

Cambridge University Press
052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative
Approach with Mathematica™ Support
P. C. Gregory
Frontmatter
[More information](#)

Bayesian Logical Data Analysis for the Physical Sciences

A Comparative Approach with *Mathematica*™ Support

Increasingly, researchers in many branches of science are coming into contact with Bayesian statistics or Bayesian probability theory. By encompassing both inductive and deductive logic, Bayesian analysis can improve model parameter estimates by many orders of magnitude. It provides a simple and unified approach to all data analysis problems, allowing the experimenter to assign probabilities to competing hypotheses of interest, on the basis of the current state of knowledge.

This book provides a clear exposition of the underlying concepts with large numbers of worked examples and problem sets. The book also discusses numerical techniques for implementing the Bayesian calculations, including an introduction to Markov chain Monte Carlo integration and linear and nonlinear least-squares analysis seen from a Bayesian perspective. In addition, background material is provided in appendices and supporting *Mathematica* notebooks are available from www.cambridge.org/052184150X, providing an easy learning route for upper-undergraduate, graduate students, or any serious researcher in physical sciences or engineering.

PHIL GREGORY is Professor Emeritus at the Department of Physics and Astronomy at the University of British Columbia.

Cambridge University Press

052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support

P. C. Gregory

Frontmatter

[More information](#)

Bayesian Logical Data Analysis for the Physical Sciences

A Comparative Approach with *Mathematica*[™] Support

P. C. Gregory

Department of Physics and Astronomy, University of British Columbia



Cambridge University Press
 052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative
 Approach with Mathematica™ Support
 P. C. Gregory
 Frontmatter
[More information](#)

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
 The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
 The Edinburgh Building, Cambridge CB2 2RU, UK
 40 West 20th Street, New York, NY 10011-4211, USA
 477 Williamstown Road, Port Melbourne, VIC 3207, Australia
 Ruiz de Alarcón 13, 28014 Madrid, Spain
 Dock House, The Waterfront, Cape Town 8001, South Africa

<http://www.cambridge.org>

© Cambridge University Press 2005

This book is in copyright. Subject to statutory exception
 and to the provisions of relevant collective licensing agreements,
 no reproduction of any part may take place without
 the written permission of Cambridge University Press.

First published 2005

Printed in the United Kingdom at the University Press, Cambridge

Typeface in Times 10/13pt System Advent 3B2 8.01 [TB]

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

Library of Congress Cataloging in Publication data

Gregory, P. C. (Philip Christopher), 1941–

Bayesian logical data analysis for the physical sciences: a comparative approach with
Mathematica support / P.C. Gregory.

p. cm.

Includes bibliographical references and index.

ISBN 0 521 84150 X (hardback)

1. Bayesian statistical decision theory. 2. Physical sciences – Statistical methods.

3. Mathematica (Computer file) I. Title.

QA279.5.G74 2005

519.5'42 – dc22 2004045930

ISBN 0 521 84150 X hardback

The publisher has used its best endeavors to ensure that the URLs for external websites referred to in this book are correct and active at the time of going to press. However, the publisher has no responsibility for the websites and can make no guarantee that a site will remain live or that the content is or will remain appropriate.

Disclaimer of warranty

We make no warranties, express or implied, that the programs contained in this volume are free of error, or are consistent with any particular standard of merchantability, or that they will meet your requirements for any particular application. They should not be relied on for solving a problem whose incorrect solution could result in injury to a person or loss of property. If you do use the programs in such a manner, it is at your own risk. The authors and publisher disclaim all liability for direct or consequential damages resulting from your use of the programs.

