

(原书第 6 版)

商务与经济统计

Statistics for Business and Economics (6th Edition)



附光盘

(英) 保罗·纽博尔德 (Paul Newbold)
诺丁汉大学
(美) 威廉 L. 卡尔森 (William L. Carlson) 著
圣奥拉夫学院
(美) 贝蒂 M. 索恩 (Betty M. Thorne)
斯泰森大学
庄新田 王世权 刘汝萍 等译



机械工业出版社
China Machine Press

F-7

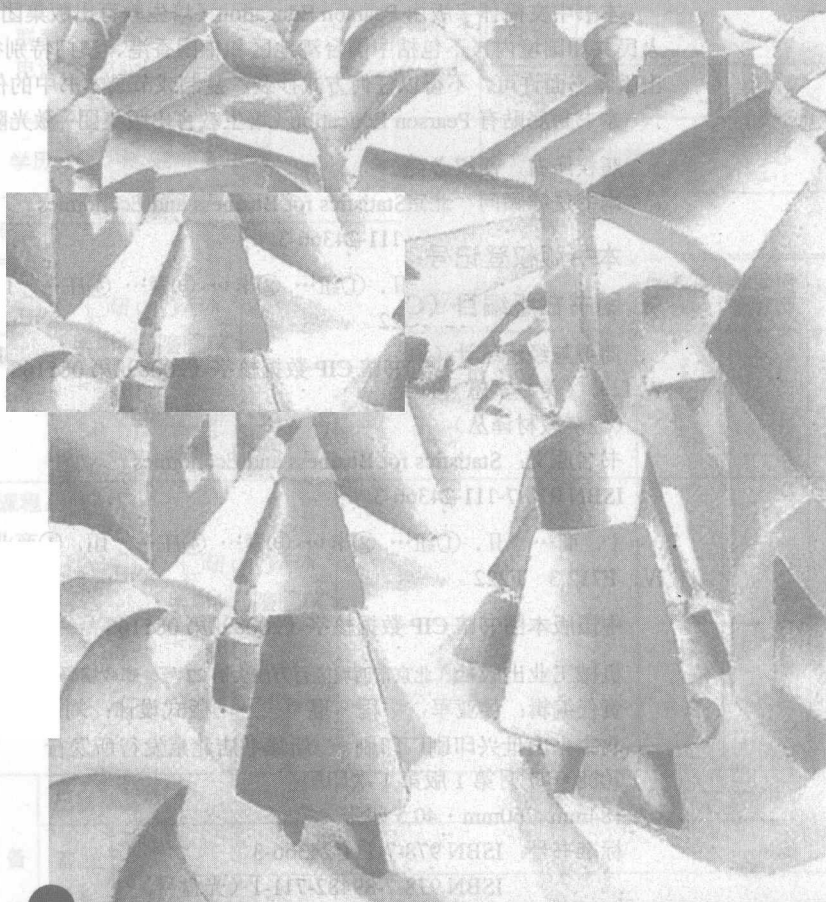
经

F712.3
12

(原书第 6 版)

商务与经济统计

Statistics for Business and Economics (6th Edition)



机械工业出版社
China Machine Press

(英) 保罗·纽博尔德 (Paul Newbold)
诺丁汉大学
(美) 威廉 L. 卡尔森 (William L. Carlson) 著
圣奥拉夫学院
(美) 贝蒂 M. 索恩 (Betty M. Thorne)
斯泰森大学
庄新田 王世权 刘汝萍 等译

本书是一本统计学入门教科书,以收集数据、整理数据、分析数据、根据数据所提供的信息进行决策为主线,通过对应用统计方法进行入门性的介绍,使大家能够在许多商务和经济领域中进行正确的统计分析。本书强化读者对概念的理解,强调用商务和经济领域中的实例来阐述统计学的基本概念与方法,并将统计知识与大量的实例和练习结合在一起,论述如何理解各种统计方法及其假设,从而更好地解决商务和经济领域中的问题。

读者对象:经济管理类专业本科生、研究生、MBA 学生和 EMBA 学生等。

Paul Newbold, William L. Carlson, Betty M. Thorne. Statistics for Business and Economics, 6th Edition.

ISBN 0-13-188090-X

Copyright © 2007 by Pearson Education, Inc.

Simplified Chinese Edition Copyright © 2008 by China Machine Press.

