

统计学关我什么事：生活中的极简统计学(贝叶斯统计学的超级入门书，日本畅销10万册。)

作者：小岛宽之

图书在版编目（CIP）数据

统计学关我什么事：生活中的极简统计学/（日）小岛宽之著；罗梦迪译。--北京：北京时代华文书局，2018.4

ISBN 978-7-5699-2304-9

I. ①统... II. ①小...②罗... III. ①统计学—通俗读物 IV. ①C8-49

中国版本图书馆CIP数据核字（2018）第055781号

KANZEN DOKUSHU BAYES TOKEIGAKU NYUMON by HIROYUKI KOJIMA

Copyright 2015 HIROYUKI KOJIMA

Chinese（in simplified character only）translation copyright 2018 by Beijing Time-Chinese Publishing House Co.，Ltd.

All rights reserved.

Original Japanese language edition published by Diamond，Inc.

Chinese（in simplified character only）translation rights arranged with Diamond，Inc.

through BARDON-CHINESE MEDIA AGENCY.

北京市版权著作权合同登记号 字：01-2017-2435

统计学关我什么事：生活中的极简统计学

TONGJIXUE GUANWO SHENMESHU: SHENGHUO ZHONG DE JIJIAN TONGJIXUE

作 者 （日）小岛宽之

译 者 罗梦迪

出版人 王训海

选题策划 胡俊生

责任编辑 张超峰

装帧设计 红杉林文化 嘉承设计

责任印制 刘 银

出版发行 北京时代华文书局 <http://www.bjsdsj.com.cn>

北京市东城区安定门外大街138号皇城国际大厦A座8楼

邮 编：100011

电 话：010-64267955 64267677

印 刷 三河市祥达印刷包装有限公司 0316-3656589

（如发现印装质量问题，请与印刷厂联系调换）

开 本 880mm×1230mm 1/32

印 张 8.25 字 数 180千字

版 次 2018年6月第1版

印 次 2018年6月第1次印刷

书 号 978-7-5699-2304-9

版权所有，侵权必究

目录

[第0讲 只要会做四则运算，便可掌握贝叶斯统计学](#)

[本书的特点](#)

[第1部 快速学习！理解贝叶斯统计学的精髓](#)

[第1讲 信息增加导致概率变化 “贝叶斯推理”的基本方法](#)

[小结](#)

[练习题](#)

[第2讲 贝叶斯推理的结果，有时与直觉大相径庭① 使用客观数据时的注意事项](#)

[小结](#)

[练习题](#)

[第3讲 根据主观数字也可以进行推理 疑惑时分的“理由不充分原理”](#)

[小结](#)

[练习题](#)

[第4讲 运用“概率的概率”，拓宽推理范围](#)

[小结](#)

[练习题](#)

[专栏 贝叶斯是何许人也？](#)

[第5讲 从推算过程开始，逐渐明确的贝叶斯推理的特征](#)

[小结](#)

[练习题](#)

[第6讲 明快而严格，但其使用场合受到限制的内曼-皮尔逊式推理](#)

[小结](#)

[练习题](#)

[第7讲 通过少量信息得出切实结论的贝叶斯推理与内曼-皮尔逊式推理的差异](#)

[小结](#)

[练习题](#)

[第8讲 贝叶斯推理的基础：极大似然原理 贝叶斯统计学与内曼-皮尔逊统计学的衔接点](#)

[小结](#)

[练习题](#)

[第9讲 贝叶斯推理的结果，有时与直觉大相径庭② 蒙蒂霍尔问题与三个囚犯的问题](#)

[小结](#)

[练习题](#)

[专栏 关于“幸运”的两条法则](#)

[第10讲 掌握多条信息时的推理① 运用“独立试验的概率乘法公式”](#)

[小结](#)

[练习题](#)

[第11讲 掌握多条信息时的推理② 以垃圾邮件过滤器为例](#)

[小结](#)

[练习题](#)

[第12讲 在贝叶斯推理中可以依次使用信息 “序贯理性”](#)

[小结](#)

[练习题](#)

[第13讲 每获得一条信息，贝叶斯推理就变得更精确一些](#)

[小结](#)

[练习题](#)

[专栏 帮助贝叶斯复兴的学者们](#)

[第2部 完全自学！从“概率论”到“正态分布”](#)

[第14讲 “概率”与“面积”的性质相同 概率论的基础](#)

[小结](#)

[练习题](#)