Contents

<i>Preface</i>	page	xiii
Software support		xv
<i>Acknowledgements</i>		xvii
1 Role of probability theory in science		1
1.1 Scientific inference		1
1.2 Inference requires a probability theory		2
1.2.1 The two rules for manipulating probabilities		4
1.3 Usual form of Bayes' theorem		5
1.3.1 Discrete hypothesis space		5
1.3.2 Continuous hypothesis space		6
1.3.3 Bayes' theorem – model of the learning process		7
1.3.4 Example of the use of Bayes' theorem		8
1.4 Probability and frequency		10
1.4.1 Example: incorporating frequency information		11
1.5 Marginalization		12
1.6 The two basic problems in statistical inference		15
1.7 Advantages of the Bayesian approach		16
1.8 Problems		17
2 Probability theory as extended logic		21
2.1 Overview		21
2.2 Fundamentals of logic		21
2.2.1 Logical propositions		21
2.2.2 Compound propositions		22
2.2.3 Truth tables and Boolean algebra		22
2.2.4 Deductive inference		24
2.2.5 Inductive or plausible inference		25
2.3 Brief history		25
2.4 An adequate set of operations		26
2.4.1 Examination of a logic function		27
2.5 Operations for plausible inference		29

Cambridge University Press

052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support

P. C. Gregory

Frontmatter

[More information](#)

vi	<i>Contents</i>	
	2.5.1 The desiderata of Bayesian probability theory	30
	2.5.2 Development of the product rule	30
	2.5.3 Development of sum rule	34
	2.5.4 Qualitative properties of product and sum rules	36
	2.6 Uniqueness of the product and sum rules	37
	2.7 Summary	39
	2.8 Problems	39
3	The how-to of Bayesian inference	41
	3.1 Overview	41
	3.2 Basics	41
	3.3 Parameter estimation	43
	3.4 Nuisance parameters	45
	3.5 Model comparison and Occam's razor	45
	3.6 Sample spectral line problem	50
	3.6.1 Background information	50
	3.7 Odds ratio	52
	3.7.1 Choice of prior $p(T M_1, I)$	53
	3.7.2 Calculation of $p(D M_1, T, I)$	55
	3.7.3 Calculation of $p(D M_2, I)$	58
	3.7.4 Odds, uniform prior	58
	3.7.5 Odds, Jeffreys prior	58
	3.8 Parameter estimation problem	59
	3.8.1 Sensitivity of odds to T_{\max}	59
	3.9 Lessons	61
	3.10 Ignorance priors	63
	3.11 Systematic errors	65
	3.11.1 Systematic error example	66
	3.12 Problems	69
4	Assigning probabilities	72
	4.1 Introduction	72
	4.2 Binomial distribution	72
	4.2.1 Bernoulli's law of large numbers	75
	4.2.2 The gambler's coin problem	75
	4.2.3 Bayesian analysis of an opinion poll	77
	4.3 Multinomial distribution	79
	4.4 Can you really answer that question?	80
	4.5 Logical versus causal connections	82
	4.6 Exchangeable distributions	83
	4.7 Poisson distribution	85

Cambridge University Press

052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support

P. C. Gregory

Frontmatter

[More information](#)

<i>Contents</i>		vii
4.7.1	Bayesian and frequentist comparison	87
4.8	Constructing likelihood functions	89
4.8.1	Deterministic model	90
4.8.2	Probabilistic model	91
4.9	Summary	93
		94
5	Frequentist statistical inference	96
5.1	Overview	96
5.2	The concept of a random variable	96
5.3	Sampling theory	97
5.4	Probability distributions	98
5.5	Descriptive properties of distributions	100
5.5.1	Relative line shape measures for distributions	101
5.5.2	Standard random variable	102
5.5.3	Other measures of central tendency and dispersion	103
5.5.4	Median baseline subtraction	104
5.6	Moment generating functions	105
5.7	Some discrete probability distributions	107
5.7.1	Binomial distribution	107
5.7.2	The Poisson distribution	109
5.7.3	Negative binomial distribution	112
5.8	Continuous probability distributions	113
5.8.1	Normal distribution	113
5.8.2	Uniform distribution	116
5.8.3	Gamma distribution	116
5.8.4	Beta distribution	117
5.8.5	Negative exponential distribution	118
5.9	Central Limit Theorem	119
5.10	Bayesian demonstration of the Central Limit Theorem	120
5.11	Distribution of the sample mean	124
5.11.1	Signal averaging example	125
5.12	Transformation of a random variable	125
5.13	Random and pseudo-random numbers	127
5.13.1	Pseudo-random number generators	131
5.13.2	Tests for randomness	132
5.14	Summary	136
5.15	Problems	137
6	What is a statistic?	139
6.1	Introduction	139
6.2	The χ^2 distribution	141

