

国外实用统计丛书

贝叶斯统计方法

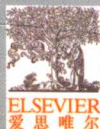
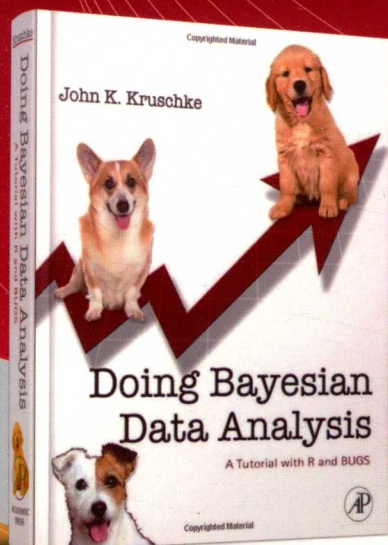
——R和BUGS软件数据分析示例

(影印版)

Doing Bayesian Data Analysis

A Tutorial with R and BUGS

[美] 约翰 K. 克鲁斯克 (John K. Kruschke) 著



 **机械工业出版社**
CHINA MACHINE PRESS

贝叶斯统计方法

——R和BUGS软件数据分析示例

(影印版)

[美] 约翰 K. 克鲁斯克 (John K. Kruschke) 著



机械工业出版社

图书在版编目 (CIP) 数据

贝叶斯统计方法: R和BUGS软件数据分析示例: 英文/(美)克鲁斯克
(Kruschke, J. K.) 著. — 影印本. — 北京: 机械工业出版社, 2015.5

书名原文: Doing Bayesian Data Analysis: A Tutorial with R and BUGS
ISBN 978-7-111-50446-7

I. ①贝… II. ①克… III. ①贝叶斯方法—统计方法—英文 IV. ①0212.8

中国版本图书馆CIP数据核字 (2015) 第123599号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 汤 嘉 责任编辑: 汤 嘉

封面设计: 张 静 责任印制: 乔 宇

保定市中国画美凯印刷有限公司印刷

2015 年 7 月第 1 版第 1 次印刷

184mm × 240mm · 45.5 印张 · 1 插页 · 844 千字

标准书号: ISBN 978-7-111-50446-7

定价: 98.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

电话服务

网络服务

服务咨询热线: (010) 88379833 机工官网: www.cmpbook.com

读者购书热线: (010) 88379649 机工官博: weibo.com/cmp1952

教育服务网: www.cmpedu.com

封面无防伪标均为盗版

金 书 网: www.golden-book.com

Doing Bayesian Data Analysis

John K. Kruschke

Department of Psychology & Neuroscience Center

University of Tübingen

Tübingen, Germany

Doing Bayesian Data Analysis

A Tutorial with R and BUGS

John K. Kruschke

*Department of Psychological & Brain Sciences
Indiana University
Bloomington, IN*

Dedicated to my mother, Marilyn A. Kruschke,
and to the memory of my father, Earl R. Kruschke,
who both brilliantly exemplified and taught sound reasoning.

And, in honor of my father,
who dedicated his first book to his children,

I also dedicate this book to mine:

Claire A. Kruschke and Loren D. Kruschke

目 录

第 1 章	关于本书	1
	1.1 目标读者	1
	1.2 预备知识	2
	1.3 本书结构	3
	1.3.1 重点章节	4
	1.3.2 与贝叶斯方法对应的传统检验方法	4
	1.4 期待反馈	6
	1.5 致谢	6
第 1 部分	基础篇：关于参数、概率、贝叶斯法则及 R 软件	
第 2 章	我们所信的模式	9
	2.1 观测模型与信念模型	10
	2.1.1 先验信念与后验信念	12
	2.2 统计推断的三个目标	12
	2.2.1 参数估计	13
	2.2.2 数值预测	13
	2.2.3 模型比较	13
	2.3 R 编程基础	14
	2.3.1 软件的获取和安装	15
	2.3.2 激活 R 和命令行使用	15
	2.3.3 应用实例	16
	2.3.4 获取帮助	17
	2.3.5 编程	18
	2.4 练习	20

