

Christian P. Robert

# The Bayesian Choice

From Decision-Theoretic Foundations  
to Computational Implementation

Second Edition

 Springer

Christian P. Robert  
CEREMADE  
Universite Paris Dauphine  
Place du Maréchal de Lattre de Tassigny  
75775 Paris cedex 16  
France  
xian@ceremade.dauphine.fr

Library of Congress Control Number: 2007926596

ISBN 978-0-387-71598-8 e-ISBN 978-0-387-71599-5

Printed on acid-free paper.

©2007 Springer Science+Business Media, LLC

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

9 8 7 6 5 4 3 2 1

springer.com

*To my reference prior,  
Brigitte,  
and to two major updates,  
Joachim and Rachel.*

# Preface to the Paperback Edition

What could not be changed must be endured.

**Robert Jordan**, *New Spring, Prequel to The Wheel of Time*

## THANKS

While this paperback edition is almost identical to the second edition of *The Bayesian Choice*, published in 2001, and thus does not require a specific introduction, it offers me the opportunity to thank several groups of people for their contributions that made this edition possible.

First, the changes, when compared with the second edition, are only made of corrections of typographical and conceptual errors (whose updated list can be found on my Webpage<sup>1</sup>). Almost all errors have been pointed out to me by friends, colleagues, unknown lecturers or anonymous readers who (always kindly and sometimes apologetically) sent me emails asking me to clarify a specific paragraph, a formula or a problem that did not make sense to them. Needless to say, I am very grateful to those numerous contributors for making the book more accurate and I obviously encourage all contributors who think there could be an error in the current edition to contact me because they cannot be wrong! Either there indeed is a mistake that needs to be set right or there is no mistake but the context is ambiguous at best and the corresponding text needs to be rewritten. Thanks, then, to Guido Consonni, Estelle Dauchy, Arnaud Doucet, Pierre Druihlet, Ed Green, Feng Liang, Jean-Michel Marin, M.R.L.N. Panchanana, Fabrice Pautot, and Judith Rousseau.

Second, working with my colleague Jean-Michel Marin on the design of a course for teaching Bayesian Statistics from a practical and computational perspective (a venture now published as *Bayesian Core* by Springer in early 2007) was a very important moment in that I realized that the material in this very book, *The Bayesian Choice*, was essential in communicating the essential relevance and coherence of the Bayesian approach

<sup>1</sup> <http://www.ceremade.dauphine.fr/~xian/books.html>

through its decision-theoretic foundations, while the message contained in the other book and transmitted only through processing datasets is that the Bayesian methodology is a universal and multifaceted tool for data analysis. While introducing wider and less mathematical audiences to the elegance and simplicity of the Bayesian methodology in a shorter and therefore more focussed volume was also necessary, if only because some learn better from examples than from theory, I came to the conclusion that there was no paradox in insisting on those foundations in another book! I am therefore immensely thankful to Jean-Michel Marin for initiating this epiphany (if I may rightly borrow this expression from Joyce!), as well as for several years of intense collaboration. Similarly, the DeGroot Prize committee of the ISBA—International Society for Bayesian Analysis—World meeting of 2004 in Valparaiso, Chile, greatly honored me by attributing to *The Bayesian Choice* this prestigious prize. In doing so, this committee highlighted the relevance of both foundations and implementation for the present and future of Bayesian Statistics, when it stated that the “*book sets a new standard for modern textbooks dealing with Bayesian methods, especially those using MCMC techniques, and that it is a worthy successor to DeGroot’s and Berger’s earlier texts*”. I am quite indebted to the members of the committee for this wonderful recognition.

Third, it has been more than 18 years since I started working with John Kimmel from Springer New York (on a basic Probability textbook with Arup Bose that never materialized), and I always appreciated the support he provided over the various editions of the books. So, when he presented me with the possibility to publish this paperback edition, I first got some mixed feelings, because he made me feel like a classics author! This caused my kids poking endless fun at me and, in the end, I am quite grateful to John for the opportunity to teach from this book to a wider audience and thus hopefully exposing them to the beauty of Bayesian theory and methodology. Short of embarking upon a translation of *The Bayesian Choice* into Chinese or Arabic, I do not think there is much more he could do to support the book!

