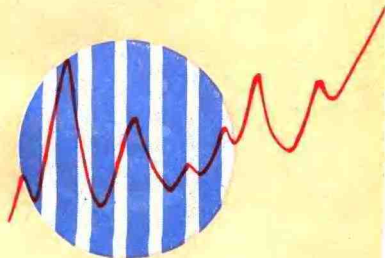


贝叶斯统计学

原理、模型及应用



中国统计出版社

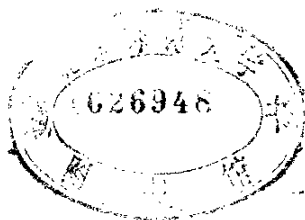
贝叶斯统计学

——原理、模型及应用

S·詹姆斯·普雷斯 著

廖文 陈安贵 等译

袁 卫 校



中国统计出版社

(京)新登字041号

S. James Press

**BAYESIAN STATISTICS, PRINCIPLES,
MODELS, AND APPLICATIONS**

First edition

John Wiley & Sons, Inc., 1989

本书根据1989年版本译出

贝叶斯统计学

——原理、模型及应用

BEIYES TONGJI XUE

S·詹姆斯·普雷斯 著

廖文 陈安贵 等译

袁卫 校

中国统计出版社出版

新华书店北京发行所发行

北京通县永乐印刷厂印刷

850×1168毫米 32开本 8印张 19万字

1992年2月第1版 1992年2月北京第1次印刷

印数: 1—3,000

ISBN7-5037-0670-8/O·11

定价: 8.00元

译者的话

本书根据美国加利福尼亚大学河岸分校统计学系教授S·詹姆斯·普雷斯(S. James Press)的1989年版《*Bayesian Statistics: Principles, Models, and Applications*》一书译出。

这是一本贝叶斯统计学的入门书。它把有关的基本论题集纳于一源,首次对贝叶斯统计学的理论和实践作了论据广泛的概述。它为贝叶斯统计学追根溯源——从贝叶斯以前的时期到现今,并介绍了当今应用的许多贝叶斯计算机程序。

《贝叶斯统计学:原理、模型及应用》阐述的论题涉及最基本的原理、实际应用和案例研究。本书通过描述“贝叶斯原文献”(参见附录3、附录4)的特点,即利用完整的贝叶斯传记梗概再现“贝叶斯原文献”的全貌,继而阐明了贝叶斯思想的来龙去脉。为此,作者采用了大量的图表。例如,在兰德公司进行的估算八十年代核战争爆发的可能性的大众实验结果以及对作者归属有争议的论文进行贝叶斯分析的结果。

上编重点论述贝叶斯统计学的理论,它以詹姆斯·伯努利(James Bernoulli)在1713年所作的研究为肇始,然后转入基本原理的详尽阐述,其中包括概率和决策的公理化系统、一致性和概率定义,“贝叶斯定理”和决策理论的原理以及重要的de Finetti定理。

该编完整地包括了逼近方法、数字方法和其它任何书中都未曾有过的对贝叶斯计算机程序的描述。

下编逐一详细地提供了模型和应用实例。论题包括简单线性回归的贝叶斯推断、贝叶斯MANOVA和MANCOVA程序,归类与判别的贝叶斯推断,等等。

本书既适用于具有基础概率和统计学知识的学生,也同样适

用于从事实际工作的统计工作人员、工程技术人员、心理学和社会科学工作者。

S·詹姆斯·普雷斯教授在过去的25年中，曾在耶鲁大学、加利福尼亚大学洛杉矶分校、伦敦大学学院、伦敦经济与政治学院、不列颠哥伦比亚大学（温哥华）等校任教授和讲师。他是《应用多元分析：运用贝叶斯和频率学派的推断方法》一书的作者（该书有两种版本）。同时还是美国统计协会、数理统计研究所、美国科学促进协会和英国皇家统计学会会员。普雷斯博士在斯坦福大学获得了统计学博士学位。

本书由中国人民大学统计学系88级研究生集体翻译，译者有：黄萍（第一章）、廖文、郭向军（第二章）、杨玉华（第三章）、胡忠兵（第四章）、张泰（第五章）、王志强（第六章）、黎樟林（第七章及论题索引）、孙显福（第八章及附录1.2.3）、陈安贵（附录4）。廖文、陈安贵和黎樟林对个别章节译稿做了修改。中国人民大学统计学系袁卫博士翻译了序言并审校全书。

在武汉大学管理学院院长张尧庭教授和中国人民大学统计学系袁卫博士的热忱帮助下，我们十分荣幸地获得了普雷斯教授的慷慨赠书。张尧庭教授还亲自为本书作序，在此一并表示衷心的感谢！