Published by arrangement with the original publisher, Prentice-hall, Inc., a Pearson Education company. This edition is authorized for sale and distribution in the People's Republic of China exclusively (except Taiwan, Hong Kong SAR and Macao SAR).

All rights reserved.

本书中文简体字版由 Pearson Education (培生教育出版集团)授权机械工业出版社在中华人民共和国境内(不包括中国台湾地区和香港、澳门特别行政区)独家出版发行。未经出版者书面许可,不得以任何方式抄袭、复制或节录本书中的任何部分。

本书封底贴有 Pearson Education (培生教育出版集团)激光防伪标签,无标签者不得销售。

版权所有,侵权必究

本书法律顾问 北京市展达律师事务所

本书版权登记号:图字:01-2007-2282

图书在版编目(CIP)数据

商务与经济统计(原书第6版)/(英)纽博尔德(Newbold, P.), (美)卡尔森(Carlson, W. L.), (美)索恩(Thorne, B. M.)著;庄新田等译. —北京:机械工业出版社, 2008.7
(经济教材译丛)

书名原文: Statistics for Business and Economics

ISBN 978-7-111-24366-3

I. 商… II. ①纽… ②卡… ③索… ④庄… III. ①商业统计-教材 ②经济统计-教材
IV. F712.3 F222

中国版本图书馆 CIP 数据核字(2008)第 085166 号

机械工业出版社(北京市西城区百万庄大街 22 号 邮政编码 100037)

责任编辑:吴亚军 程琨 版式设计:刘永青

北京牛山世兴印刷厂印刷 · 新华书店北京发行所发行

2008 年 7 月第 1 版第 1 次印刷

184mm×260mm · 40.5 印张

标准书号: ISBN 978-7-111-24366-3

ISBN 978-7-89482-711-1 (光盘号)

定价: 90.00 元(附光盘)

凡购本书,如有倒页、缺页、脱页,由本社发行部调换

本社购书热线:(010) 68326294

投稿热线:(010) 88379007

中国机械工业出版社

新 告

谨以此书献给我的妻子夏洛特 (Charlotte)，我们的孩子安德烈亚 (Andrea)、道格拉斯 (Douglas)、拉里 (Larry)，以及我们的孙子埃兹拉 (Ezra)、萨万娜 (Savanah)、海伦娜 (Hellana)、安娜 (Anna)、伊娃·罗斯 (Eva Rose) 和艾米莉 (Emily)。

——威廉 L. 卡尔森

谨以此书献给我的丈夫吉姆 (Jim)，以及我们家的珍妮 (Jennie)、安 (Ann)、勒妮 (Renee)、乔恩 (Jon)、克里斯 (Chris)、乔恩 (Jon)、马里厄斯 (Marius)、米哈埃拉 (Mihaela)、恺撒 (Cezara) 和安达 (Anda)。

——贝蒂 M. 索恩

译者序

统计是获得信息的手段和源泉。统计学作为一门收集、整理和分析统计数据的方法论科学,其目的是探索客观事物内在的规律性,以达到对客观事物的科学认识。因此,统计学在众多的专业、学科领域中都发挥着重要作用。在商务和经济领域,几乎所有的问题都要用到统计。随着信息技术的发展,人们可以更方便、更容易地获得大量商业数据,但是数据本身并没有实际意义,只有对数据进行统计分析,才能为我们的决策提供有用的信息。本书正是基于统计学理论与商务和经济实践的结合,阐述如何应用统计方法从大量数据中提炼出决策所需要的信息。

本书是由美国圣奥拉夫学院的威廉 L. 卡尔森教授和斯泰森大学的贝蒂 M. 索恩教授在英国诺丁汉大学保罗·纽博尔德教授所编写的《商务与经济统计》的基础上改编而成的。保罗·纽博尔德教授、威廉 L. 卡尔森教授和贝蒂 M. 索恩教授都是统计学领域的资深专家,在统计学方面造诣颇深。不仅在教学方面享有盛誉,而且还注重将统计学理论与实践相结合,并取得了可喜的成果。