#### IN MEMORIAM

This is a sheer consequence of time moving on, unfortunately, but I lost another dear friend since the last publication of *The Bayesian Choice*. José Sam Lazaro passed away last Spring: a mathematician, a professor and a colleague at the Université de Rouen, a music addict and a movie aficionado that made me discover *Der Tod und das Mädchen* as well as *The Night of The Hunter*, an intense piano player, a memorable tale teller, he was above all a philosopher and a friend. Although he would have made a joke out of it, I would like to dedicate this edition to his memory and wish him well to play this final and endless sonata...

**Valencia and Paris**  
**February 2007**

**Christian P. Robert**

# Preface to the Second Edition

“You can never know everything,” Lan said quietly, “and part of what you know is always wrong. Perhaps even the most important part. A portion of wisdom lies in knowing that. A portion of courage lies in going on anyway.”

**Robert Jordan**, *Winter’s Heart*, Book IX of the *Wheel of Time*.

## OVERVIEW OF CHANGES

Why a second edition? When thinking about it, this is more like a third edition, since the previous edition of *The Bayesian Choice* was the translation of the French version, and already included updates and corrections. The reasons for a new edition of the book are severalfold. The Bayesian community has grown at an incredible pace since 1994. The previous version not only overlooks important areas in the field but misses the significant advances that have taken place in the last seven years.

Firstly, the MCMC<sup>2</sup> revolution has fueled considerable advances in Bayesian modeling, with applications ranging from medical Statistics, to signal processing, to finance. While present in the 1994 edition, these methods were not emphasized enough: for instance, MCMC methods were not presented until the penultimate chapter.

Another significant advance that needed attention is the development of new testing approaches and, more generally, of model choice tools in connection with, and as a result of, MCMC techniques such as reversible jump. Other important expansions include hierarchical and dynamic models, whose processing only began to emerge in the early 1990s.

This second edition is not revolutionary, compared with the 1994 edition. It includes, however, important advances that have taken place since then. The only new chapter deals with model choice (Chapter 7) and is isolated from general testing theory (Chapter 5), since model choice is indeed a different problem and also because it calls for new, mostly computational,

<sup>2</sup> MCMC stands for *Markov chain Monte Carlo*, a simulation methodology which was (re)discovered in the early 1990s by the Bayesian community.

tools. For this reason, and also to emphasize the increasing importance of computational techniques, Chapter 6—previously Chapter 9—has been placed earlier in the book, after the presentation of the fundamentals of Bayesian Statistics. Chapter 6 could almost be called a new chapter in that its presentation has been deeply renovated in the light of ten years of MCMC practice. In Chapter 3, the material on noninformative priors has been expanded and includes, in particular, matching priors, since the research activity has been quite intense in this area in the past few years. Chapter 4 still deals with general estimation problems, but I have incorporated a new section on dynamic models, since those are quite central to the development of Bayesian Statistics in applied fields such as signal processing, finance and econometrics. Despite Delampady's criticisms of Chapter 11 in *The Mathematical Reviews*, I have decided to leave this chapter in: it does not hurt, when one is finished reading a book, to take an overall and more philosophical view of the topic because the reader has very likely acquired enough perspective to understand such arguments. (In a strictly textbook implementation, this chapter can be suggested as an additional reading, comparable with the Notes.)

Another noteworthy change from the previous edition is the decreased emphasis on decision-theoretic principles. Although I still believe that statistical procedures must be grounded on such principles, the developments in the previous decade have mainly focused on methodology, including computational methodology, rather than attacking broader and more ambitious decision problems (once again, including computational methodology). The second part of the book (starting with Chapter 6) is therefore less decision-theoretic and, in contrast to others, chapters such as Chapters 8 and 9 have hardly been changed.

At a more typographical level, subsections and separations have been introduced in many sections to improve visibility and reading, and more advanced or more sketchy parts have been relegated to a *Notes* section at the end of each chapter, following the approach adopted in *Monte Carlo Statistical Methods*, written with George Casella. The end of an example is associated with the || symbol, while the end of a proof is indicated by the □□ symbol.