[第15讲 在获得信息之后，概率的表示方法 “条件概率”的基本性质](#)

[小结](#)

[练习题](#)

[第16讲 “概率分布图”帮助我们进行更加通用的推理](#)

[小结](#)

[练习题](#)

[第17讲 “贝塔分布”的性质由两个数字决定](#)

[小结](#)

[练习题](#)

[第18讲 决定概率分布性质的“期待值”](#)

[小结](#)

[练习题](#)

[专栏 何为“主观概率”?](#)

[第19讲 在“贝塔分布”中使用概率分布图进行高级推理](#)

[小结](#)

[练习题](#)

[第20讲 在抛硬币或天体观测时观察到的“正态分布”](#)

[小结](#)

[练习题](#)

[第21讲 在“正态分布”中使用概率分布图进行高级推理](#)

[小结](#)

[练习题](#)

[补讲 贝塔分布的积分计算](#)

[结语](#)

[参考文献](#)

[练习题参考答案](#)

第0讲 只要会做四则运算，便可掌握贝叶斯统计学

本书的特点

0-1 从零基础达到应用水平

本书是“贝叶斯统计学”的超级入门书。“超级”的含义：

- 从零基础开始学习
- 抛开烦琐的符号和计算过程，学习运用贝叶斯统计
- 不只是随便说说，而是毫无保留地传授知识

对贝叶斯统计学感兴趣的人不在少数。然而此前的教科书，导入部分编写浅显，中途却难度骤增，这使很多读者大受挫折。这是因为在尚未理解贝叶斯统计的本质时，就被灌输了一大堆概率符号，使得理解起来更为困难。

为了不再重蹈这样的覆辙，本书编写之时做了一些功课，具体会在下节进行说明：

0-2 仅使用面积图和简单算术

贝叶斯统计的基础是概率公式——“贝叶斯公式”，它立足于“条件概率”的发展事项。“贝叶斯公式”是高等数学中很难理解的一个概念，原因有二：第一，公式复杂而不够直观；第二，条件概率在某种程度上属于“不可靠”的概念，对于思维缜密的人来说总觉得“哪里有些奇怪”。

事实上，上述第二点在贝叶斯统计中是至关重要的。因为正是这份“不可靠”，才是贝叶斯统计的本质，它与便利性息息相关。后面我们会讲到，贝叶斯统计在20世纪初曾因为其“不可靠”而遭到批判，一度被斥于统计学之外。但由于贝叶斯统计的“不可靠”与“便利性”为表里一致的关系，“正因为不可靠才得以运用”。在一部分学者对于这种“便利性”的关注下，贝叶斯统计于20世纪后半期恢复了其应有的地位。在21世纪的今天，贝叶斯统计已经成为统计学的主流。

笔者着重考虑了这两点，在编写过程中也有所侧重，并做了如下功课。

功课1 将不出现“贝叶斯公式”（极少一部分除外）的方针贯彻到底

以“通过面积图进行图解”的方针作为贝叶斯公式的替代。从本质上来讲，二者是相同的，然而对于大多数读者而言，图解的方式更加直观且易于理解。同时，通过“面积图”可以更清晰地看出“贝叶斯公式”的“不可靠”和“便利性”究竟体现在哪里。

功课2 只需简单算术的计算水平即可

这意味着，只需要会做四则运算就可以掌握了，连开方和文字式计算都不需要。而且这四则运算，即使是不擅长手算的人也可以借助计算器轻而易举地完成。

当然，在本书末尾会出现“贝塔分布”“正态分布”这些有难度的概念。因为如果不介绍这些概念，是无法达到前文所述“毫无保留的传授”程度的。全面理解这些概念，需要用到大学的微分积分知识，这对于许多读者来说是很大的负担。因此在本书中也只能作一些相对简单的解说。

这也就是说，本书的方针——向读者灌输仅通过四则运算就能掌握的公式。这也是本书编写时所做的功课之一。在这个意义上，本书并非“充分齐全”的教材。然而如果想要“充分理解”贝叶斯统计学的人，不妨在读过本书之后再试着挑战一下专业书籍。本书的目的是抛开烦琐的数学概念，将“贝叶斯统计学隐藏的本质”剖析呈现出来。

0-3 比尔·盖茨也在关注它！贝叶斯统计在商业活动中的应用

随着因特网的普及和同步技术的发展，贝叶斯统计开始运用于商业领域。通过互联网可以实现自动收集顾客的购买和检索记录，从而推测顾客类别。在这一点上，贝叶斯统计学完胜传统意义上的统计学。