本书共分为 21 章,涵盖了商务与经济统计的所有内容。全书以收集数据、整理数据、分析数据、根据数据所提供的信息进行决策为主线,分别介绍了收集数据的抽样方法;描述数据的图表和数值方法;分析数据的参数估计、非参数估计、假设检验、拟合优度检验、方差分析等方法;利用数据进行预测的简单线性回归、多元回归和时间序列分析与预测;统计方法在质量管理中的运用;统计决策理论。其主要内容包括:为什么要学习统计学;描述数据的图表方法;描述数据的数值方法;概率;离散型随机变量及其概率分布;连续型随机变量及其概率分布;抽样和抽样分布;单个总体的估计;参数估计的进一步讨论;假设检验;假设检验 II;简单线性回归;多元回归;回归分析的进一步讨论;非参数估计;拟合优度检验和列联表;方差分析;质量简介;时间序列分析与预测;抽样的进一步讨论;统计决策理论。

本书在介绍统计学方法时深入浅出,难度适中,并且结合大量的应用实例来说明统计方法的应用,而且这些实例对那些不太了解商务和经济活动的学生来说一样可以轻松的理解,这样就极大地方便了学生的学习。具体而言,本书的特色主要表现在以下几个方面。

1. 实用性强

本书是针对非统计专业的学生编写的,其重要目标是帮助学生成为更优秀的决策者而不是统计学家。因此,本书在结构和语言上都突出了“实用”二字。作者根据几十年的教学与科研经验,从实用的角度构造了一个在商务和经济活动中较常用的统计方法体系,既系统全面又简明清楚、通俗易懂,还加强了对实际背景的描述。所有的统计学基本概念都是在特定的商务决策环境下,通过引入真实案例、真实数据以及现实中的应用,使学生对相关的概念首先产生一个感性认识,然后再深入介绍相关的原理,这样便可以达到事半功倍的效果。

2. 与计算机相结合

随着计算机的出现和大量统计软件的广泛应用,传统的统计学教学也发生了根本性的变化,尤其是对非统计学专业的学生来说,学习统计学的主要目的就是应用。因此,学习统计学的关键是掌握统计方法的基本思想、应用特点和应用场合。在理解其基本思想之后,可以通过计算机软件进行相应的计算,最后解释计算结果并得出结论。本书正是顺应这种潮流,介绍了统计学软件的应用,书中所有的例题和实例均用 Minitab 软件与 Excel 软件进行计算和分析,以此提高学生运用统计方法分析和解决实际问题的能力。

3. 强调学与练的结合

为了帮助学生更好地理解本书所介绍的内容,在每节后面都设置了基础练习和应用练习,每章后面也设置了练习。基础练习用于帮助学生理解基本概念和原理,应用练习用于帮助学生掌握每种方法的实际应用。很多练习都有生动的现实背景。这种组织安排让人始终带着问题去学习,抱着研究的态度去解题,既掌握了方法又了解了应用背景,学完和做完之后有茅塞顿开之感。为了便于自学,本书在附录 B 中给出了偶数题的答案。

综上所述,作为教科书,本书具有较多的应用层次,不仅可以用于经济管理类专业本科生、研究生、MBA 学生和 EMBA 学生的教学,而且对于那些从事经济管理工作,但需要运用统计方法分析问题的读者来说,也颇有应用价值。可以说,这是一本值得经济、财政、金融、市场、会计、管理等领域的从业者一读的好书。如今历时近一年的时间,这部译著终于与广大读者见面了。值此之际,译者首先要感谢东北大学工商管理学院的领导与同事的支持及鼓励。其次,为了使这本书更加完美地呈现给读者,真实、准确地展现书中所揭示的思想,在翻译过程中我们请教了相关领域的专家,在此一并表示谢忱。最后,在本书翻译过程中机械工业出版社华章公司的吴亚军编辑做了大量的协调工作,并提出了很多有益的建议,也向他表示感谢。

本书是集体努力的结晶,初稿翻译分工如下:刘汝萍翻译了前言、第 1~3 章;王世权和韦福雷翻译了第 4~6 章;杨锦华翻译了第 7 章、第 8 章;王丽敏翻译了第 9~11 章;韦福雷翻译了第 12~14 章;胡彩梅翻译了第 15~17 章;庄新田和胡彩梅翻译了第 18~20 章;董入芳翻译了第 21 章、附录 A、附录 B 以及目录部分。初稿完成后,由王世权进一步对全书各章节进行了多次校阅,最后由庄新田负责全书的统稿。

由于译者水平有限,书中难免有不妥之处,敬请读者批评指正。

庄新田

2008 年 3 月于东北大学

作者简介

威廉 L. 卡尔森 (William L. Carlson)