Several books on Bayesian Statistics have appeared in the interim, among them Bernardo and Smith (1994), Carlin and Louis (1996, 2000a), Gelman et al. (1996), O'Hagan (1994), and Schervish (1995). However, these books either emphasize deeper theoretical aspects at a higher mathematical level (Bernardo and Smith (1994), O'Hagan (1996), or Schervish (1996)) and are thus aimed at a more mature audience than this book, or they highlight a different vision of the practice of Bayesian Statistics (Carlin and Louis (2000a) or Gelman et al. (1996)), missing for instance the connection with Decision Theory developed in this book.

## COURSE SCHEDULES

My advice about running a course based on this book has hardly changed. In a first course on Bayesian analysis, the basic chapters (Chapters 1–6) should be covered almost entirely, with the exception of the Notes and Sections 4.5 and 5.4, while a course focusing more on Decision Theory could skip parts of Chapters 1 and 3, and Chapter 4 altogether, to cover Chapters 7–9. For a more advanced curriculum for students already exposed to Bayesian Statistics, my suggestion is first to cover the impropriety issue in Section 1.5, the noninformative priors in Section 3.5, the dynamic models in Section 4.5 and Notes 4.7.3 and 4.7.4. I would also spend time on the testing issues of Chapter 5 (with the possible exception of Sections 5.3 and 5.4). Then, after a thorough coverage of simulation methods via Chapter 6, I would move to the more controversial topic of model choice in Chapter 7, to recent admissibility results as in Section 8.2.5 and Note 8.7.1, and to the hierarchical and empirical modelings of Chapter 10. In this later scenario, the Notes should be most helpful for setting out reading seminars.

## THANKS

I have always been of two minds about including a thank-you section in a book: on the one hand, it does not mean anything to most readers, except maybe to bring to light some of the author's idiosyncrasies that might better remain hidden! It may also antagonize some of those concerned because they are not mentioned, or because they are not mentioned according to their expectations, or even because they *are* mentioned! On the other hand, the core of ethical requirements for intellectual works is that sources should be acknowledged. This extends to suggestions that contributed to making the work better, clearer or simply different. And it is a small token of gratitude to the following people for the time spent on the successive drafts of this edition that their efforts should be acknowledged in print for all to behold!

Although this is “only” a revision, the time spent on this edition was mostly stolen from evenings, early mornings and week-ends, that is from Brigitte, Joachim and Rachel's time! I am thus most grateful to them for reading and playing (almost) quietly while I was typing furiously and searching desperately through piles of material for this or that reference, and for listening to Bartoli and Guðjónsson, rather than to Manau or Diana Krall! I cannot swear this book-writing experience will never happen again but, in the meanwhile, I promise there will be more time available for reading *Mister Bear to the Rescue*, and for besieging the Playmobil castle in full scale, for playing chess and for biking on Sunday afternoons!

I am thankful to several people for the improvements in the current edition! First, I got a steady stream of feedback and suggestions from those who taught from the book. This group includes Ed Green, Tatsuya Kubokawa, and Marty Wells. In particular, Judith Rousseau, radical biker

and Jordanite as well as Bayesian, definitely was instrumental in the re-organization of Chapter 3. I also got many helpful comments from many people, including the two “Cambridge Frenchies” Christophe Andrieu and Arnaud Doucet (plus a memorable welcome for a retreat week in Cambridge to finish Chapter 6), Jim Berger (for his support in general, and for providing preprints on model choice in particular), Olivier Cappé (who also installed Linux on my laptop, and consequently brought immense freedom for working on the book anywhere, from the sand-box to the subway, and, lately, to CREST, where Unix is now banned!), Maria DeIorio, Jean-Louis Fouley, Malay Ghosh (through his very supportive review in JASA), Jim Hobert (who helped in clarifying Chapters 6 and 10), Ana Justel, Stephen Lauritzen (for pointing out mistakes with Wishart distributions), Anne Philippe, Walter Racugno (who gave me the opportunity to teach an advanced class in model choice in Ca’liari last fall, thus providing the core of Chapter 7), Adrian Raftery, Anne Sullivan Rosen (about the style of this preface), and Jean-Michel Zakoian (for his advice on the new parts on dynamic models). I also take the opportunity to thank other friends and colleagues such as George Casella, Jérôme Dupuis, Merrilee Hurn, Kerrie Mengersen, Eric Moulines, Alain Monfort, and Mike Titterington, since working with them gave me a broader vision of the field, which is hopefully incorporated in this version. In particular, the experience of writing *Monte Carlo Statistical Methods* with George Casella in the past years left its mark on this book, not only through the style file and the inclusion of Notes, but also as a sharper focus on essentials. Manuela Delbois helped very obligingly with the transformation from T<sub>E</sub>X to L<sup>A</sup>T<sub>E</sub>X, and with the subsequent additions and indexings. And, last but not least, John Kimmel and Jenny Wolkowicki, from Springer-Verlag, have been very efficient and helpful in pushing me to write this new edition for the former, in keeping the whole schedule under control and in getting the book published on time for the latter. Needless to say, the usual *proviso* applies: all remaining typos, errors, confusions and obscure statements are mine and only mine!