由于译者水平有限，书中错误在所难免，敬请读者批评指正。

译者

1991年1月于中国人民大学

中译本序言

S·詹姆斯·普雷斯(S. James, Press)写的《贝叶斯统计学：原理、模型及应用》一书是1989年出版的贝叶斯学派的新作。这本书最鲜明的特点是着重对贝叶斯学派观点和丰富的应用实例作充分的介绍，而对一些数学上的繁杂的论证和推导只作扼要的说明。这很适宜于社会科学和应用科学研究人员和一般大学生为了解贝叶斯学派而选用。以前出版的介绍贝叶斯学派的观点和方法的书，往往因为数学推导太多而使读者望而却步。

普雷斯是贝叶斯学派的代表性人物，他并不认为统计是数学的一部分，他认为统计有自己的观念、理论和方法，数学只是一种工具。就象物理、化学要用数学，但它们并不是数学。这本书的写法是相当明显地体现了这一看法的。全书分两部分：理论；模型和应用。理论部分着重于介绍原理和方法；模型和应用部分既介绍了各种常见的统计模型：回归、方差分析，判别等内容的贝叶斯处理的方法，还介绍了各种重要的应用实例。即使在第一部分，他也用生动的例子来说明原理和方法的意义。本书中所举的一些实例，在一般的书上是难以找到的，如：生父的鉴定、估计核战争的概率、心理因素的分析和处理、调查表中拒不回答的表格如何处理、著作权的研究，等等，这些例子涉及的面相当广泛并且引人入胜。他的最后一章就是案例的研究和讨论，所以本书的编写有一个重要的倾向，即它的理论和方法是为了解决实际问题，是重视应用的，尤其是在一些难以处理的社会科学领域中的应用。这些内容对我们的学生、教师和研究工作者会有很多的启示。

这本书还有两个内容是其他的书上难以找到的：(1)全文刊

印了贝叶斯的论文，并对贝叶斯的生平作了较为详细的介绍；(2)关于概率这一概念的理解和应采用什么样的公理系统的论述。这也反映了作者想认真介绍贝叶斯学派的意图。要了解贝叶斯学派的观点，不研究贝叶斯的原作，而只看后来人的解释，这实际上是不全面的。因为后来的人对贝叶斯的论文各人有各人的体会和发展，仁者见仁、智者见智，因而各人所阐发的正是他体会最深的部分，是不是贝叶斯本人的核心思想就难说了，贝叶斯学派本身就发展成各个支派，不正是这个原因吗！因此，了解一下贝叶斯的生平，读一读他写的原文，可以让读者自己去体会贝叶斯的观点和方法。当然，因为贝叶斯写的论文是古代的英语，要翻译成中文是相当不好办的，尽管花了很大的力气，但原文本身由于受当时历史条件的限制，不可能把一些概念论述得象现在这样简明易懂，译文也就难免有些晦涩，这些请读者理解和谅解。关于概率这一概念的讨论，不可能不涉及哲学，书中虽然没有明确地谈论作者自己的哲学观点，然而他列举了种种社会的、自然科学的不确定性的例子之后，他认为概率是对这类不确定的随机现象定量化的一种聪明的办法。他明确地主张应采用条件概率的公理系统〔匈牙利数学家瑞尼(Rényi)提出的公理系统〕作为概率的基础，而不主张用目前流行的柯尔莫果洛夫的公理系统，他在书中详细地论述了这些观点，这也是本书的一个重要的组成部分。

总之，这本书具有它自身的价值，在写法上通俗易懂，一些实例确能引人入胜，在这些方面，很难有其他的书可以和它相比。我衷心感谢中国统计出版社能及时地出版这样一本教材，我相信它对于推动贝叶斯学派在中国的发展会有历史性的，不可磨灭的作用。

武汉大学 张尧庭

1990年12月

序 言

本书试图向受过传统概率统计(频率学派)训练的学生和科学研究人员展示贝叶斯统计学的全貌。读者阅读本书需要掌握微积分和线性代数的知识,并且仅此而已。本书不仅面向统计理论工作者,更面向广大应用统计人员,着重于应用贝叶斯方法解决各领域中的实际问题。对于应用统计工作人员来说,掌握贝叶斯统计的基本原理和统计思想是至关重要的。因为只有在此基础上,他们才能(1)更好地应用统计方法解决实际问题;(2)深感荣幸地向上级决策者推荐使用贝叶斯方法进行决策;(3)饶有兴趣地深入研究这种方法,进而丰富和发展贝叶斯统计学。正是基于应用贝叶斯方法去解决实际问题,本书将尽可能全面地介绍贝叶斯统计的基本原理、思想及各种应用,而尽可能少地给予理论的证明。

全书分为上下两编,每编四章。上编介绍贝叶斯理论,下编给出了各种模型和应用案例。正文之后附有介绍贝叶斯生平和贝叶斯最初提出该理论的论文原稿的四个附录。此外,还有文献目录和论题索引。