圣奥拉夫学院荣誉退休的经济学教授。他在该学院执教 31 年，并多次担任系主任以及包括计算机学院院长在内的多项行政职务。他不仅主持过美国政府和明尼苏达大学委托的课题，而且还在很多大学举办过讲座。他被选为美国大学优等生荣誉学会 (Phi Beta Kappa) 的荣誉会员。在到圣奥拉夫学院任教之前，他已经在私人产业和契约方面研究了 10 年。他在密歇根理工大学获得工程学士学位，在伊利诺伊理工学院获得硕士学位，在密歇根大学雷克汉姆 (Rackham) 研究生院获得质量管理博士学位。他的研究包括高速公路安全、统计学教育以及与管理相关的众多领域。在此之前，他已经出版了两本统计学教科书。他曾经带领很多学生团体对不同的国家进行研究。他现在是 Cannon Valley Elder Collegium 的执行董事。他喜爱他的孙子、木工、旅行、读书和在威斯康星州北部的工作。



贝蒂 M. 索恩 (Betty M. Thorne)

作者、研究人员和荣誉教师。她是佛罗里达州斯泰森大学商务管理学院的决策科学教授，并担任本科部的主任。索恩博士曾获斯泰森大学颁发给教学成绩优异教师的 McEniry 奖，该奖是斯泰森大学教师的最高荣誉。她也是斯泰森大学商务管理学院年度突出教师奖和年度教授奖的获得者。她为斯泰森大学在奥地利因斯布鲁克的夏季项目、斯泰森大学法学院、斯泰森大学高级 MBA 项目和斯泰森大学行政



公务项目讲授课程。她在 2004 年和 2005 年被选为斯泰森大学法学院 JD/MBA 项目的最佳教授。她在日内瓦学院获得理科学士学位和文科硕士学位，在印第安纳大学获得博士学位。索恩博士是“如何在学校和商务中更加有效地利用统计”（Making Statistics More Effective in Schools and Business）会议计划委员会的成员，并担任秘书和财务管理员，她每年都与统计学科的同行举行一次会议来讨论统计学的研究和教学问题。她还是决策科学学会、美国质量学会和美国统计学会的成员。索恩博士参与了北美 Fareston 与 Tamoxifen 乳腺癌辅助治疗临床试验（NAFTA）长达 5 年的跟踪研究（<http://www.naftatrial.com>）。

索恩和丈夫吉姆有 4 个孩子。他们经常到世界各地旅游，加入并致力于帮助弱势儿童的国际组织。

前言

本书是作者多年从事统计教学工作的经验总结。在2004年和2005年分别被评为“全国优秀教材”和“全国优秀教材一等奖”。本书在2004年和2005年分别被评为“全国优秀教材”和“全国优秀教材一等奖”。本书在2004年和2005年分别被评为“全国优秀教材”和“全国优秀教材一等奖”。

目标读者

本书是一本统计学入门教科书，强化对概念的理解、强调用商务和经济领域中的实例来阐述问题是如何得到解决的。本书适用于：

- 开设商务统计学课程的管理类专业本科生及 MBA 学生。
- 经济类专业本科生及研究生。
- EMBA 学生。
- 开设商务统计学课程的研究生。

本书的宗旨

本书通过对应用统计方法的入门性介绍，使大家能够在许多商务和经济领域中进行正确的统计分析。我们认为对于专业的统计分析，理解相关的假设是非常必要的。对于许多统计方法而言，利用现代计算机可以很容易地根据数据计算出所需要的结果，因此许多教科书仅仅应用简单的“规则”来使用这些输出结果。我们的方法是将知识与大量的例子和学生练习结合在一起，介绍如何理解各种方法及其假设，从而更好地理解商务和经济领域中的问题。

第 6 版的变化

我们对本书进行了更新和扩展，以更好地满足读者的需求并且更加灵活。本书主要有以下几个方面的重大变化和补充，具体如下：

- (1) 对描述统计学的介绍进行了重新设计。
- (2) 在每一节的应用练习之前增加了基础练习。
- (3) 增加了一些应用练习，这些练习将学生置于实际的商务环境中并且强调计算机的应用。
- (4) 根据读者和评论者的反馈，将单个总体、两个总体的置信区间和假设检验分开讨论，以便于进行比较。
- (5) 对简单线性回归和多元回归方法的介绍进行了修正。

