette Cedex, France

James Long Department of Statistics, University of California, Berkeley, CA, USA

Thomas J. Loredo Department of Astronomy, Cornell University, Ithaca, NY, USA

Yu Lu Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA, USA

Daniel Pereira Machado Service d'Astrophysique IRFU, Centre d'Études Atomiques de Saclay, Paris, France

Barry Madore Carnegie Observatories, Pasadena, CA, USA

Kaisey S. Mandel Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

Adam Mantz NASA Goddard Space Flight Center, Greenbelt, MD, USA

Domenico Marinucci Department of Mathematics, University of Rome, Roma, Italy

Elizabeth Martinez-Gomez Center for Astrostatistics, Penn State University, University Park, PA, USA

Kushal Mehta Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

Attila Mészáros Faculty of Mathematics and Physics, Charles University, Prague 8, Czech Republic

Brendan Miller Department of Astronomy, University of Michigan, Ann Arbor, MI, USA

H.J. Mo Department of Astronomy, University of Massachusetts, Amherst, MA, USA

Debashis Mondal Department of Statistics, University of Chicago, Chicago, IL, USA

Adam N. Morgan Department of Astronomy, University of California, Berkeley, CA, USA

Kellen Murphy Department of Physics and Astronomy, Ohio University, Athens, OH, USA

J. Newman Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA, USA

Tim Naylor School of Physics, University of Exeter, Exeter, UK

Mark Neyrinck Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD, USA

Fabio Noviello CNRS and Université Paris-Sud, Orsay, France

Martin Paegert 6301 Stevenson Ctr, Vanderbilt University, Nashville, TN, USA

Paykari Paniez Service d'Astrophysique, Centre dÉtudes Atomiques de Saclay, Orme des Merisiers, Gif-sur-Yvette, France

M.A. Papa Max Planck Institute for Gravitational Physics, Potsdam, Germany

Donald B. Percival Department of Statistics, University of Washington, Seattle, WA, USA

John Peterson Department of Physics, Purdue University, West Lafayette, IN, USA

Rosati Piero European Southern Observatory, Garching, Germany

Andrew Pollock XMM-Newton Science Operations Center, European Space Astronomy Centre, Madrid, Spain

Adrian Pope Argonne National Laboratory, Santa Fe, NM, USA

Jennifer Posson-Brown Harvard-Smithsonian Center for Astrophysics, Cambridge MA, USA

R. Prix Max Planck Institute for Gravitational Physics, Hannover, Germany

Pavlos Protopapas Harvard-Smithsonian Center for Astrophysics, Harvard University, Cambridge, MA, USA

David Rapetti Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

Kathy Reeves Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

Christian L. Reichardt Department of Physics, University of California, Berkeley, CA, USA

John Rice Department of Statistics, University of California, Berkeley, CA, USA

Joseph W. Richards Department of Statistics and Astronomy, University of California, Berkeley, CA, USA

Erik Rosolowsky University of British Columbia, Kelowna, BC, Canada

Gerald Ruch Department of Physics, University of St. Thomas, Saint Paul, MN, USA

David Ruppert School of Operations Research and Information Engineering, Cornell University, Ithaca, NY, USA

Claes-Erik Rydberg Department of Astronomy, Stockholm University, Stockholm, Sweden

Steve Saar Harvard-Smithsonian Center for Astrophysics, Cambridge MA, USA

Gendith Sardane Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA, USA

Sanat K. Sarkar Department of Statistics, Fox School of Business and Management, Temple University, Philadelphia, PA

Jeffrey D. Scargle Planetary Systems Branch, NASA Ames Research Center, Moffett Field, CA, USA

Chad M. Schafer Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, USA

Sam Schmidt Department of Physics, University of California, Davis, CA, USA

P. Schneider Argelander-Institut für Astronomie, University of Bonn, Bonn, Germany

Carolyn Sealfon Department of Physics, West Chester University, Merion Science Center, West Chester, PA, USA

