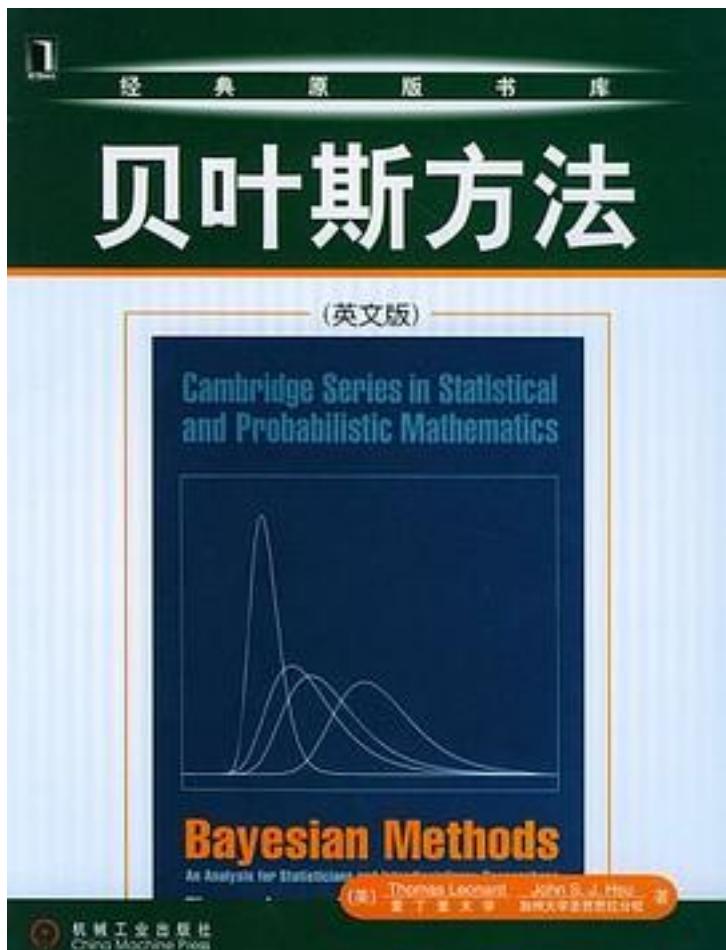


贝叶斯方法



[贝叶斯方法 下载链接1](#)

著者:[加] Cameron Davidson-Pilon

出版者:人民邮电出版社

出版时间:2017-1-1

装帧:平装

isbn:9787115438805

贝叶斯推理的方法非常自然和极其强大。然而，大多数图书讨论贝叶斯推理，依赖于非常复杂的数学分析和人工的例子，使没有强大数学背景的人无法接触。不过，现在好了，卡梅伦的这本书从编程、计算的角度来介绍贝叶斯推理，把贝叶斯理论和编程实践结

合起来，使大多数程序员都可以入门并掌握。

本书通过强大的Python语言库PyMC，以及相关的Python工具，包括NumPy\SciPy\Matplotlib讲解了概率编程。通过本书介绍的方法，读者只需付出很少的努力，就能掌握有效的贝叶斯分析方法。

作者介绍:

作者

Cameron

Davidson-Pilon，接触过数学在多个领域的应用——从基因和疾病的动态演化，到金融价格的随机模型。他对于开源社区主要的贡献包括这本书以及lifelines项目。Cameron成长于加拿大的安大略省圭尔夫市，而就读于滑铁卢大学以及莫斯科独立大学。如今他住在安大略省渥太华市，并在电商领军者Shopify工作。

译者

辛愿，浙江大学硕士毕业，腾讯公司基础研究高级工程师，舆情系统开发经理。曾在百度从事推荐系统、用户画像、数据采集等相关研究工作，拥有多项专利，组织过上海大数据技术沙龙。目前专注于文本挖掘、舆情分析、智能聊天机器人等相关领域。

钟黎，腾讯公司研究员。曾在中国科学院、微软亚洲研究院、IBM研究院（新加坡）从事图像处理、语音处理、机器学习等相关研究工作，拥有多项专利，目前聚焦在自然语言处理、深度学习和人工智能等相关领域。

欧阳婷，华南理工大学硕士毕业，腾讯公司后台策略工程师。在电信、互联网行业参与过推荐系统、资源优化、KPI预测、用户画像等相关项目，拥有多项专利，目前聚焦在欺诈检测、时序分析、业务安全等相关领域。