Cambridge University Press

052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support

P. C. Gregory

Frontmatter

[More information](#)

viii	<i>Contents</i>	
6.3	Sample variance S^2	143
6.4	The Student's t distribution	147
6.5	F distribution (F -test)	150
6.6	Confidence intervals	152
6.6.1	Variance σ^2 known	152
6.6.2	Confidence intervals for μ , unknown variance	156
6.6.3	Confidence intervals: difference of two means	158
6.6.4	Confidence intervals for σ^2	159
6.6.5	Confidence intervals: ratio of two variances	159
6.7	Summary	160
6.8	Problems	161
7	Frequentist hypothesis testing	162
7.1	Overview	162
7.2	Basic idea	162
7.2.1	Hypothesis testing with the χ^2 statistic	163
7.2.2	Hypothesis test on the difference of two means	167
7.2.3	One-sided and two-sided hypothesis tests	170
7.3	Are two distributions the same?	172
7.3.1	Pearson χ^2 goodness-of-fit test	173
7.3.2	Comparison of two-binned data sets	177
7.4	Problem with frequentist hypothesis testing	177
7.4.1	Bayesian resolution to optional stopping problem	179
7.5	Problems	181
8	Maximum entropy probabilities	184
8.1	Overview	184
8.2	The maximum entropy principle	185
8.3	Shannon's theorem	186
8.4	Alternative justification of MaxEnt	187
8.5	Generalizing MaxEnt	190
8.5.1	Incorporating a prior	190
8.5.2	Continuous probability distributions	191
8.6	How to apply the MaxEnt principle	191
8.6.1	Lagrange multipliers of variational calculus	191
8.7	MaxEnt distributions	192
8.7.1	General properties	192
8.7.2	Uniform distribution	194
8.7.3	Exponential distribution	195
8.7.4	Normal and truncated Gaussian distributions	197
8.7.5	Multivariate Gaussian distribution	202

Cambridge University Press

052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support

P. C. Gregory

Frontmatter

[More information](#)

<i>Contents</i>		ix
8.8	MaxEnt image reconstruction	203
8.8.1	The kangaroo justification	203
8.8.2	MaxEnt for uncertain constraints	206
8.9	Pixon multiresolution image reconstruction	208
8.10	Problems	211
9	Bayesian inference with Gaussian errors	212
9.1	Overview	212
9.2	Bayesian estimate of a mean	212
9.2.1	Mean: known noise σ	213
9.2.2	Mean: known noise, unequal σ	217
9.2.3	Mean: unknown noise σ	218
9.2.4	Bayesian estimate of σ	224
9.3	Is the signal variable?	227
9.4	Comparison of two independent samples	228
9.4.1	Do the samples differ?	230
9.4.2	How do the samples differ?	233
9.4.3	Results	233
9.4.4	The difference in means	236
9.4.5	Ratio of the standard deviations	237
9.4.6	Effect of the prior ranges	239
9.5	Summary	240
9.6	Problems	241
10	Linear model fitting (Gaussian errors)	243
10.1	Overview	243
10.2	Parameter estimation	244
10.2.1	Most probable amplitudes	249
10.2.2	More powerful matrix formulation	253
10.3	Regression analysis	256
10.4	The posterior is a Gaussian	257
10.4.1	Joint credible regions	260
10.5	Model parameter errors	264
10.5.1	Marginalization and the covariance matrix	264
10.5.2	Correlation coefficient	268
10.5.3	More on model parameter errors	272
10.6	Correlated data errors	273
10.7	Model comparison with Gaussian posteriors	275
10.8	Frequentist testing and errors	279
10.8.1	Other model comparison methods	281
10.9	Summary	283
10.10	Problems	284