Marco Selig Max Planck Institut für Astronomie, Garching, Germany

Jonathan Sick Department of Physics, Queen’s University, Kingston, ON, Canada

Jennifer Siegal-Gaskins Center for Cosmology and Astro-Particle Physics, Ohio State University, Columbus, OH, USA

Aneta Siemiginowska Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

Maham Siddiqi Max Planck Institute for Gravitational Physics, Hannover, Germany

Dejan Slepćev Department of Mathematics, Carnegie Mellon University, Pittsburgh, PA, USA

Kunlaya Soiaporn School of Operations Research and Information Engineering, Cornell University, Ithaca, NY, USA

Jean-Luc Starck Service d'Astrophysique, Centre d'Études Atomiques de Saclay, Orme des Merisiers, Gif-sur-Yvette, France

Dan L. Starr Astronomy Department, University of California, Berkeley, CA, USA

Nathan Stein Department of Statistics, Harvard University, Science Center, Cambridge, MA, USA

David Stenning Department of Statistics, University of California, Irvine, CA, USA

Eric R. Switzer Kavli Institute Cosmological Physics, University of Chicago, Chicago, IL, USA

Dorottya Szécsi Eötvös Loránd University, Budapest, Hungary

Ismael Tereno Department de Física, Universidade de Lisboa, Lisbon, Portugal

Luke Tierney Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA, USA

Paolo Tozzi Osservatorio Astronomico di Trieste, Trieste, Italy

Roberto Trotta Astrophysics Group, Blackett Laboratory, Imperial College London, London, UK

Irina S. Udaltsova Department of Statistics, University of California, Davis CA, USA

David A. van Dyk Department of Mathematics, Imperial College London, London, UK

Alexander van Engelen McGill University, Montréal, QC, Canada

Roland Vavrek European Space Astronomy Centre, Villanueva de la Cañada, Madrid, Spain

Arun Vedachalam Department of Computational and Data Sciences, George Mason University, Fairfax, MD, USA

Licia Verde Instituto de Ciencias del Cosmos, Facultad de Fisica, Universidad de Barcelona, Barcelona, Spain

Péter Veres Department of Astronomy & Astrophysics, Pennsylvania State University, University Park, PA, USA

Ricardo Vilalta Department of Computer Science, University of Houston, Houston, TX, USA

Benjamin D. Wandelt Institut d'Astrophysique, Paris, France

Ira Wasserman Department of Astronomy, Cornell University, Ithaca, NY, USA

Michael Way NASA Ames Research Center, Moffett Field, CA, USA

Martin D. Weinberg Department of Astronomy, Lederle Graduate Research Center, University of Massachusetts, Amherst, MA, USA

Philipp Wilking Institut für Astronomie, University of Bonn, Bonn, Germany

Jin Xu Department of Statistics, University of California, Irvine, CA, USA

Ilsang Yoon Department of Astronomy, University of Massachusetts, Amherst, MA, USA

C. Alex Young NASA Goddard Space Flight Center, Greenbelt, MD, USA

Andreas Zezas Department of Physics, University of Crete, Heraklion, Greece

Part I
Statistics in Cosmology

Chapter 1
Likelihood-Free Inference in Cosmology: Potential for the Estimation of Luminosity Functions

Chad M. Schafer and Peter E. Freeman

Abstract Statistical inference of cosmological quantities of interest is complicated by significant observational limitations, including heteroscedastic measurement error and irregular selection effects. These observational difficulties exacerbate challenges posed by the often-complex relationship between estimands and the distribution of observables; indeed, in some situations it is only possible to simulate realizations of observations under various assumed cosmological theories. When faced with these challenges, one is naturally led to consider utilizing repeated simulations of the full data generation process, and then comparing observed and simulated data sets to constrain the parameters. In such a scenario, one would not have a likelihood function relating the parameters to the observable data. This paper will present an overview of methods that allow a likelihood-free approach to inference, with emphasis on approximate Bayesian computation, a class of procedures originally motivated by similar inference problems in population genetics.

