

TURING

图灵新知

统计学大师之路

乔治·博克斯回忆录

An Accidental Statistician

The Life and Memories of George E.P. Box

[英] 乔治·博克斯 (George E.P. Box) 著

朱磊磊 译

所有模型都是错误的，
但有些是有用的。



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

统计学大师之路

乔治·博克斯回忆录 An Accidental Statistician
The Life and Memories of George E.P. Box

[英] 乔治·博克斯 (George E.P. Box) 著
朱磊磊 译



George E.P. Box

人民邮电出版社
北京

图书在版编目(CIP)数据

统计学大师之路：乔治·博克斯回忆录 / (英) 乔治·博克斯 (George E. P. Box) 著；朱磊磊译. — 北京：人民邮电出版社，2018.7

(图灵新知)

ISBN 978-7-115-48655-4

I. ①统… II. ①乔… ②朱… III. ①乔治·博克斯 (1919-2013)—回忆录 IV. ①K837.125.31

中国版本图书馆CIP数据核字(2018)第125679号

内 容 提 要

乔治·博克斯(1919—2013)是二十世纪下半叶的统计学大师之一，在实验设计、时间序列分析、统计控制和贝叶斯推断等方面做出了重要贡献，深刻影响了统计学、工程学、化学、经济学和环境科学等领域的理论和实践。本书是他在晚年所写的回忆录，其中回想了他的成长经历以及偶然踏上统计学家生涯并从产业界走向学术界的过程，追忆了他与二十世纪的一些重要统计学家的交往和合作，并分享了他对于统计学、教学科研、创新等话题的思考。

◆ 著 [英] 乔治·博克斯 (George E.P. Box)

译 朱磊磊

责任编辑 楼伟珊

责任印制 周昇亮

◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号

邮编 100164 电子邮件 315@ptpress.com.cn

网址 <http://www.ptpress.com.cn>

三河市君旺印务有限公司印刷

◆ 开本：880×1230 1/32

印张：10.5

字数：245千字

2018年7月第1版

印数：1—4 000册

2018年7月河北第1次印刷

著作权合同登记号 图字：01-2014-0501号

定价：49.00元

读者服务热线：(010)51095186转600 印装质量热线：(010)81055316

反盗版热线：(010)81055315

广告经营许可证：京东工商广登字 20170147 号



站在巨人的肩上

Standing on the Shoulders of Giants

版权声明

All Rights Reserved. This translation published under license. Authorized translation from the English language edition, entitled *An Accidental Statistician: The Life and Memories of George E.P. Box*, ISBN 978-1-118-40088-3, by George E.P. Box, Published by John Wiley & Sons. Copyright © 2013 by John Wiley & Sons, Inc. No part of this book may be reproduced in any form without the written permission of the original copyrights holder.

Simplified Chinese translation edition published by POSTS & TELECOM PRESS Copyright © 2018.

本书简体中文版由 John Wiley & Sons, Inc. 授权人民邮电出版社独家出版。

本书封底贴有 John Wiley & Sons, Inc. 激光防伪标签，无标签者不得销售。

版权所有，侵权必究。

谨以本书献给我的学生们，
能与他们合作，并成为朋友是我的荣幸

序 一

弗吉尼亚·伍尔夫曾写过这样一个角色，她的头脑“从深处不断喷吐出各种场景、名字、说法、记忆和想法，就仿佛一座喷泉”。乔治·博克斯就是这样一种活跃思维的化身。与乔治一起吃晚饭，也是见证各种有关他的工作和朋友的故事、诗歌、歌曲和轶事的喷涌而出。本书就将带你跃入那座思想喷泉。它读来非常有趣，哪怕关于统计学和科学的书通常并不在你的阅读范围之内。

毫无疑问，很多读者会认为（就像我曾认为的）这个主题又难又无趣。所幸在这里，我们有一位充满魅力的讲故事的人，他会在一次讨论方差分析时引用约吉·贝拉，会让墨菲（墨菲定律的那个墨菲）在质量有提高余地时提醒我们。其中还有种种故事，涉及组合设计、时间序列预测、调优操作、干预分析法等统计学方法。但它们并不是数学化的，绝大多数都包含与这些方法和应用相关的人士的个人轶事。所以 you 不仅可以在此学到统计学和科学知识，你也是同时在阅读一部统计学家和科学家的名人录（其中还有英国女王哦）。

我第一次遇见乔治·博克斯是在 1968 年，在“周一晚啤酒讨论会”上。这是一个长期在乔治家地下室举行的非正式讨论会。当时我正在上比尔·亨特教授的非线性模型建模课程。比尔建议我到