Contents

CHAPTER 1	This Book's Organization: <i>Read Me First!</i>	1
1.1	Real People Can Read This Book.....	1
1.2	Prerequisites	2
1.3	The Organization of This Book	3
1.3.1	What Are the Essential Chapters?	4
1.3.2	Where's the Equivalent of Traditional Test X in This Book?	4
1.4	Gimme Feedback (Be Polite)	6
1.5	Acknowledgments	6
Part 1	The Basics: <i>Parameters, Probability, Bayes' Rule, and R</i>	
CHAPTER 2	Introduction: <i>Models We Believe In</i>	9
2.1	Models of Observations and Models of Beliefs.....	10
2.1.1	Prior and Posterior Beliefs	12
2.2	Three Goals for Inference from Data.....	12
2.2.1	Estimation of Parameter Values	13
2.2.2	Prediction of Data Values	13
2.2.3	Model Comparison.....	13
2.3	The R Programming Language	14
2.3.1	Getting and Installing R	15
2.3.2	Invoking R and Using the Command Line	15
2.3.3	A Simple Example of R in Action	16
2.3.4	Getting Help in R.....	17
2.3.5	Programming in R.....	18
2.4	Exercises	20

第 3 章	概率究竟是什么?	23
	3.1 所有可能事件的集合	24
	3.1.1 抛硬币实验	25
	3.2 概率: 意识之外	25
	3.2.1 意识之外: 长期相对频率	25
	3.2.2 意识以内: 主观信念	28
	3.2.3 概率: 量化可能性	29
	3.3 概率分布	30
	3.3.1 离散分布: 概率质量	30
	3.3.2 连续分布: 密度初探	31
	3.3.3 分布的均值与方差	37
	3.3.4 反映信念不确定性的方差	39
	3.3.5 最高密度区间 (HDI)	40
	3.4 双变量联合分布	42
	3.4.1 边际概率	43
	3.4.2 条件概率	44
	3.4.3 独立事件	46
	3.5 R 代码	47
	3.5.1 图 3.1 的 R 代码	47
	3.5.2 图 3.3 的 R 代码	48
	3.6 练习	49
第 4 章	贝叶斯公式	51
	4.1 贝叶斯公式简介	52
	4.1.1 从条件概率的定义导出	53
	4.1.2 受双因素表的启发	54
	4.1.3 连续情形下的积分表达	56
	4.2 在模型和数据中的应用	56
	4.2.1 数据的顺序不变性	59
	4.2.2 一个例子: 抛硬币	60
	4.3 推断的三个目标	63
	4.3.1 参数估计	63
	4.3.2 数值预测	63
	4.3.3 模型比较	64
	4.3.4 为什么贝叶斯推断是困难的	67
	4.3.5 贝叶斯推断在日常生活中的应用	68

CHAPTER 3	What Is This Stuff Called Probability?	23
3.1	The Set of All Possible Events	24
3.1.1	Coin Flips: Why You Should Care	25
3.2	Probability: Outside or Inside the Head	25
3.2.1	Outside the Head: Long-Run Relative Frequency	25
3.2.2	Inside the Head: Subjective Belief	28
3.2.3	Probabilities Assign Numbers to Possibilities	29
3.3	Probability Distributions	30
3.3.1	Discrete Distributions: Probability Mass	30
3.3.2	Continuous Distributions: Rendezvous with Density	31
3.3.3	Mean and Variance of a Distribution	37
3.3.4	Variance as Uncertainty in Beliefs	39
3.3.5	Highest Density Interval (HDI)	40
3.4	Two-Way Distributions	42
3.4.1	Marginal Probability	43
3.4.2	Conditional Probability	44
3.4.3	Independence of Attributes	46
3.5	R Code	47
3.5.1	R Code for Figure 3.1	47
3.5.2	R Code for Figure 3.3	48
3.6	Exercises	49
CHAPTER 4	Bayes' Rule	51
4.1	Bayes' Rule	52
4.1.1	Derived from Definitions of Conditional Probability	53
4.1.2	Intuited from a Two-Way Discrete Table	54
4.1.3	The Denominator as an Integral over Continuous Values	56
4.2	Applied to Models and Data	56
4.2.1	Data Order Invariance	59
4.2.2	An Example with Coin Flipping	60
4.3	The Three Goals of Inference	63
4.3.1	Estimation of Parameter Values	63
4.3.2	Prediction of Data Values	63
4.3.3	Model Comparison	64
4.3.4	Why Bayesian Inference Can Be Difficult	67
4.3.5	Bayesian Reasoning in Everyday Life	68