1.1 Introduction

The ever-increasing efforts to build catalogs of astronomical objects, and to measure key properties of these objects, is, in large part, motivated by the goal of inferring unknown constants that characterize the Universe. This paper seeks to present an example of such a problem, and to describe some of the features of the data and their collection that complicates what is otherwise a standard statistical inference problem. To an outsider of this field, it can be surprising the extent to which

---

C.M. Schafer (✉) • P.E. Freeman
Department of Statistics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
e-mail: cschafer@stat.cmu.edu; pfreeman@stat.cmu.edu

E.D. Feigelson and G.J. Babu (eds.), *Statistical Challenges in Modern Astronomy V*,
Lecture Notes in Statistics 209, DOI 10.1007/978-1-4614-3520-4_1,
© Springer Science+Business Media New York 2013

Table 1.1 Examples of key cosmological parameters

Parameter	Description	In Fig. 1.1[a]
Ω_m	Ratio of total matter density to that needed for a flat Universe	0.266
Ω_Λ	Similar to Ω_m, but for dark energy density	0.734
H_0	Hubble constant: the current expansion rate of the Universe	71.0 km/s/Mpc

[a] Estimates based on WMAP7 [2]

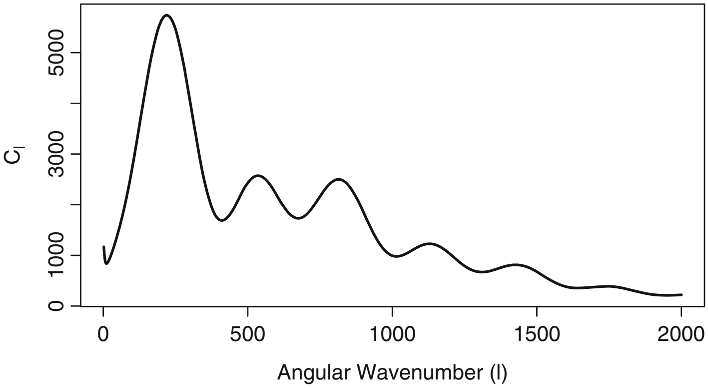

Fig. 1.1 Power spectrum, a function of cosmological parameters, of fluctuations in the temperature of photons that comprise the cosmic microwave background (CMB). The parameter values are fixed to those shown in Table 1.1

many questions regarding the nature of Universe have been boiled down to the estimation of a relatively small number of *cosmological parameters*. Table 1.1 gives some examples of these physical constants. Carefully-derived cosmological theory posits relationships between these parameters and the distribution of observables. In (relatively) simple situations, the distribution of the data is of a “standard” form, and the likelihood function can be derived. This allows for utilization of well-established methods of inference, including finding maximum likelihood estimates or exploring the posterior distribution of these parameters given the observed data.

One of the most important inference problems that fits into this framework is the estimation of cosmological parameters using fluctuations in the temperature of photons that comprise the cosmic microwave background (CMB). These photons are remnants of the time, only 300,000 years after the Big Bang, when the temperature of the Universe had cooled sufficiently for light to travel freely. The slight variation in the temperature of these photons encodes important information regarding the nature of the early Universe; the amount of correlation on different angular scales has been characterized as a function of cosmological parameters. Figure 1.1 shows the *power spectrum* that describes the Gaussian process on the sphere used to model the process; this power spectrum corresponds to the parameter values shown in Table 1.1. A succession of experiments has observed this background radiation to greater precision, and hence has achieved stronger constraints on the unknowns. The estimates in Table 1.1 are based on the recent WMAP 7 data release [2].