第一章首先讨论贝叶斯统计推断与决策独自形成的一套完整的逻辑公理体系。如果违背贝叶斯原理,所进行的统计推断与决策将产生较大的偏误。换言之,其后果将导致风险决策既不经济又缺乏合理性。

第二章介绍贝叶斯统计学的基本原理。贝叶斯定理既适用于离散随机变量,也同样适用于完全连续的随机变量。我们将讨论贝叶斯估计、假设检验和决策理论。在这一章,我们要介绍先验分布、贝叶斯因子、重要的 de Finetti 定理、似然原则和预报分布。

第三章给出了贝叶斯方法应用中各种复杂后验分布的近似方法。我们向读者展示大样本理论的结果和拉普拉斯后验密度积分的近似解法。我们还将讨论高维积分时后验密度近似解法中样本信息和分布模拟的重要性及其相互关系。在这一章的最后，我们还对现代贝叶斯方法的各种计算机应用软件进行了全面介绍。

第四章介绍应用专家意见主观地评价先验分布的方法。作为应用案例，我们讨论如何使用专家估价法估计出本世纪80年代世界爆发核战争的多元先验分布。

第五章讨论一元和多元回归模型中的贝叶斯推断。在该章中，我们在回归模型中使用无差异先验分布和预报分布去预报未来的观察值。

第六章继续讨论第五章提出的一般线性统计模型问题，并在这章给出多元情况下的贝叶斯方差分析和协方差分析。在这些题目中，我们将讨论 de Finetti 的总体均值向量分布中的“可交换性”概念。

第七章介绍贝叶斯归类与判别过程的理论与应用。在抽样调查中，对于首次未能做出明确回答的调查问题，我们还需要进一步的研究分析。在这一章中，将应用贝叶斯归类与判别理论解决上述抽样调查中的问题。

第八章给出一个应用贝叶斯方法解决美国建国初期《联邦党人论文集》中有争议的著作权问题的案例，我们相信它会引起读者浓厚的兴趣。

本书可作为高年级本科生或研究生的贝叶斯统计学教材。学时可安排一学期。上半学期可讲授前两章及附录，下半学期学习第三章至第八章。如果课时不够，某些章节可略去不讲。至于略去哪些章节，要视教师和学生的兴趣而决。我在这里所提到的兴趣是指该课堂之教学目的是侧重于理论还是侧重于模型和应用。

读者手中的这本书源于1984年7月我在南澳大利亚Adelaide大学授课所用的讲义。当时的讲座是由澳大利亚联邦科学与工业

研究组织(CSIRO)的数学与统计学部和 Adelaide 大学经济学与统计学系共同举办的。我要感谢 CSIRO 的全体同事，没有他们的鼓励、支持和帮助，旧讲稿不可能得到修改和完善，这本书也不可能与广大读者见面。（下略）

S·詹姆斯·普雷斯

1989年1月于加利福尼亚

目 录

上编 理 论

第一章 基础知识.....	(3)
1.1 引言.....	(3)
1.2 随机性.....	(3)
1.3 概率.....	(7)
1.3.1 公理化系统.....	(7)
1.3.2 相合性.....	(12)
1.3.3 主观概率信念的操作.....	(13)
1.3.4 各种概率定义的比较.....	(15)
1.4 托马斯·贝叶斯.....	(16)
1.5 小结.....	(19)
第一章补充: Leonard J. Savage 决策的公理化 基础.....	(20)
练习.....	(21)
第二章 原理.....	(23)
2.1 引言.....	(23)
2.2 离散参数的贝叶斯定理.....	(23)
2.2.1 例子 (零件次品).....	(23)
2.2.2 贝叶斯定理 (离散参数).....	(24)
2.2.3 评注.....	(24)
2.2.4 贝叶斯定理的解释.....	(25)
2.3 贝叶斯方法.....	(25)
2.3.1 估计: 决策理论——效用.....	(25)
区间估计: 可信区间 (连续参数).....	(30)

2.3.2 假设检验	(34)
简单对简单	(34)
简单对复合	(36)
Lindley 悖论	(38)
2.4 离散参数二项分布	(39)
2.4.1 例子	(39)
2.4.2 零件次品例子的求解	(40)
2.5 贝叶斯定理 (连续参数)	(41)
2.6 例: 连续参数二项分布	(42)
2.6.1 一致先验	(42)
2.6.2 零件次品例子	(42)
2.7 先验分布	(44)
2.7.1 贝叶斯与非贝叶斯推断方法: 经验贝叶斯估计	(44)
2.7.2 模糊 (无偏好) 先验	(50)
平稳估计原理	(53)
2.7.3 自然共轭先验; g ——先验	(54)
基于数据的先验	(55)
g ——先验	(56)
2.7.4 估计完备先验分布	(57)
2.8 似然原理	(58)
似然原理与限定条件	(59)
2.9 预报分布; de Finetti 定理; 预报学派的基本原理	(59)
2.9.1 预报分布	(59)
定义	(60)
零件次品例子的预报密度	(60)
2.9.2 可交换性	(61)
定义	(61)