#### IN MEMORIAM

A most personal word about two people whose *absence* has marked this new edition: in the summer 1997, I lost my friend Costas Goutis in a diving accident in Seattle. By no means am I the only one to feel keenly his absence, but, beyond any doubt, this book would have benefited from his vision, had he been around. Two summers later, in 1999, Bernhard K. Flury died in a mountain accident in the Alps. While the criticisms of our respective books always focussed on the cover colors, even to the extent of sending one another pirated versions in the “right” color, the disappearance of his unique humor has taken a measure of fun out of the world.

Paris, France  
March 2001

Christian P. Robert

# Preface to the First Edition

From where we stand, the rain seems random.  
If we could stand somewhere else, we would see the order in it.

— **T. Hillerman** (1990) *Coyote Waits*.

This book stemmed from a translation of a French version that was written to supplement the gap in the French statistical literature about Bayesian Analysis and Decision Theory. As a result, its scope is wide enough to cover the two years of the French graduate Statistics curriculum and, more generally, most graduate programs. This book builds on very little prerequisites in Statistics and only requires basic skills in calculus, measure theory, and probability. Intended as a preparation of doctoral candidates, this book goes far enough to cover advanced topics and modern developments of Bayesian Statistics (complete class theorems, the Stein effect, hierarchical and empirical modelings, Gibbs sampling, etc.). As usual, what started as a translation eventually ended up as a deeper revision because of the comments of French readers, adjustments to the different needs of American programs, and because my perception of things has changed slightly in the meantime. As a result, this new version is quite adequate for a general graduate audience of an American university.

In terms of level and existing literature, this book starts at a level similar to those of the introductory books of Lee (1989) and Press (1989), but it also goes further and keeps up with most of the recent advances in Bayesian Statistics, while justifying the theoretical appeal of the Bayesian approach on decision-theoretic grounds. Nonetheless, this book differs from the reference book of Berger (1985a) by including the more recent developments of the Bayesian field (the Stein effect for spherically symmetric distributions, multiple shrinkage, loss estimation, decision theory for testing and confidence regions, hierarchical developments, Bayesian computation, mixture estimation, etc.). Moreover, the style is closer to that of a textbook in the sense that the progression is intended to be linear. In fact, the exposition of

the advantages of a Bayesian approach and of the existing links with other axiomatic systems (fiducial theory, maximum likelihood, frequentist theory, invariance, etc.) does not prevent an overall unity in the discourse. This should make the book easier to read by students; through the years and on both sides of the blackboard(!), I found most Statistics courses disturbing because a wide scope of methods was presented simultaneously with very little emphasis on ways of discriminating between competing approaches. In particular, students with a strong mathematical background are quite puzzled by this multiplicity of theories since they have not been exposed previously to conflicting systems of axioms. A unitarian presentation that includes other approaches as limiting cases is thus more likely to reassure the students, while giving a broad enough view of Decision Theory and even of parametric Statistics.