11 Nonlinear model fitting	287
11.1 Introduction	287
11.2 Asymptotic normal approximation	288
11.3 Laplacian approximations	291
11.3.1 Bayes factor	291
11.3.2 Marginal parameter posteriors	293
11.4 Finding the most probable parameters	294
11.4.1 Simulated annealing	296
11.4.2 Genetic algorithm	297
11.5 Iterative linearization	298
11.5.1 Levenberg–Marquardt method	300
11.5.2 Marquardt’s recipe	301
11.6 <i>Mathematica</i> example	302
11.6.1 Model comparison	304
11.6.2 Marginal and projected distributions	306
11.7 Errors in both coordinates	307
11.8 Summary	309
11.9 Problems	309
12 Markov chain Monte Carlo	312
12.1 Overview	312
12.2 Metropolis–Hastings algorithm	313
12.3 Why does Metropolis–Hastings work?	319
12.4 Simulated tempering	321
12.5 Parallel tempering	321
12.6 Example	322
12.7 Model comparison	326
12.8 Towards an automated MCMC	330
12.9 Extrasolar planet example	331
12.9.1 Model probabilities	335
12.9.2 Results	337
12.10 MCMC robust summary statistic	342
12.11 Summary	346
12.12 Problems	349
13 Bayesian revolution in spectral analysis	352
13.1 Overview	352
13.2 New insights on the periodogram	352
13.2.1 How to compute $p(f D, I)$	356
13.3 Strong prior signal model	358
13.4 No specific prior signal model	360

Cambridge University Press

052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support

P. C. Gregory

Frontmatter

[More information](#)

<i>Contents</i>		xi
13.4.1	X-ray astronomy example	362
13.4.2	Radio astronomy example	363
13.5	Generalized Lomb–Scargle periodogram	365
13.5.1	Relationship to Lomb–Scargle periodogram	367
13.5.2	Example	367
13.6	Non-uniform sampling	370
13.7	Problems	373
14	Bayesian inference with Poisson sampling	376
14.1	Overview	376
14.2	Infer a Poisson rate	377
14.2.1	Summary of posterior	378
14.3	Signal + known background	379
14.4	Analysis of ON/OFF measurements	380
14.4.1	Estimating the source rate	381
14.4.2	Source detection question	384
14.5	Time-varying Poisson rate	386
14.6	Problems	388
Appendix A	Singular value decomposition	389
Appendix B	Discrete Fourier Transforms	392
B.1	Overview	392
B.2	Orthogonal and orthonormal functions	392
B.3	Fourier series and integral transform	394
B.3.1	Fourier series	395
B.3.2	Fourier transform	396
B.4	Convolution and correlation	398
B.4.1	Convolution theorem	399
B.4.2	Correlation theorem	400
B.4.3	Importance of convolution in science	401
B.5	Waveform sampling	403
B.6	Nyquist sampling theorem	404
B.6.1	Astronomy example	406
B.7	Discrete Fourier Transform	407
B.7.1	Graphical development	407
B.7.2	Mathematical development of the DFT	409
B.7.3	Inverse DFT	410
B.8	Applying the DFT	411
B.8.1	DFT as an approximate Fourier transform	411
B.8.2	Inverse discrete Fourier transform	413