一个不独立的(多元)可交换性的例子.....	(61)
2.9.3 de Finetti 定理.....	(62)
2.10 稳健与非参数贝叶斯推断: 贝叶斯自助法.....	(64)
2.11 例子: 生父鉴定.....	(66)
观察数据.....	(67)
2.12 小结: 贝叶斯方法的优点.....	(68)
练习.....	(69)
第三章 逼近法、数值法与计算机程序.....	(71)
3.1 引言.....	(71)
3.2 大样本的后验分布.....	(72)
3.3 贝叶斯积分的逼近求值.....	(74)
3.3.1 低维(数值逼近法),	(75)
Lindley逼近法.....	(75)
Tierney—Kadane 逼近法.....	(78)
Naylor—Smith 逼近法.....	(81)
3.3.2 高维(Monte Carlo 积分)	(83)
3.4 贝叶斯分布的模拟.....	(85)
3.5 贝叶斯计算机程序.....	(87)
3.5.1 引言.....	(87)
3.5.2 程序概要.....	(87)
3.6 小结.....	(104)
练习.....	(104)
第四章 多元先验分布的估计: 估计核战争爆发的概率.....	(106)
4.1 引言.....	(106)
4.2 多元主观估计.....	(107)
4.3 一般看法.....	(107)
4.4 模型.....	(107)
4.5 多元密度估计.....	(108)
正态核.....	(109)

4.6 群体估计方法概要	(109)
4.7 经验应用: 核战争	(110)
4.7.1 回答的一致性	(114)
4.7.2 含义	(115)
4.7.3 直方图	(115)
4.7.4 修匀的先验密度 (调整后密度)	(119)
4.8 与主观概率估计有关的心理因素	(121)
关于心理因素的结论	(123)
4.9 小结	(124)
练习	(124)

下编 模型与应用

第五章 回归中的贝叶斯推断	(127)
5.1 引言	(127)
5.2 简单线性回归	(127)
5.2.1 模型	(127)
似然函数	(128)
先验	(128)
5.2.2 预报分布	(131)
5.2.3 关于标准差的后验推断	(132)
5.3 多元回归模型	(133)
5.3.1 多元模糊先验	(133)
5.3.2 多元回归	(134)
似然函数	(135)
模糊先验	(136)
后验分析	(136)
5.3.3 关于协方差矩阵的后验推断	(137)
5.3.4 预报密度	(138)
5.4 小结	(138)

练习	(138)
第六章 贝叶斯多元方差与协方差分析	(141)
6.1 引言	(141)
6.2 MANOVA 模型: 单向型	(141)
对回归形式进行简化	(142)
6.3 似然	(142)
6.4 先验	(143)
6.5 在 MANOVA 问题中可交换假设的实际含义	(143)
其它含义	(144)
6.6 后验	(145)
6.6.1 联合后验	(145)
6.6.2 条件后验	(145)
6.6.3 边缘后验	(146)
6.6.4 例子: 均衡设计	(146)
6.7 例子: 考试分数	(147)
6.8 效应的后验分布	(149)
6.9 小结	(150)
练习	(151)
第七章 归类与判别中的贝叶斯推断	(152)
7.1 引言	(152)
7.2 似然函数	(153)
7.3 先验密度	(154)
7.4 后验密度	(154)
7.5 预报密度	(154)
7.6 后验归类概率	(155)
7.7 例子: 两个总体	(157)
7.8 对不确定回答者的事后推测: 一种应用	(158)
7.8.1 问题	(158)
解答	(158)

7.8.2例证	(160)
7.9小结	(161)
练习	(162)
第八章 贝叶斯推断应用的一个案例研究	(164)
8.1引言	(164)
8.2联邦党人论文集	(165)
8.3似然函数	(165)
8.4著作权争议问题的后验机会	(166)
8.5参数的先验分布	(168)
8.6最终结果	(171)
8.7小结	(172)
练习	(172)

附 录

附录1 托马斯·贝叶斯 作者 Hilary L. Seal	(173)
附录2 托马斯·贝叶斯——一第传记性摘记 作者 G. A. Barnard	(181)
附录3 论机会学说中的一个问题 作者: The Late Rev. Mr Bayes, F.R.S.	(185)
附录4 论机会学说中的一个问题 作者: Reverend Thomas Bayes	(188)
文献目录	(212)
论题索引	(231)

上 编

理 论

• 1 •