(6) 通过大量的应用练习介绍了如何利用相关数值进行投资组合分析。

(7) 增加了用图表描述数据的新方法。

致学生

本书附带的光盘包含了整本书中问题和练习所需要的所有数据文件、TreePlan 软件和文档。

致谢

我们对下面为本书提供评论和深刻见解的个人表示感谢：

- | | |
|-------------------------------------|----------------------------------|
| C. Patrick Kohrman——宾夕法尼亚州立大学，伯克利分校 | Robert Lemke——森林湖学院 |
| James Thorson——南康涅狄格州立大学 | Michael Gordinier——波士顿大学 |
| Mamnoon Jamil——罗格斯大学，坎顿分校 | Fred Wenstop——挪威管理学院 |
| Zhimin Huang——阿德菲大学 | Sheri Aggarwal——弗吉尼亚大学 |
| Renee Fontenot——得克萨斯大学，玻苾恩分校 | Jorgen Lauridsen——南丹麦大学 |
| Allen Lynch——莫瑟尔大学 | Robert Gillette——肯塔基大学 |
| Bulent Uyar——北艾奥瓦大学 | Peter Boatwright——卡内基梅隆大学 |
| David Hudgins——俄克拉何马大学 | Mark Kamstra——西蒙菲沙大学 |
| Allan Lacayo——戴波罗谷学院 | Albert Madansky——芝加哥大学 |
| J. Morgan Jones——北卡罗来纳大学 | Jeff Russell——芝加哥大学 |
| Eugene Allevato——伍德伯里大学 | Nick Polsen——芝加哥大学 |
| Patricia Odell——布莱恩特大学 | Aaron Smith——弗吉尼亚大学 |
| Jay DeVore——加利福尼亚州立理工大学 | Yu-Chi Cheng——圣母大学 |
| Valerie Bencivenga——得克萨斯大学 | Professor Mohanty——加利福尼亚大学，洛杉矶分校 |
| Myles J. Callan——弗吉尼亚大学 | Ken Alexander——南加利福尼亚大学 |
| Andrew Narwold——圣迭戈大学 | Mendy Fygenon——南加利福尼亚大学 |
| Anthony Smith——卡内基梅隆大学 | Matthew White——斯坦福大学 |
| Peter Baxendale——南加利福尼亚大学 | Stefanos Zenios——斯坦福大学 |
| Steen Anderson——奥胡斯商学院，丹麦 | Lawrence Brown——宾夕法尼亚州立大学 |
| Eric Bentzen——哥本哈根商学院，丹麦 | Abba Krieger——宾夕法尼亚州立大学 |
| Hans Geilnkirchen——伊拉兹马斯大学，荷兰 | Harvey Singer——乔治梅森大学 |
| Peter Reiss——斯坦福大学 | William Hausman——艾奥瓦威廉玛莉大学 |
| David Hudgins——俄克拉何马大学 | Jim Swanson——中央密苏里大学 |
| | C. Barry Pfitzner——伦道夫-梅肯学院 |

我们还要感谢 Annie Puciloski，她对本书进行了仔细的校对。同时，向提供帮助和专门技

术以及 GGS 生产服务的 Sandra Krausman 表示感谢。

我们必须感谢圣奥拉夫学院的行政管理助手 Priscilla Hall 女士，她参与了本书各个部分的工作并对参与本书工作的学生进行管理，这些学生包括：Michael Loop、Holly Malcomson、Erin McMurtry、Kelly Schwinghammer、Catharina Zuber。没有他们的帮助，本书是不可能完成的。

感谢斯泰森大学的 Jim Scheiner、Paul Dascher、Marie Gilotti、Sean A. Thomas、John Tichenor 和 Emma Astrom。同时，对 Jennie Bishop（计算机程序员 II，佛罗里达州沃鲁萨乡村健康部门）表示特别的感谢。

另外，对我们的家庭成员表示特别的感谢，他们在本书漫长的写作过程中给予了极大的支持。威廉 L. 卡尔森向妻子 Charlotte 以及成年的孩子 Andrea、Douglas 和 Larry 表示特别的感谢。贝蒂 M. 索恩向丈夫吉姆和成年的孩子 Jennie Bishop、Ann Thorne、Renee Payne、Jon Thorne 以及 Marius、Mihaela、Cezara 和 Anda Sabou 表示感谢。