Cambridge University Press

052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support

P. C. Gregory

Frontmatter

[More information](#)

xii	<i>Contents</i>	
B.9	The Fast Fourier Transform	415
B.10	Discrete convolution and correlation	417
	B.10.1 Deconvolving a noisy signal	418
	B.10.2 Deconvolution with an optimal Wiener filter	420
	B.10.3 Treatment of end effects by zero padding	421
B.11	Accurate amplitudes by zero padding	422
B.12	Power-spectrum estimation	424
	B.12.1 Parseval's theorem and power spectral density	424
	B.12.2 Periodogram power-spectrum estimation	425
	B.12.3 Correlation spectrum estimation	426
B.13	Discrete power spectral density estimation	428
	B.13.1 Discrete form of Parseval's theorem	428
	B.13.2 One-sided discrete power spectral density	429
	B.13.3 Variance of periodogram estimate	429
	B.13.4 Yule's stochastic spectrum estimation model	431
	B.13.5 Reduction of periodogram variance	431
B.14	Problems	432
Appendix C Difference in two samples		434
C.1	Outline	434
C.2	Probabilities of the four hypotheses	434
	C.2.1 Evaluation of $p(C, S D_1, D_2, I)$	434
	C.2.2 Evaluation of $p(C, \bar{S} D_1, D_2, I)$	436
	C.2.3 Evaluation of $p(\bar{C}, S D_1, D_2, I)$	438
	C.2.4 Evaluation of $p(\bar{C}, \bar{S} D_1, D_2, I)$	439
C.3	The difference in the means	439
	C.3.1 The two-sample problem	440
	C.3.2 The Behrens–Fisher problem	441
C.4	The ratio of the standard deviations	442
	C.4.1 Estimating the ratio, given the means are the same	442
	C.4.2 Estimating the ratio, given the means are different	443
Appendix D Poisson ON/OFF details		445
D.1	Derivation of $p(s N_{\text{on}}, I)$	445
	D.1.1 Evaluation of Num	446
	D.1.2 Evaluation of Den	447
D.2	Derivation of the Bayes factor $B_{\{s+b, b\}}$	448
Appendix E Multivariate Gaussian from maximum entropy		450
	<i>References</i>	455
	<i>Index</i>	461

Cambridge University Press

052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support

P. C. Gregory

Frontmatter

[More information](#)

Preface

The goal of science is to unlock nature's secrets. This involves the identification and understanding of nature's observable structures or patterns. Our understanding comes through the development of theoretical models which are capable of explaining the existing observations as well as making testable predictions. The focus of this book is on what happens at the interface between the predictions of scientific models and the data from the latest experiments. The data are always limited in accuracy and incomplete (we always want more), so we are unable to employ deductive reasoning to prove or disprove the theory. How do we proceed to extend our theoretical framework of understanding in the face of this? Fortunately, a variety of sophisticated mathematical and computational approaches have been developed to help us through this interface, these go under the general heading of statistical inference. Statistical inference provides a means for assessing the plausibility of one or more competing models, and estimating the model parameters and their uncertainties. These topics are commonly referred to as "data analysis" in the jargon of most physicists.

We are currently in the throes of a major paradigm shift in our understanding of statistical inference based on a powerful theory of extended logic. For historical reasons, it is referred to as Bayesian Inference or Bayesian Probability Theory. To get a taste of how significant this development is, consider the following: probabilities are commonly quantified by a real number between 0 and 1. The end-points, corresponding to absolutely false and absolutely true, are simply the extreme limits of this infinity of real numbers. Deductive logic, which is based on axiomatic knowledge, corresponds to these two extremes of 0 and 1. Ask any mathematician or physicist how important deductive logic is to their discipline! Now try to imagine what you might achieve with a theory of extended logic that encompassed the whole range from 0 to 1. This is exactly what is needed in science and real life where we never know anything is absolutely true or false. Of course, the field of probability has been around for years, but what is new is the appreciation that the rules of probability are not merely rules for manipulating random variables. They are now recognized as uniquely valid principles of logic, for conducting inference about any proposition or hypothesis of interest. Ordinary deductive logic is just a special case in the idealized limit of complete information. The reader should be warned that most books on Bayesian statistics

do not make the connection between probability theory and logic. This connection, which is captured in the book by physicist E. T. Jaynes, *Probability Theory – The Logic of Science*,¹ is particularly appealing because of the unifying principles it provides for scientific reasoning.