审校者

余凯博士，地平线机器人技术创始人、CEO，国际杰出机器学习专家，中国人工智能学会副秘书长。余博士是前百度研究院执行院长，创建了百度深度学习研究院。他在百度所领导的团队在广告变现、搜索排序、语音识别、计算机视觉等领域做出杰出贡献，创纪录地连续三次获得公司高荣誉——“百度奖”。他还创建了中国公司自动驾驶项目，后发展为百度自动驾驶事业部。

岳亚丁博士，腾讯公司专家研究员，腾讯技术职级评委会基础研究岗位的负责委员。岳博士拥有19年在金融、电信、互联网行业的数据挖掘经验，主导或参与过用户画像、在线广告、推荐系统、CRM、欺诈检测、KPI预测等多种项目。他曾在微软（加拿大）从事行为定向广告的模型研发，另有11年的工程结构、海洋水文气象的力学研究及应用的工作经历。

目录: 目录

第1章 贝叶斯推断的哲学 1

1.1 引言 1

1.1.1 贝叶斯思维 1

1.1.2 贝叶斯推断在实践中的运用 3

1.1.3 频率派的模型是错误的吗？ 4

1.1.4 关于大数据 4

1.2 我们的贝叶斯框架 5

1.2.1 不得不讲的实例：抛硬币	5
1.2.2 实例：图书管理员还是农民	6
1.3 概率分布	8
1.3.1 离散情况	9
1.3.2 连续情况	10
1.3.3 什么是	12
1.4 使用计算机执行贝叶斯推断	12
1.4.1 实例：从短信数据推断行为	12
1.4.2 介绍我们的第一板斧：PyMC	14
1.4.3 说明	18
1.4.4 后验样本到底有什么用？	18
1.5 结论	20
1.6 补充说明	20
1.6.1 从统计学上确定两个值是否真的不一样	20
1.6.2 扩充至两个转折点	22
1.7 习题	24
1.8 答案	24
第2章 进一步了解PyMC	27
2.1 引言	27
2.1.1 父变量与子变量的关系	27
2.1.2 PyMC变量	28
2.1.3 在模型中加入观测值	31
2.1.4 最后……	33
2.2 建模方法	33
2.2.1 同样的故事，不同的结局	35
2.2.2 实例：贝叶斯A/B测试	38
2.2.3 一个简单的场景	38
2.2.4 A和B一起	41
2.2.5 实例：一种人类谎言的算法	45
2.2.6 二项分布	45
2.2.7 实例：学生作弊	46
2.2.8 另一种PyMC模型	50
2.2.9 更多的PyMC技巧	51
2.2.10 实例：挑战者号事故	52
2.2.11 正态分布	55
2.2.12 挑战者号事故当天发生了什么？	61
2.3 我们的模型适用吗？	61
2.4 结论	68
2.5 补充说明	68
2.6 习题	69
2.7 答案	69
第3章 打开MCMC的黑盒子	71
3.1 贝叶斯景象图	71
3.1.1 使用MCMC来探索景象图	77
3.1.2 MCMC算法的实现	78
3.1.3 后验的其他近似解法	79
3.1.4 实例：使用混合模型进行无监督聚类	79
3.1.5 不要混淆不同的后验样本	88
3.1.6 使用MAP来改进收敛性	91
3.2 收敛的判断	92
3.2.1 自相关	92
3.2.2 稀释	95
3.2.3 pymc.Matplot.plot()	97
3.3 MCMC的一些秘诀	98

3.3.1 聪明的初始值	98
3.3.2 先验	99
3.3.3 统计计算的无名定理	99
3.4 结论	99
第4章 从未言明的最伟大定理	101
4.1 引言	101
4.2 大数定律	101
4.2.1 直觉	101
4.2.2 实例：泊松随机变量的收敛	102
4.2.3 如何计算 $\text{Var}(Z)$	106
4.2.4 期望和概率	106
4.2.5 所有这些与贝叶斯统计有什么关系呢	107
4.3 小数据的无序性	107
4.3.1 实例：地理数据聚合	107
4.3.2 实例：Kaggle的美国人口普查反馈比例预测比赛	109
4.3.3 实例：如何对Reddit网站上的评论进行排序	111
4.3.4 排序！	115
4.3.5 但是这样做的实时性太差了	117
4.3.6 推广到评星系统	122
4.4 结论	122
4.5 补充说明	122
4.6 习题	123
4.7 答案	124
第5章 失去一只手臂还是一条腿	127
5.1 引言	127
5.2 损失函数	127
5.2.1 现实世界中的损失函数	129
5.2.2 实例：优化“价格竞猜”游戏的展品出价	130
5.3 机器学习中的贝叶斯方法	138
5.3.1 实例：金融预测	139
5.3.2 实例：Kaggle观测暗世界大赛	144
5.3.3 数据	145
5.3.4 先验	146
5.3.5 训练和PyMC实现	147
5.4 结论	156
第6章 弄清楚先验	157
6.1 引言	157
6.2 主观与客观先验	157
6.2.1 客观先验	157
6.2.2 主观先验	158
6.2.3 决策，决策……	159
6.2.4 经验贝叶斯	160
6.3 需要知道的有用的先验	161
6.3.1 Gamma分布	161
6.3.2 威沙特分布	162
6.3.3 Beta分布	163
6.4 实例：贝叶斯多臂老虎机	164
6.4.1 应用	165
6.4.2 一个解决方案	165
6.4.3 好坏衡量标准	169
6.4.4 扩展算法	173
6.5 从领域专家处获得先验分布	176
6.5.1 试验轮盘赌法	176
6.5.2 实例：股票收益	177