	4.4 R 代码.....	69
	4.4.1 图 4.1 的 R 代码.....	69
	4.5 练习.....	71
第 2 部分	用于二元比例推断的基本理论	
第 5 章	二元比例推断的精确数学分析方法.....	77
	5.1 伯努利分布的似然函数.....	78
	5.2 贝塔分布简介.....	80
	5.2.1 先验贝塔分布.....	81
	5.2.2 后验贝塔分布.....	84
	5.3 推断的三个目标.....	85
	5.3.1 二元比例的估计.....	85
	5.3.2 预测数据.....	87
	5.3.3 模型比较.....	88
	5.4 总结：如何做贝叶斯推断.....	90
	5.5 R 代码.....	91
	5.5.1 图 5.2 的 R 代码.....	91
	5.6 练习.....	95
第 6 章	二元比例推断的格点估计法.....	101
	6.1 θ 取值离散时的贝叶斯准则.....	102
	6.2 连续先验密度的离散化.....	102
	6.2.1 离散化先验密度的例子.....	104
	6.3 估计.....	106
	6.4 序贯数据的预测.....	107
	6.5 模型比较.....	108
	6.6 总结.....	109
	6.7 R 代码.....	109
	6.7.1 图 6.2 及类似图形的 R 代码.....	109
	6.8 练习.....	112
第 7 章	二元比例推断的 Metropolis 算法.....	117
	7.1 Metropolis 算法的简单例子.....	119
	7.1.1 政治家巧遇 Metropolis 算法.....	119
	7.1.2 随机游走.....	120

4.4	R Code	69
4.4.1	R Code for Figure 4.1	69
4.5	Exercises	71

Part 2 **All the Fundamentals Applied to Inferring a Binomial Proportion**

CHAPTER 5	Inferring a Binomial Proportion via Exact Mathematical Analysis	77
5.1	The Likelihood Function: Bernoulli Distribution.....	78
5.2	A Description of Beliefs: The Beta Distribution.....	80
5.2.1	Specifying a Beta Prior.....	81
5.2.2	The Posterior Beta	84
5.3	Three Inferential Goals.....	85
5.3.1	Estimating the Binomial Proportion	85
5.3.2	Predicting Data	87
5.3.3	Model Comparison.....	88
5.4	Summary: How to Do Bayesian Inference	90
5.5	R Code	91
5.5.1	R Code for Figure 5.2.....	91
5.6	Exercises	95
CHAPTER 6	Inferring a Binomial Proportion via Grid Approximation	101
6.1	Bayes' Rule for Discrete Values of θ	102
6.2	Discretizing a Continuous Prior Density.....	102
6.2.1	Examples Using Discretized Priors	104
6.3	Estimation	106
6.4	Prediction of Subsequent Data.....	107
6.5	Model Comparison.....	108
6.6	Summary	109
6.7	R Code	109
6.7.1	R Code for Figure 6.2 and the Like.....	109
6.8	Exercises	112
CHAPTER 7	Inferring a Binomial Proportion via the Metropolis Algorithm	117
7.1	A Simple Case of the Metropolis Algorithm.....	119
7.1.1	A Politician Stumbles on the Metropolis Algorithm	119
7.1.2	A Random Walk	120