The plan<sup>3</sup> of the book is as follows: Chapter 1 is an introduction to statistical models, including the Bayesian model and some connections with the Likelihood Principle. The book then proceeds with Chapter 2 on Decision Theory, considered from a classical point of view, this approach being justified through the axioms of rationality and the need to compare decision rules in a coherent way. It also includes a presentation of usual losses and a discussion of the Stein effect. Chapter 3 gives the corresponding analysis for prior distributions and deals in detail with conjugate priors, mixtures of conjugate priors, and noninformative priors, including a concluding section on prior robustness. Classical statistical models are studied in Chapter 4, paying particular attention to normal models and their relations with linear regression. This chapter also contains a section on sampling models that allows us to include the pedagogical example of capture-recapture models. Tests and confidence regions are considered separately in Chapter 5, since we present the usual construction through 0 – 1 losses, but also include recent advances in the alternative decision-theoretic evaluations of testing problems. The second part of the book dwells on more advanced topics and can be considered as providing a basis for a more advanced graduate course. Chapter 8 covers complete class results and sufficient/necessary admissibility conditions. Chapter 9 introduces the notion of invariance and its relations with Bayesian Statistics, including a heuristic section on the Hunt–Stein theorem. Hierarchical and empirical extensions of the Bayesian approach, including some developments on the Stein effect, are treated in Chapter 10. Chapter 6 is quite appealing, considering the available literature, as it incorporates in a graduate textbook an introduction to state-of-the-art computational methods (Laplace, Monte Carlo and, mainly, Gibbs sampling). In connection with this chapter, a short appendix provides the usual pseudo-random generators. Chapter 11 is a more personal conclusion on the advantages of Bayesian theory, also mentioning the most common criticisms of the Bayesian approach. French readers may appreciate that

<sup>3</sup> The chapter and section numbers have been adapted to the current edition.

a lot of effort has been put into the exercises of each chapter in terms of volume and difficulty. They now range from very easy to difficult, instead of being uniformly difficult! The most difficult exercises are indexed by asterisks and are usually derived from research papers (covering subjects such as *spherically symmetric distributions* (1.1), *the Pitman nearness criticism* (2.57–2.62), *marginalization paradoxes* (3.44–3.50), *multiple shrinkage* (10.38), etc.). They should benefit most readers by pointing out new directions of Bayesian research and providing additional perspectives.

A standard one-semester course should cover the first five chapters (with the possible omission of Note 2.8.2, §2.5.4, §2.6, §3.4, Note 4.7.1, §4.3.3, and §5.4). More advanced (or longer) courses can explore the material presented in Chapters 8, 9, and 10, bearing in mind that a detailed and rigorous treatment of these topics requires additional reading of the literature mentioned in those chapters. In any case, I would advise against entirely forgoing Chapter 6. Even a cursory reading of this chapter may be beneficial to most students, by illustrating the practical difficulties related to the computation of Bayesian procedures and the corresponding answers brought by simulation methods.

This book took many excruciatingly small steps and exacted a heavy toll on evenings, weekends, and vacations. . . It is thus only a small indication of my gratitude that this book be dedicated to Brigitte (although she might take this as a propitiatory attempt for future books!!!). Many persons are to be thanked for the present version of this book. First and foremost, Jim Berger’s “responsibility” can be traced back to 1987 when he invited me to Purdue University for a year and, as a result, considerably broadened my vision of Statistics; he emphasized his case by insisting very vigorously that I translate the French version and urging me along the whole time. My gratitude to Jim goes very deep when I consider his strong influence in my “coming-of-age” as a statistician. Mary-Ellen Bock, Anirban Das Gupta, Edward George, Gene (formerly Jiunn) Hwang, and Marty Wells were also very instrumental in my progression towards the Bayesian choice, although they do not necessarily support this choice. In this regard, George Casella must be singled out for his strong influence through these years of intense collaboration and friendship, even during his most severe (and “unbearable”) criticisms of the Bayesian paradigm! I am also quite grateful to Jean-François Angers, Dean Foster, and Giovanni Parmigiani for taking the risk of using a preliminary version of these notes in their courses, as well as for their subsequent comments. Thanks to Teena Seele for guiding my first steps in T<sub>E</sub>X, as well as some delicate points in this book—never use `\def` as an abbreviation of **definition**! I am also grateful to Elsevier North-Holland for granting me permission to use Diaconis and Ylvisaker’s (1985) figures in §3.3. Last, and definitely not least, Kerrie Mengersen and Costas Goutis put a lot of time and effort reading through a preliminary version and provided many helpful comments on content, style, and clarity, while adding a

touch of Ausso-Greek accent to the tone. (In addition, Costas Goutis saved the subject index from utter destruction!) They are thus partly responsible for the improvements over previous versions (but obviously not for the remaining defects!), and I am most grateful to them for their essential help.