6.5.3 对于威沙特分布的专业提示	184
6.6 共轭先验	185
6.7 杰弗里斯先验	185
6.8 当N增加时对先验的影响	187
6.9 结论	189
6.10 补充说明	190
6.10.1 带惩罚的线性回归的贝叶斯视角	190
6.10.2 选择退化的先验	192
第7章 贝叶斯A B测试	195
7.1 引言	195
7.2 转化率测试的简单重述	195
7.3 增加一个线性损失函数	198
7.3.1 收入期望的分析	198
7.3.2 延伸到A B测试	202
7.4 超越转化率: t检验	204
7.4.1 t检验的设定	204
7.5 增幅的估计	207
7.5.1 创建点估计	210
7.6 结论	211
术语表	213
• • • • • (收起)	

[贝叶斯方法](#) [下载链接1](#)

标签

贝叶斯方法

Python

贝叶斯

统计学

机器学习

概率编程与贝叶斯推断

计算机科学

数据挖掘

评论

注重讲例子，重点向读者展示“结果管用”，但很多应该讲所以然的时候就一带而过了。特别是，有很多代码也没有解释，如果不熟悉scipy的话，会突然看得莫名其妙。图漂亮。

两天时间翻完了，代码占一半，还不解释代码函数都什么意思，理论方面也是只提概念不解释。不适合新手入门读，理论好的看这本书也学不到pymc库多少知识。真不知道这本书面向的读者到底是哪一类。看完收获不大，倒是学了画各种图的方法。最后奉劝大家，看见书名和序言里有余凯俩字就不要浪费时间了。

理论欠缺，都是应用。

缺相关理论知识

用概率面拉扯解释贝叶斯很明白；翻译错误不少，译者显然在细节上功夫不够

耳目一新，案例引起兴趣，但是解释的不清楚

给的代码很齐全，几乎不介绍理论，对于只是想使用概率编程的同学倒是省了不少麻烦事。可惜我现在在用PyMC3，还是多少有些区别，思想是一样的。

写的最细的书了。

大致了解 有点模糊

理论上这个书是比较简单易懂的，但是我看他就死活看不明白，最后下定决心看的韦来生的贝叶斯统计才学懂

特别适合我这种侧重方法应用的小白…十分感谢作者！简直舍不得读完，但还是读完了。

太简略

最近的感觉是，看再多的书也不如实践一个模型来得惊心动魄。

内容还行，面向小白挺好的。语言特别绕，啰嗦又散，或许一些老外写书就是这样吧，但是翻译的也的确一般。

写的很好，不过看完感觉没啥用，太浅了。直接用还好，想理解贝叶斯还是不够

现代的编程书,github过万赞,代码协作实时更新jupyter notebook交互编程,配套中文翻译,快速上手贝叶斯方法

译本还在讲pymc, pymc3的版本在github上有了。本书权当离线中文版看。。

适合贝叶斯推断领域的入门和初步感知

幼稚的内容蹩脚的翻译。我又被豆瓣骗了 还我50元

理论讲的不够清晰

[贝叶斯方法_下载链接1](#)

书评

Jupyter

<http://nbviewer.jupyter.org/github/CamDavidsonPilon/Probabilistic-Programming-and-Bayesian-Methods-for-Hackers/tree/master/> GitHub

<https://github.com/CamDavidsonPilon/Probabilistic-Programming-and-Bayesian-Methods-for-Hackers>

只有六章 [<https://download.csdn.net/download/wizardforcel/10307474>]

[贝叶斯方法_下载链接1](#)