7.1.3	随机游走的性质	122
7.1.4	为什么关注随机游走	125
7.1.5	Metropolis 算法是如何起作用的	126
7.2	Metropolis 算法的详细介绍	129
7.2.1	预烧、效率和收敛	131
7.2.2	术语: 马尔可夫链 - 蒙特卡罗方法	132
7.3	从抽样后验分布到推断的三个目标	132
7.3.1	估计	134
7.3.2	预测	136
7.3.3	模型比较: $p(D)$ 的估计	137
7.4	BUGS 的 MCMC	139
7.4.1	用 BUGS 估计参数	140
7.4.2	用 BUGS 预测	143
7.4.3	用 BUGS 进行模型比较	145
7.5	结论	145
7.6	R 代码	146
7.6.1	作者编写的 Metropolis 算法的 R 代码	146
7.7	练习	150
第 8 章	使用 Gibbs 抽样推断两个二元比例	155
8.1	两个比例的先验、似然和后验	157
8.2	后验分布的常规分析	159
8.3	使用格点估计近似后验分布	163
8.4	使用 MCMC 推断后验分布	165
8.4.1	Metropolis 算法	165
8.4.2	Gibbs 抽样	167
8.5	BUGS 实现	173
8.5.1	在 BUGS 中抽样获取先验分布	174
8.6	潜在偏差有何差异?	175
8.7	总结	177
8.8	R 代码	178
8.8.1	格点估计的 R 代码 (图 8.1 和图 8.2)	178
8.8.2	Metropolis 抽样的 R 代码 (图 8.3)	181
8.8.3	BUGS 抽样的 R 代码 (图 8.6)	184
8.8.4	画后验直方图的 R 代码	186
8.9	练习	188

7.1.3	General Properties of a Random Walk	122
7.1.4	Why We Care.....	125
7.1.5	Why It Works.....	126
7.2	The Metropolis Algorithm More Generally	129
7.2.1	“Burn-in,” Efficiency, and Convergence	131
7.2.2	Terminology: Markov Chain Monte Carlo.....	132
7.3	From the Sampled Posterior to the Three Goals	132
7.3.1	Estimation	134
7.3.2	Prediction	136
7.3.3	Model Comparison: Estimation of $p(D)$	137
7.4	MCMC in BUGS.....	139
7.4.1	Parameter Estimation with BUGS.....	140
7.4.2	BUGS for Prediction	143
7.4.3	BUGS for Model Comparison	145
7.5	Conclusion.....	145
7.6	R Code	146
7.6.1	R Code for a Home-Grown Metropolis Algorithm	146
7.7	Exercises	150

CHAPTER 8	Inferring Two Binomial Proportions via Gibbs Sampling	155
8.1	Prior, Likelihood, and Posterior for Two Proportions	157
8.2	The Posterior via Exact Formal Analysis	159
8.3	The Posterior via Grid Approximation	163
8.4	The Posterior via Markov Chain Monte Carlo	165
8.4.1	Metropolis Algorithm	165
8.4.2	Gibbs Sampling.....	167
8.5	Doing It with BUGS.....	173
8.5.1	Sampling the Prior in BUGS.....	174
8.6	How Different Are the Underlying Biases?.....	175
8.7	Summary	177
8.8	R Code	178
8.8.1	R Code for Grid Approximation (Figures 8.1 and 8.2)	178
8.8.2	R Code for Metropolis Sampler (Figure 8.3).....	181
8.8.3	R Code for BUGS Sampler (Figure 8.6)	184
8.8.4	R Code for Plotting a Posterior Histogram	186
8.9	Exercises	188