The relationship between the cosmological parameters and the power spectrum of the CMB fluctuations is complex: It is highly nonlinear, and there are strong degeneracies between some the parameters. The complexity of this relationship presents its own challenges. Bayesian methods dominate in cosmology, and MCMC is feasible in this situation; one only needs to make small steps in the cosmological parameter space, and the parameter vectors are mapped into the corresponding power spectrum, which in turn defines the likelihood function for the data. Schafer and Stark [3] presents a Monte Carlo method for constructing confidence regions of optimal expected size that is specifically motivated by this type of situation. Yet, both of these methods rely upon knowledge of the likelihood function of the data. Increasingly, we are faced with situations in which this is not a reasonable assumption. This may be because the distribution of the data is inherently complex, or it may be because of data corrupted by irregular truncation effects and/or heteroscedastic measurement error with complex dependence structure.

This paper describes *likelihood-free* approaches to inference, in particular, *approximate Bayesian computation* (ABC). The term "likelihood-free" is not intended to imply that a likelihood function does not exist in these applications; instead, it is the case that the likelihood function is too complex to admit a form that can be evaluated reliably for different values of the parameters of interest. These procedures will instead be built upon repeated simulation of the data-generating process (allowing for the incorporation of any complex computer models, data contamination, or selection effects) and then comparing simulated with observed data. Implementation of these approaches presents their own set of challenges. The difficulty of deriving an appropriate likelihood function is replaced with that of finding an approximate *sufficient statistic* for the parameter of interest. There are also computational challenges to implementing these procedures, but these can be mitigated via the design of efficient algorithms. This paper will present a brief introduction to some techniques and directions for addressing these challenges.

Another objective of this paper is to allow a reader familiar with statistical inference, but not with astronomy, the chance to learn some background on a relatively simple cosmological inference problem that possesses some of the aforementioned challenges. In the next section we will present two examples, with background information. The first is a stylized example of estimating cosmological parameters using observations of Type Ia supernovae. This example serves largely to introduce important concepts and methods. The second is the problem of estimating a bivariate *luminosity function*, the distribution of astronomical objects of interest as a function of their distance and the amount of light they emit. We will utilize the quasar catalog of [4] to motivate a promising approach to estimating the bivariate luminosity function which relies upon forward simulation of the full data generation process.

1.2 Examples and Astronomical Background

In this section we will present two examples of statistical inference using astronomical data. The first is relatively simple and will serve only to demonstrate basic likelihood-free techniques. The second application possesses the type of complications that motivate the consideration of these approaches. Both of these build upon the same astronomical background, including the following key quantities described below.

Key Quantities in the Examples

1. **Redshift** (often denoted z)—Because the Universe is expanding, light emitted by an astronomical object is shifted to longer wavelengths prior to reaching the observer: the ratio of the wavelength at which the light is observed to the wavelength when emitted equals $1+z$. Since the magnitude of this shift increases as a function of the time since the light was emitted, redshift is often taken as a (nonlinear) proxy for time (or distance). For the current epoch, $z=0$; for quasars, $z \leq 7$; and for the CMB, the most distant structure yet observed in the Universe, $z \approx 1089$.

2. **Apparent magnitude** (m)—The brightness of the object as measured by the observer. Magnitudes are measured on a logarithmic scale such that *decreasing* the magnitude by five corresponds to changing the brightness by a factor of 100. The root of the magnitude system was the classification of stars by the Greek astronomer Hipparchus, who used one for the brightest stars and six for the faintest.

3. **Absolute magnitude** (M)—The apparent magnitude of that an object would have if it were located 10 pc (or about 32 light-years) from Earth. The relationship between m and M in a flat Universe can be written as

$$M = m - \frac{(1+z)}{c\,H_0} \int_0^z \left(\Omega_m (1+u)^3 + \Omega_\Lambda\right)^{-0.5} du, \tag{1.1}$$

where c is the speed of light, and H_0, Ω_m, and Ω_Λ are among the cosmological parameters shown in Table 1.1.

Equation 1.1 establishes a relationship between a measurable property of astronomical objects (the apparent magnitude), and a scientifically useful quantity (the absolute magnitude). Note how this transformation depends not only on the redshift of the object, but on the values of unknown physical constants. In the examples that follow, this expression will be utilized in different ways. In the first case, Type Ia