Paris, France  
May 1994

Christian P. Robert

# Contents

<b>Preface to the Paperback Edition</b>	<b>vii</b>
<b>Preface to the Second Edition</b>	<b>ix</b>
<b>Preface to the First Edition</b>	<b>xiii</b>
<b>List of Tables</b>	<b>xxiii</b>
<b>List of Figures</b>	<b>xxv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Statistical problems and statistical models	1
1.2 The Bayesian paradigm as a duality principle	8
1.3 Likelihood Principle and Sufficiency Principle	13
1.3.1 Sufficiency	13
1.3.2 The Likelihood Principle	15
1.3.3 Derivation of the Likelihood Principle	18
1.3.4 Implementation of the Likelihood Principle	19
1.3.5 Maximum likelihood estimation	20
1.4 Prior and posterior distributions	22
1.5 Improper prior distributions	26
1.6 The Bayesian choice	31
1.7 Exercises	31
1.8 Notes	45
<b>2 Decision-Theoretic Foundations</b>	<b>51</b>
2.1 Evaluating estimators	51
2.2 Existence of a utility function	54
2.3 Utility and loss	60
2.4 Two optimalities: minimaxity and admissibility	65
2.4.1 Randomized estimators	65
2.4.2 Minimaxity	66
2.4.3 Existence of minimax rules and maximin strategy	69
2.4.4 Admissibility	74
2.5 Usual loss functions	77
2.5.1 The quadratic loss	77
2.5.2 The absolute error loss	79

2.5.3	The 0 – 1 loss	80
2.5.4	Intrinsic losses	81
2.6	Criticisms and alternatives	83
2.7	Exercises	85
2.8	Notes	96
<b>3</b>	<b>From Prior Information to Prior Distributions</b>	<b>105</b>
3.1	The difficulty in selecting a prior distribution	105
3.2	Subjective determination and approximations	106
3.2.1	Existence	106
3.2.2	Approximations to the prior distribution	108
3.2.3	Maximum entropy priors	109
3.2.4	Parametric approximations	111
3.2.5	Other techniques	113
3.3	Conjugate priors	113
3.3.1	Introduction	113
3.3.2	Justifications	114
3.3.3	Exponential families	115
3.3.4	Conjugate distributions for exponential families	120
3.4	Criticisms and extensions	123
3.5	Noninformative prior distributions	127
3.5.1	Laplace’s prior	127
3.5.2	Invariant priors	128
3.5.3	The Jeffreys prior	129
3.5.4	Reference priors	133
3.5.5	Matching priors	137
3.5.6	Other approaches	140
3.6	Posterior validation and robustness	141
3.7	Exercises	144
3.8	Notes	158
<b>4</b>	<b>Bayesian Point Estimation</b>	<b>165</b>
4.1	Bayesian inference	165
4.1.1	Introduction	165
4.1.2	MAP estimator	166
4.1.3	Likelihood Principle	167
4.1.4	Restricted parameter space	168
4.1.5	Precision of the Bayes estimators	170
4.1.6	Prediction	171
4.1.7	Back to Decision Theory	173
4.2	Bayesian Decision Theory	173
4.2.1	Bayes estimators	173
4.2.2	Conjugate priors	175
4.2.3	Loss estimation	178
4.3	Sampling models	180
4.3.1	Laplace succession rule	180