如今，许多互联网企业都在实际应用贝叶斯统计。其中，微软由于很早就开始在商业活动中运用贝叶斯统计学而闻名。Windows的操作系统帮助功能中就导入了贝叶斯统计。此外，在网上搜索“小孩病症”的时候，优先显示可靠结果的软件也已经开发出来。微软的前董事长比尔·盖茨在1996年曾在报纸内容中称，微软之所以在激烈的市场竞争中胜出，正是由于采用了贝叶斯统计。比尔·盖茨还在2001年关于基本方针的演讲中称，微软的21世纪战略正是贝叶斯统计战略，公开表示，已经在全世界范围内挖到了许多贝叶斯统计研究人才。该发言引起了很大关注。

谷歌搜索引擎的自动翻译系统中也引入了贝叶斯统计技术。

当然，贝叶斯统计技术在IT企业之外的各个领域也有着广泛应用。例如，消除传真图像中的杂音就运用了贝叶斯统计技术。此外，医疗领域的“自动诊断系统”等也需要用到贝叶斯统计。

通过阅读本书可以得知，贝叶斯统计的优势在于，“在数据少的情况下也可以进行推测，数据越多，推测结果越准确”，以及“对所获的信息可做出瞬时反应，自动升级推测”的学习功能。了解了这一点之后，就完全可以理解为什么贝叶斯统计是非常适合应用于高端商业的技术了。

从事商业活动的人，如果能够熟练使用贝叶斯统计，那是再好不过的。本书中的案例和解说，为这一类人群提供了很好的参考。

0-4 贝叶斯统计依存于人的心理

在0-2节中有提到，“贝叶斯统计在某种程度上是不可靠的”。究其原因，是由于贝叶斯统计中所涉及的概率是“主观的”。换言之，通过贝叶斯统计得到的概率并非客观的数值，而是依存于人的心理的主观数值。在从这个意义上讲，贝叶斯统计具备了一定的“思想”。也正是因此，注重客观性的科学界为贝叶斯统计打上了“假冒伪劣”的烙印，并导致它一度消亡。

然而，遗憾的是，关于贝叶斯统计学的绝大部分书籍中，并未对这一问题进行记载。也许是作者们不愿将其公之于众，抑或只是因为他们对此不甚了解罢了。实际上，几乎没有一本教科书对于这个问题正面进行过阐述。然而，所谓的“主观性”和“思想性”，才正是贝叶斯统计学的本质和它具有便利性的根本原因所在。因此，在解说贝叶斯统计学的时候，如果忽视掉这一点，是难以将贝叶斯统计学的本质传达给读者的。

本书不刻意避开贝叶斯统计的“主观性”和“思想性”，而是将这些特点展现出来进行解说，特别是对于贝叶斯统计学与传统的统计学之间的差异进行详解。希望众多读者能够为贝叶斯统计学的神奇和有趣拍手称赞。

0-5 附带简单的填空练习题，适合自学

本书沿袭之前出版的《完全自学 统计学入门》（钻石社）的编写方法，用最详尽的语言解释说明，并在每一讲之后设置简单的填空练习题。学习数学的最佳方法是做一些简单的练习题。本书中收录的练习题并非应用题的形式，而是用来对讲义内容进行巩固的，希望各位读者认真练习，加深理解。

读完这本书，您一定会产生这样的想法：

“咦？明明没有经过登山训练，却不知不觉到了山顶呢！”

那么，就让我们向着山顶，出发吧。

第1部 快速学习！理解贝叶斯统计学的精髓

在第1部中，将为您解说关于“贝叶斯统计的推算应该用何种方法来思考，具有什么样的性质”的问题。解说中采用了我们身边的许多事例，如“这位顾客是来买东西，还是随便逛逛”“收到的是真命巧克力？还是义理巧克力”对于读者来说，这些例子应当是很容易想象和理解的。另一方面，本书内容涉及贝叶斯统计学与“序贯理性”“内曼-皮尔逊统计学”的区别，这对于贝叶斯统计学的特征，已经探讨得相当深入了。

欢迎访问：电子书学习和下载网站 (<https://www.shgis.cn>)

文档名称：《统计学关我什么事：生活中的极简统计学》小岛宽之 著.epub

请登录 <https://shgis.cn/post/1108.html> 下载完整文档。

手机端请扫码查看：