What are the important consequences of this development? We are only beginning to see the tip of the iceberg. Already we have seen that for data with a high signal-to-noise ratio, a Bayesian analysis can frequently yield many orders of magnitude improvement in model parameter estimation, through the incorporation of relevant prior information about the signal model. For several dramatic demonstrations of this point, have a look at the first four sections of Chapter 13. It also provides a more powerful way of assessing competing theories at the forefront of science by quantifying Occam’s razor, and sheds a new light on systematic errors (e.g., Section 3.11). For some problems, a Bayesian analysis may simply lead to a familiar statistic. Even in this situation it often provides a powerful new insight concerning the interpretation of the statistic. But most importantly, Bayesian analysis provides an elegantly simple and rational approach for answering any scientific question for a given state of information.

This textbook is based on a measurement theory course which is aimed at providing first year graduate students in the physical sciences with the tools to help them design, simulate and analyze experimental data. The material is presented at a mathematical level that should make it accessible to physical science undergraduates in their final two years. Each chapter begins with an overview and most end with a summary. The book contains a large number of problems, worked examples and 132 illustrations.

The Bayesian paradigm is becoming very visible at international meetings of physicists and astronomers (e.g., *Statistical Challenges in Modern Astronomy III*, edited by E. D. Feigelson and G. J. Babu, 2002). However, the majority of scientists are still not at home with the topic and much of the current scientific literature still employs the conventional “frequentist” statistical paradigm. This book is an attempt to help new students to make the transition while at the same time exposing them in Chapters 5, 6, and 7 to some of the essential ideas of the frequentist statistical paradigm that will allow them to comprehend much of the current and earlier literature and interface with his or her research supervisor. This also provides an opportunity to compare and contrast the two different approaches to statistical inference. No previous background in statistics is required; in fact, Chapter 6 is entitled “What is a statistic?” For the reader seeking an abridged version of Bayesian inference, Chapter 3 provides a stand-alone introduction on the “How-to of Bayesian inference.”

¹ Early versions of this much celebrated work by Jaynes have been in circulation since at least 1988. The book was finally submitted for publication in 2002, four years after his death, through the efforts of his former student G. L. Bretthorst. The book is published by Cambridge University Press (Jaynes, 2003, edited by G. L. Bretthorst).

Cambridge University Press

052184150X - Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica™ Support

P. C. Gregory

Frontmatter

[More information](#)*Preface*

xv

The book begins with a look at the role of statistical inference in the scientific method and the fundamental ideas behind Bayesian Probability Theory (BPT). We next consider how to encode a given state of information into the form of a probability distribution, for use as a prior or likelihood function in Bayes' theorem. We demonstrate why the Gaussian distribution arises in nature so frequently from a study of the Central Limit Theorem and gain powerful new insight into the role of the Gaussian distribution in data analysis from the Maximum Entropy Principle. We also learn how a quantified Occam's razor is automatically incorporated into any Bayesian model comparison and come to understand it at a very fundamental level.

Starting from Bayes' theorem, we learn how to obtain unique and optimal solutions to any well-posed inference problem. With this as a foundation, many common analysis techniques such as linear and nonlinear model fitting are developed and their limitations appreciated. The Bayesian solution to a problem is often very simple in principle, however, the calculations require integrals over the model parameter space which can be very time consuming if there are a large number of parameters. Fortunately, the last decade has seen remarkable developments in practical algorithms for performing Bayesian calculations. Chapter 12 provides an introduction to the very powerful Markov chain Monte Carlo (MCMC) algorithms, and demonstrates an application of a new automated MCMC algorithm to the detection of extrasolar planets.