第 9 章	多层先验下的伯努利似然	191
9.1	单个铸币厂生产的单枚硬币	192
9.1.1	通过网格近似得到后验估计 1	196
9.2	单个铸币厂生产的多枚硬币	200
9.2.1	通过网格近似得到后验估计 2	203
9.2.2	通过蒙特卡罗抽样得到后验估计	206
9.2.3	单枚铸币估计的离群和收缩	212
9.2.4	案例研究: 触摸治疗	217
9.2.5	硬币数量及每枚硬币的抛掷次数	219
9.3	多个铸币厂生产的多枚硬币	219
9.3.1	独立铸币厂	219
9.3.2	非独立铸币厂	224
9.3.3	个体间差异及 Meta 分析	227
9.4	总结	228
9.5	R 代码	228
9.5.1	触摸治疗实验的 分析代码	228
9.5.2	过滤冷凝实验的 分析代码	231
9.6	练习	235
第 10 章	分层建模和模型比较	241
10.1	多层模型的模型比较	241
10.2	BUGS 中的模型比较	244
10.2.1	一个简单的例子	244
10.2.2	带有伪先验的真实例子	246
10.2.3	在使用带有伪先验的 跨维度 MCMC 时的 一些建议	253
10.3	嵌套模型的模型比较	254
10.4	模型比较的分层框架 回顾	256
10.4.1	MCMC 模型比较的比较方法	257
10.4.2	总结和警告	258
10.5	练习	259
第 11 章	原假设显著性检验 (NHST)	265
11.1	硬币是否均匀的 NHST	267
11.1.1	固定 N 的情况	267

CHAPTER 9	Bernoulli Likelihood with Hierarchical Prior	191
9.1	A Single Coin from a Single Mint	192
9.1.1	Posterior via Grid Approximation	196
9.2	Multiple Coins from a Single Mint.....	200
9.2.1	Posterior via Grid Approximation	203
9.2.2	Posterior via Monte Carlo Sampling	206
9.2.3	Outliers and Shrinkage of Individual Estimates	212
9.2.4	Case Study: Therapeutic Touch	217
9.2.5	Number of Coins and Flips per Coin.....	219
9.3	Multiple Coins from Multiple Mints.....	219
9.3.1	Independent Mints	219
9.3.2	Dependent Mints.....	224
9.3.3	Individual Differences and Meta-Analysis	227
9.4	Summary	228
9.5	R Code	228
9.5.1	Code for Analysis of Therapeutic-Touch Experiment	228
9.5.2	Code for Analysis of Filtration-Condensation Experiment	231
9.6	Exercises	235
CHAPTER 10	Hierarchical Modeling and Model Comparison	241
10.1	Model Comparison as Hierarchical Modeling.....	241
10.2	Model Comparison in BUGS	244
10.2.1	A Simple Example	244
10.2.2	A Realistic Example with “Pseudopriors”	246
10.2.3	Some Practical Advice When Using Transdimensional MCMC with Pseudopriors	253
10.3	Model Comparison and Nested Models	254
10.4	Review of Hierarchical Framework for Model Comparison	256
10.4.1	Comparing Methods for MCMC Model Comparison.....	257
10.4.2	Summary and Caveats	258
10.5	Exercises	259
CHAPTER 11	Null Hypothesis Significance Testing	265
11.1	NHST for the Bias of a Coin.....	267
11.1.1	When the Experimenter Intends to Fix N	267

11.1.2	固定 z 的情况	270
11.1.3	自我反省	272
11.1.4	贝叶斯分析	274
11.2	关于硬币的先验信息	274
11.2.1	NHST 分析	275
11.2.2	贝叶斯分析	275
11.3	置信区间和最高密度区间 (HDI)	277
11.3.1	NHST 置信区间	277
11.3.2	贝叶斯 HDI	280
11.4	多重假设	281
11.4.1	对实验误差的 NHST 修正	282
11.4.2	唯一的贝叶斯后验结论	284
11.4.3	贝叶斯分析如何减少误报	285
11.5	怎样的抽样分布才是好的	286
11.5.1	确定实验方案	286
11.5.2	探索模型预测 (后验预测校验)	287
11.6	练习	288
第 12 章	单点检验的贝叶斯方法	295
12.1	单一先验的估计方法	296
12.1.1	参数的原假设值是否在可信范围内?	297
12.1.2	差异的原假设值是否在可信范围内?	297
12.1.3	实际等值区域 (ROPE)	301
12.2	两个模型的先验比较方法	303
12.2.1	两枚硬币的均匀性是否相同?	303
12.2.2	不同组之间是否有差异?	307
12.3	模型比较的估计	310
12.3.1	原假设值为真的概率是多少?	310
12.3.2	建议	311
12.4	R 代码	312
12.4.1	图 12.5 的 R 代码	312
12.5	练习	314