讨论会上来讨论一下我们当时正在做的某个研究。一想到要与著名的博克斯教授讨论一个建模问题，我不免有些不安。但比尔说，不用担心，乔治很喜欢工程师。比尔和其他几位讨论会参与者都是化学工程师，并且乔治早期与奥拉夫·霍根（当时是威斯康星大学的化学工程学教授）的合作对于当时刚起步的统计系起到了很好的示范和推动作用。所以我鼓起勇气，在一个晚上来到了讨论会，然后坐在后排，犹豫着要不要大着胆子去取一瓶啤酒（本地产的福尔巴赫牌啤酒，这对做统计学来说是个恰当的选择，因为这个牌子的啤酒没有两瓶的味道是一样的）。在后来的将近三十年时间里，我又许多次参加了这个讨论会；而在这些年里，数以百计的讨论会参与者见证了乔治如何为任何想要使用统计学解决问题的人提供指导、建议和鼓励。没有什么问题是太小的，也没有什么是太难的。乔治所给出的始终是友善和有益的建议，而从来不会是打击。一周又一周，我们见识了迭代实验和科学发现相互促进的循环。我们目睹了“所有模型都是错误的，但有些是有用的”一说的真实例子。我们看到了统计学如何催化和催生新的科学方法，以及科学问题如何反过来又催化和催生新的统计学思想。这是怎样一番享受啊。

我的研究方向是水质工程。在一个晚上，我想讨论一个涉及生化需氧量的问题。乔治问道，能否让他来解释一下这个生化需氧量测试。然后他给出了一个我这辈子听过的最棒的解释。我问他怎么会恰巧知道这些，然后我被告知，在十六岁时，他在一家城市污水处理厂担任助理化学师。在 1939 年，在我出生前一年，年仅十九岁的他发表了一篇关于活性污泥法中的生化需氧量的论文。这种污水处理方法在当时刚出现不久，人们对此理解得尚不深入。乔治的这篇论文可与同时期威斯康星的某些知名工程师在这个领域的研究成果相媲美。在五十五年后，在 20 世纪 90 年代，乔治和我尝试

通过多元非平稳时间序列分析预测活性污泥法的污水处理效果。想象一下，这可是与一位世界知名的统计学家，同时也是最早一批研究这种现在已被广泛使用的污水处理方法的研究者展开合作。

乔治和我还有一项小小的未竟事宜。在威斯康星大学麦迪逊分校的工程学大楼和统计学大楼隔壁是兰德尔营公园，其石拱门入口的一侧矗立着一座六米多高的内战士兵雕像。乔治曾想过，应该给这个士兵佩戴一枚勋章。在 1993 年，我们在伦敦海德公园的一个路边摊上发现了一枚合适的黄铜勋章，但把这枚勋章挂到士兵脖子上的计划一直未能实现。现在我觉得乔治应该保留这枚勋章，以奖励他在回忆录写作上的成就。他当之无愧。（并且我们这两个老伙计也就不用爬到士兵身上了。）

昨天（2012 年 5 月 10 日）晚上，我的妻子和我与乔治和克莱尔在肖尔伍德希尔斯的乔治住所一起用晚餐。乔治说：“回忆录已经完成了。”我问道：“你的下一个项目是什么？很难想象你会每天不写点东西。”他答道：“我正在酝酿一篇关于费希尔论多因子实验设计的论文。”

回忆录已经完结，但显然乔治并没有。而这是个好消息。谢谢你，乔治。

P. 马克·贝尔图

（威斯康星大学土木与环境工程系荣退教授）

序 二

很高兴能见到这部自传的面世，它让人有机会了解这位非凡的学者和绅士：乔治·爱德华·佩勒姆·博克斯。我与乔治·博克斯漫长而活跃的统计生涯的交错始于 1952 年夏，当时我正在马里兰州阿伯丁试验场为弗兰克·格拉布斯博士担任研究员。格拉布斯博士和我一同拜读了乔治·博克斯关于响应面法的精彩论文。^①我还记得当时我们如何深深折服于其简洁……当然，是在传达旨意后的简洁。原创性和深入浅出是博克斯之天才的表现。在 9 月回到罗利的统计学研究所继续研究生学习时，我得知了乔治·博克斯要来罗利做一年访问学者。在 1953 年 1 月，我成为了他的首位博士生。

很高兴能读到乔治的童年生活。他的父亲努力在一个能为有才能的年轻人提供最好机会的社会中营造出一个良好的家庭环境。我们会了解到，乔治如何有幸遇到慧眼识珠的老师，从而得以发现他在写作以及数学上的天赋。他在成年后的首份工作是作为一位助理化学师，收集污水处理数据，正是从这里，他的统计学家生涯开始了。而后来的事情证明了这是怎样一个卓越的职业生涯啊！

① G.E.P. Box and K.B. Wilson, "On the Experimental Attainment of Optimum Conditions," *Jour. Royal Stat. Soc. Series B*, 13 (1951): 1-45.