4.3.2	The tramcar problem	181
4.3.3	Capture-recapture models	182
4.4	The particular case of the normal model	186
4.4.1	Introduction	186
4.4.2	Estimation of variance	187
4.4.3	Linear models and G-priors	190
4.5	Dynamic models	193
4.5.1	Introduction	193
4.5.2	The AR model	196
4.5.3	The MA model	198
4.5.4	The ARMA model	201
4.6	Exercises	201
4.7	Notes	216
<b>5</b>	<b>Tests and Confidence Regions</b>	<b>223</b>
5.1	Introduction	223
5.2	A first approach to testing theory	224
5.2.1	Decision-theoretic testing	224
5.2.2	The Bayes factor	227
5.2.3	Modification of the prior	229
5.2.4	Point-null hypotheses	230
5.2.5	Improper priors	232
5.2.6	Pseudo-Bayes factors	236
5.3	Comparisons with the classical approach	242
5.3.1	UMP and UMPU tests	242
5.3.2	Least favorable prior distributions	245
5.3.3	Criticisms	247
5.3.4	The $p$ -values	249
5.3.5	Least favorable Bayesian answers	250
5.3.6	The one-sided case	254
5.4	A second decision-theoretic approach	256
5.5	Confidence regions	259
5.5.1	Credible intervals	260
5.5.2	Classical confidence intervals	263
5.5.3	Decision-theoretic evaluation of confidence sets	264
5.6	Exercises	267
5.7	Notes	279
<b>6</b>	<b>Bayesian Calculations</b>	<b>285</b>
6.1	Implementation difficulties	285
6.2	Classical approximation methods	293
6.2.1	Numerical integration	293
6.2.2	Monte Carlo methods	294
6.2.3	Laplace analytic approximation	298
6.3	Markov chain Monte Carlo methods	301
6.3.1	MCMC in practice	302

6.3.2	Metropolis–Hastings algorithms	303
6.3.3	The Gibbs sampler	307
6.3.4	Rao–Blackwellization	309
6.3.5	The general Gibbs sampler	311
6.3.6	The slice sampler	315
6.3.7	The impact on Bayesian Statistics	317
6.4	An application to mixture estimation	318
6.5	Exercises	321
6.6	Notes	334
<b>7</b>	<b>Model Choice</b>	<b>343</b>
7.1	Introduction	343
7.1.1	Choice between models	344
7.1.2	Model choice: motives and uses	347
7.2	Standard framework	348
7.2.1	Prior modeling for model choice	348
7.2.2	Bayes factors	350
7.2.3	Schwartz’s criterion	352
7.2.4	Bayesian deviance	354
7.3	Monte Carlo and MCMC approximations	356
7.3.1	Importance sampling	356
7.3.2	Bridge sampling	358
7.3.3	MCMC methods	359
7.3.4	Reversible jump MCMC	363
7.4	Model averaging	366
7.5	Model projections	369
7.6	Goodness-of-fit	374
7.7	Exercises	377
7.8	Notes	386
<b>8</b>	<b>Admissibility and Complete Classes</b>	<b>391</b>
8.1	Introduction	391
8.2	Admissibility of Bayes estimators	391
8.2.1	General characterizations	391
8.2.2	Boundary conditions	393
8.2.3	Inadmissible generalized Bayes estimators	395
8.2.4	Differential representations	396
8.2.5	Recurrence conditions	398
8.3	Necessary and sufficient admissibility conditions	400
8.3.1	Continuous risks	401
8.3.2	Blyth’s sufficient condition	402
8.3.3	Stein’s necessary and sufficient condition	407
8.3.4	Another limit theorem	407
8.4	Complete classes	409
8.5	Necessary admissibility conditions	412
8.6	Exercises	416