Although the primary emphasis is on the role of probability theory in inference, there is also focus on an understanding of how to simulate the measurement process. This includes learning how to generate pseudo-random numbers with an arbitrary distribution (in Chapter 5). Any linear measurement process can be modeled as a convolution of nature's signal with the measurement point-spread-function, a process most easily dealt with using the convolution theorem of Fourier analysis. Because of the importance of this material, I have included Appendix B on the Discrete Fourier Transform (DFT), the Fast Fourier Transform (FFT), convolution and Weiner filtering. We consider the limitations of the DFT and learn about the need to zero pad in convolution to avoid aliasing. From the Nyquist Sampling Theorem we learn how to minimally sample the signal without losing information and what prefiltering of the signal is required to prevent aliasing.

In Chapter 13, we apply probability theory to spectral analysis problems and gain a new insight into the role of the DFT, and explore a Bayesian revolution in spectral analysis. We also learn that with non-uniform data sampling, the effective bandwidth (the largest spectral window free of aliases) can be made much wider than for uniform sampling. The final chapter is devoted to Bayesian inference when our prior information leads us to model the probability of the data with a Poisson distribution.

Software support

The material in this book is designed to empower the reader in his or her search to unlock nature's secrets. To do this efficiently, one needs both an understanding of the principles of extended logic, and an efficient computing environment for visualizing

and mathematically manipulating the data. All of the course assignments involve the use of a computer. An increasing number of my students are exploiting the power of integrated platforms for programming, symbolic mathematical computations, and visualizing tools. Since the majority of my students opted to use *Mathematica* for their assignments, I adopted *Mathematica* as a default computing environment for the course. There are a number of examples in this book employing *Mathematica* commands, although the book has been designed to be complete without reference to these *Mathematica* examples. In addition, I have developed a *Mathematica* tutorial to support this book, specifically intended to help students and professional scientists with no previous experience with *Mathematica* to efficiently exploit it for data analysis problems. This tutorial also contains many worked examples and is available for download from <http://www.cambridge.org/052184150X>.

In any scientific endeavor, a great deal of effort is expended in graphically displaying the results for presentation and publication. To simplify this aspect of the problem, the *Mathematica* tutorial provides a large range of easy to use templates for publication-quality plotting.

It used to be the case that interpretative languages were not as useful as compiled languages such as C and Fortran for numerically intensive computations. The last few years have seen dramatic improvements in the speed of *Mathematica*. Wolfram Research now claims² that for most of *Mathematica*'s numerical analysis functionality (e.g., data analysis, matrix operations, numerical differential equation solvers, and graphics) *Mathematica* 5 operates on a par³ with Fortran or MATLAB code. In the author's experience, the time required to develop and test programs with *Mathematica* is approximately 20 times shorter than the time required to write and debug the same program in Fortran or C, so the efficiency gain is truly remarkable.

² <http://www.wolfram.com/products/mathematica/newin5/performance/numericallinear.html>.

³ Look up *Mathematica* *gigaNumerics* on the Web.

Acknowledgements

Most of the Bayesian material presented in this book I have learned from the works of Ed Jaynes, Larry Bretthorst, Tom Lored, Steve Gull, John Skilling, Myron Tribus, Devinder Sivia, Jim Berger, and many others from the international community devoted to the study of Bayesian inference. On a personal note, I encountered Bayesian inference one day in 1989 when I found a monograph lying on the floor of the men's washroom entitled *Bayesian Spectrum Analysis and Parameter Estimation*, by Larry Bretthorst. I was so enthralled with the book that I didn't even try to find out whose it was for several weeks. Larry's book led me to the work of his Ph.D. supervisor, Edwin T. Jaynes. I became hooked on this simple, elegant and powerful approach to scientific inference. For me, it was a breath of fresh air providing a logical framework for tackling any statistical inference question in an optimal way, in contrast to the recipe or cookbook approach of conventional statistical analysis.

I would also like to acknowledge the proof reading and suggestions made by many students who were exposed to early versions of this manuscript, in particular, Iva Cheung for her very careful proof reading of the final draft. Finally, I am really grateful to my partner, Jackie, and our children, Rene, Neil, Erin, Melanie, Ted, and Laura, for their encouragement over the many years it took to complete this book.