乔治·博克斯的显著特点之一是，他只是偶尔才单独发表论文。这部回忆录就为我们介绍了他的众多学生和研究员，生动勾勒出他的这些合作者的各具特色的个性。大约十年前，他的一帮学生将他在质量管理、实验设计、统计控制，以及稳健性方面的统计学论文结集成书。^①不论是统计学界的新手，还是专家，他们都会发现这许多论文当中的论述极其精彩。当然，他还合著了许多教材，涉及实验设计^②、时间序列分析^③、贝叶斯统计推断^④以及统计控制^⑤等，他对相关领域的理论和应用的深入浅出的阐释一再被证明是具有原创性和富有启迪的。对于乔治·博克斯来说，首字母缩写“KISS”代表着“Keep It Sophisticatedly Simple”（要深入浅出）。

我发现乔治·博克斯回忆的一段早期统计工作经历尤为吸引人。在第二次世界大战临近结束时，英国政府掌握了一批包含未知致命气体的德军毒气弹。乔治是当时最早确定这些新反应物（实际上是神经毒气）的致命毒性的小组成员之一。这些毒气弹从未被投入战争，但要是当初它们被投放到主要城市，伤亡人数将轻松达到广岛和长崎的伤亡人数。曾工作于生死一线以保护世界免遭更多灾难，这着实令人印象深刻。

① George E.P. Box, *Box on Quality and Discovery: With Design, Control, and Robustness*, eds. by George C. Tiao, Søren Bisgaard, William J. Hill, Daniel Peña and Stephen M. Stigler (2000) John Wiley & Sons.

② G.E.P. Box, William G. Hunter, and J. Stuart Hunter, *Statistics for Experiments* (1978) John Wiley & Sons.

③ G.E.P. Box and Gwilym M. Jenkins, *Time Series Analysis: Forecasting and Control* (1970) Holden Day.

④ G.E.P. Box and George C. Tiao, *Bayesian Inference in Statistical Analysis* (1973) Addison Wesley.

⑤ G.E.P. Box and Alberto Luceño, *Statistical Control: By Feedback and Adjustment* (1997) John Wiley & Sons.

我们还可以了解到那些被乔治·博克斯视为导师的人，听到他早期与那些统计学界鼎鼎大名的人物的合作故事。并且通过这部自传，我们会逐渐意识到博克斯对于现代统计学理论和实践所做出的广大贡献。他的论文和图书帮助扩展了人们对于贝叶斯统计方法和时间序列建模的认知。现在我们认识到，富含信息的数据的生产是一个专门的统计领域，初始猜想经由实验和数据分析，再形成新的猜想，这样的过程将能够帮助推进科学进步。而在统计学的专业讨论之外，我们还能体会到他对于这样一个（或者说一些）能够孕育出新的模型和非同寻常的猜想的头脑的深深欣赏。看上去乔治·博克斯（George Box）对于所有有问题要解决的人建议是，要跳出盒子（box）思考。

J. 斯图尔特·亨特

（普林斯顿大学工程与应用科学学院荣退教授）

前 言

有个故事说，一个个子很高的人和他的四岁儿子去门口捡当天的报纸。父亲突然意识到儿子难以跟上自己的步伐。他说：“抱歉，汤米，是不是我走得太快了？”小男孩则说：“不，爸爸，是我走得太快了。”

这个内容可以从两个角度来看：作为一个有趣的故事或笑话，或者作为对于科学发现本质的很好揭示。小男孩对于当时情况的判断是正确的，但不是显见的。父亲的判断是显见的，却是错误的。

所以科学洞见和幽默在这里不谋而合或许不是巧合。好的科学是一种机智，是认出了大自然在我们身上所开的玩笑。

在我九十三岁时，我还是能从过去举出不少这样的例子的。



1895 年左右的博克斯家族（从左往右顺时针）：伯蒂大伯、我的祖父和祖母、
我的父亲、黛西姑姑、佩勒姆二伯，以及莉娜姑姑

致 谢

在写作这部回忆录的大部分时间里，我都病得很重。这就把一个重担搁在了我的妻子克莱尔肩上，而她自己在这个时期的大部分时间里身体状况也不好。要是没有她的无私帮助，这部回忆录将永远不可能完成。而这样的帮助完全有赖于克莱尔聪明务实、训练有素的头脑。每当有危机出现时，她不仅知道如何应对，更是以完全理解和专业的方式加以完成。