8.7	Notes	425
<b>9</b>	<b>Invariance, Haar Measures, and Equivariant Estimators</b>	<b>427</b>
9.1	Invariance principles	427
9.2	The particular case of location parameters	429
9.3	Invariant decision problems	431
9.4	Best equivariant noninformative distributions	436
9.5	The Hunt–Stein theorem	441
9.6	The role of invariance in Bayesian Statistics	445
9.7	Exercises	446
9.8	Notes	454
<b>10</b>	<b>Hierarchical and Empirical Bayes Extensions</b>	<b>457</b>
10.1	Incompletely Specified Priors	457
10.2	Hierarchical Bayes analysis	460
10.2.1	Hierarchical models	460
10.2.2	Justifications	462
10.2.3	Conditional decompositions	465
10.2.4	Computational issues	468
10.2.5	Hierarchical extensions for the normal model	470
10.3	Optimality of hierarchical Bayes estimators	474
10.4	The empirical Bayes alternative	478
10.4.1	Nonparametric empirical Bayes	479
10.4.2	Parametric empirical Bayes	481
10.5	Empirical Bayes justifications of the Stein effect	484
10.5.1	Point estimation	485
10.5.2	Variance evaluation	487
10.5.3	Confidence regions	488
10.5.4	Comments	490
10.6	Exercises	490
10.7	Notes	502
<b>11</b>	<b>A Defense of the Bayesian Choice</b>	<b>507</b>
<b>A</b>	<b>Probability Distributions</b>	<b>519</b>
A.1	Normal distribution, $\mathcal{N}_p(\theta, \Sigma)$	519
A.2	Gamma distribution, $\mathcal{G}(\alpha, \beta)$	519
A.3	Beta distribution, $\mathcal{Be}(\alpha, \beta)$	519
A.4	Student’s $t$ -distribution, $\mathcal{T}_p(\nu, \theta, \Sigma)$	520
A.5	Fisher’s $F$ -distribution, $\mathcal{F}(\nu, \varrho)$	520
A.6	Inverse gamma distribution, $\mathcal{IG}(\alpha, \beta)$	520
A.7	Noncentral chi-squared distribution, $\chi_\nu^2(\lambda)$	520
A.8	Dirichlet distribution, $\mathcal{D}_k(\alpha_1, \dots, \alpha_k)$	521
A.9	Pareto distribution, $\mathcal{Pa}(\alpha, x_0)$	521
A.10	Binomial distribution, $\mathcal{B}(n, p)$ .	521
A.11	Multinomial distribution, $\mathcal{M}_k(n; p_1, \dots, p_k)$	521

A.12 Poisson distribution, $\mathcal{P}(\lambda)$	521
A.13 Negative Binomial distribution, $\mathcal{N}eg(n, p)$	522
A.14 Hypergeometric distribution, $\mathcal{H}yp(N; n; p)$	522
<b>B Usual Pseudo-random Generators</b>	<b>523</b>
B.1 Normal distribution, $\mathcal{N}(0, 1)$	523
B.2 Exponential distribution, $\mathcal{E}xp(\lambda)$	523
B.3 Student's $t$ -distribution, $\mathcal{T}(\nu, 0, 1)$	524
B.4 Gamma distribution, $\mathcal{G}(\alpha, 1)$	524
B.5 Binomial distribution, $\mathcal{B}(n, p)$	525
B.6 Poisson distribution, $\mathcal{P}(\lambda)$	525
<b>C Notations</b>	<b>527</b>
C.1 Mathematical	527
C.2 Probabilistic	528
C.3 Distributional	528
C.4 Decisional	529
C.5 Statistical	529
C.6 Markov chains	530
<b>References</b>	<b>531</b>
<b>Author Index</b>	<b>579</b>
<b>Subject Index</b>	<b>587</b>

# List of Tables

2.4.1 Utility function	69
3.2.1 Capture and survival information	107
3.2.2 Capture and survival priors	108
3.2.3 Ranges of posterior moments	112
3.3.1 Natural conjugate priors	121
3.5.1 Matching reference priors	140
3.8.1 Approximations by conjugate mixtures	161
4.2.1 Bayes estimators for exponential families	176
4.3.1 Probabilities of capture	182
4.3.2 Population capture division	183
4.3.3 Posterior deer distribution	185
4.3.4 Posterior mean deer population	185
4.3.5 Estimated deer population	186
5.2.1 Posterior probabilities of $p = 1/2$	232
5.2.2 Posterior probabilities of $\theta = 0$	232
5.2.3 Posterior probabilities of $\theta = 0$	233
5.2.4 Posterior probabilities of $ \theta  < 1$ .	233
5.2.5 Posterior probabilities of $\theta = 0$	234
5.2.6 Posterior probabilities of $\theta = 0$	235
5.3.1 Comparison between $p$ -values and Bayesian answers	252
5.3.2 Comparison between $p$ -values and Bayesian answers	253
5.3.3 Bayes factors and posterior probabilities	254
5.3.4 Comparison between $p$ -values and Bayesian posterior probabilities	255
5.5.1 Confidence intervals for the binomial distribution.	261
6.1.1 Parameters of a lung radiograph mode	290
6.5.1 Frequencies of car passages	333
7.1.1 Orange tree circumferences	346
7.3.1 Adequacy for the four orange models	361
7.5.1 Parameters for Kullback–Leibler divergences	371
7.5.2 Submodels for the breast-cancer dataset	373
7.7.1 Number of women in a queue	385