此外，我尤其感谢以下人士：博瓦斯·亚伯拉罕、戴维·培根、福特·巴伦坦、埃内斯托·巴里奥斯、马克·贝尔图、休·贝尔图、克莱尔·博克斯、琼·博克斯、海伦·博克斯、哈里·博克斯、罗宾·查普曼、诺曼·德雷珀、康拉德·冯·拉里·霍、玛格丽特·霍姆伍德、朱迪·亨特、斯图·亨特、布赖恩·乔伊纳、蒂姆·克雷默、凯文·利特尔、阿尔韦托·卢塞尼奥、梅尔韦·马勒、维贾伊·奈尔、拉尔斯-艾瑞克·奥列尔、朱迪·帕格尔、丹尼尔·佩纳、何塞·拉米雷斯、玛丽安·罗斯、哈维尔·托尔特，以及约翰·索尔韦·蒂瑟达尔。

感谢布伦特·尼卡斯特罗允许我在书中使用一些照片。

最后还要感谢我的朋友朱迪思·艾伦。^①

① 复旦大学管理学院统计学系黄达博士仔细审读了中文翻译，并提出了宝贵意见，特此致谢。——编者注

出版者的话

如果你还不是很了解乔治·爱德华·佩勒姆·博克斯，那么简单说，他是统计学界的一位泰山北斗。他是一位自学成才的统计学家，凭借其人生经历和统计学知识为许多领域，尤其是过程改进领域做出了独特贡献。此外，他还是一个很好的人。他几乎（如果不是根本的话）不需要加以介绍。他的光临就是我们的荣幸。

本书在乔治九十四岁生日那年出版，是他关于自己的人生、朋友和贡献的回忆录。我们知道，在他的学术生涯中，他撰写了超过两百篇论文；他单在威利出版公司就出版了十二部图书（参见后面的列表），这些书在全世界销售超过二十五万册；他还帮助创办了《技术计量学》期刊，后者是美国质量控制学会和美国统计协会的联合出版物。

我们也通过切身经验知道，乔治是一位真正的绅士、一名慈爱的父亲和一位忠诚的丈夫。他深刻影响了每一位他接触过的人，不论是初出茅庐的统计学家，还是经验老到的总编辑。当你看到他嘴角上翘时，你知道他即将说出一番妙语，其中夹杂着智慧、建议，以及不可或缺地，一个笑话。他经常能够达成他一开始所立下的目标，而不需要鲜花和掌声的刺激。他很有可能是一位研究生或总编辑有幸遇到的最令人难忘的人物。

有两个人帮助乔治完成了这本书。一位是相伴他二十七年的爱妻克莱尔·博克斯，另一位是他的朋友兼研究助理朱迪思·艾伦。

威利出版公司的所有管理层和工作人员称颂博克斯博士对统计学所做出的贡献。我们祝愿他身体康健，福寿绵长。他将始终是名垂青史和我们心中的泰山北斗。

乔治·博克斯在威利出版公司所出的图书

<i>Evolutionary Operation: A Statistical Method for Process Improvement</i> (《演化式操作：一种过程改进的统计学方法》)	Box-Draper, 1969
<i>Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building</i> (《实验员的统计学》)	Box-Hunter-Hunter, 1978
<i>Empirical Model-Building and Response Surfaces</i> (《经验建模和响应面法》)	Box-Draper, 1987
<i>Bayesian Inference in Statistical Analysis</i> (《统计分析中的贝叶斯推断》)	Box-Tiao, 1992
<i>Statistical Control: By Monitoring and Feedback Adjustment</i> (《统计控制：通过监控和反馈调整》)	Box-Luceño, 1997
<i>Evolutionary Operation: A Statistical Method for Process Improvement</i> (《演化式操作：一种过程改进的统计学方法》)	Box-Draper, 1998
<i>Statistics for Experimenters: Design, Innovation and Discovery, 2nd Edition</i> (《实验员的统计学（第二版）》)	Box-Hunter-Hunter, 2005
<i>Improving Almost Everything: Ideas and Essays, Revised Edition</i> (《改进几乎所有一切》)	Box-Friends, 2006
<i>Response Surfaces, Mixtures, and Ridge Analysis, 2nd Edition</i> (《响应面法、混料实验和岭分析（第二版）》)	Box-Draper, 2007
<i>Time Series Analysis: Forecasting and Control, 4th Edition</i> (《时间序列分析：预测与控制（第四版）》)	Box-Jenkins-Reinsel, 2008
<i>Statistical Control by Monitoring and Adjustment, 2nd Edition</i> (《统计控制：通过监控和反馈调整（第二版）》)	Box-Luceño-Paniagua-Quñones, 2009
<i>An Accidental Statistician: The Life and Memories of George E.P. Box</i> (《统计学大师之路：乔治·博克斯回忆录》)	Box, 2013