作者向原创者保罗·纽博尔德（Paul Newbold）创造的坚实基础和优良传统表示感谢。保罗懂得严格的统计分析及其基础的重要性。他认识到有许多复杂的思想需要介绍，并且他致力于为较难的思想提供清晰的解释，而且这些思想只有在解决实际问题中才会用到。因此，前几版包含了许多例子和学生应用练习。我们继续发扬该传统，撰写了这本满足信息时代未来商务领导者需求的教科书。

如果你对本书有任何建议和修正意见，请通过下面的电子邮箱联系作者：carlsoncharbill@msn.com; bthorne@stetson.edu。

Albert M. Bortolotto — 圣奥拉夫学院
Jeff Russell — 圣奥拉夫学院
Nick Polson — 圣奥拉夫学院
Aaron Smith — 弗吉尼亚大学
Yu-Chi Cheng — 圣母大学
Professor Mohanty — 加利福尼亚大学
Ken Alexander — 南加利福尼亚大学
Mendy Eysenck — 南加利福尼亚大学
Matthew White — 斯坦福大学
Stefanos Xenios — 斯坦福大学
Lawrence Brown — 南加利福尼亚大学
Abba Krieger — 南加利福尼亚大学
Harvey Singer — 圣奥拉夫学院
William Hrusman — 艾奥瓦大学
Jim Swanson — 中央密西根大学
C. Barry Pinner — 圣奥拉夫学院

Bulent Uyar — 北卡罗来纳大学
David Hudgins — 北卡罗来纳大学
Allan Lacey — 戴维逊学院
J. Morgan Jones — 北卡罗来纳大学
Eugene Allevato — 北卡罗来纳大学
Patricia Odell — 北卡罗来纳大学
Jay DeVore — 加利福尼亚州立理工大学
Valerie Bencivenga — 圣奥拉夫学院
Myles J. Callan — 弗吉尼亚大学
Andrew Newbold — 圣奥拉夫学院
Anthony Smith — 圣奥拉夫学院
Peter Baxendale — 南加利福尼亚大学
Steen Anderson — 南加利福尼亚大学
Eric Bentzen — 圣奥拉夫学院
Hans Geilker — 圣奥拉夫学院

Peter Reiss — 圣奥拉夫学院
David Hudgins — 北卡罗来纳大学

我们还要感谢 Annie Puciloski，她对本书进行了仔细的校对。同时，向提供照片和资料的作者表示感谢。

目 录

译者序	4.41
作者简介	4.42
前言	4.41
第1章 为什么要学习统计学	1
1.1 不确定环境下的决策	1
1.2 抽样	2
1.3 描述统计学和推断统计学	3
第2章 描述数据：图表	6
2.1 变量的类型	6
2.2 描述分类变量的图表	9
2.3 描述时间序列数据的图表	13
2.4 描述数值型变量的图表	16
2.5 描述变量关系的图表	21
2.6 数据表述错误	26
第3章 描述数据：数值	33
3.1 集中趋势的测量	33
3.2 离散趋势的测量值	37
3.3 加权平均数和分组数据的测量	43
3.4 变量之间关系的测量值	46
3.5 线性关系的获取	51
第4章 概率	57
4.1 随机试验、结果、事件	57
4.2 概率及其基本原理	63
4.3 概率法则	70
4.4 二元概率	78
4.5 贝叶斯定理	86
第5章 离散型随机变量及其概率分布	98
5.1 随机变量	98

5.2 离散型随机变量的概率分布	100
5.3 离散型随机变量的性质	102
5.4 二项分布	108
5.5 超几何分布	114
5.6 泊松概率分布	116
5.7 联合分布离散型随机变量	120
第6章 连续型随机变量及其概率分布	134
6.1 连续型随机变量	134
6.2 连续型随机变量的期望值	138
6.3 正态分布	140
6.4 用正态分布对二项分布进行近似	149
6.5 指数分布	152
6.6 连续型随机变量的联合分布	154
第7章 抽样和抽样分布	164
7.1 从总体中抽样	164
7.2 样本均值的抽样分布	167
7.3 样本比例的抽样分布	179
7.4 样本方差的抽样分布	182
第8章 估计：单个总体的估计	192
8.1 点估计量的性质	192
8.2 均值的置信区间：总体方差已知	197
8.3 均值的置信区间：总体方差未知	201
8.4 总体比例的置信区间（大样本）	205
第9章 参数估计的进一步讨论	211
9.1 两个正态总体均值差异的置信区间	211

9.2	两个正态分布总体均值差异的 置信区间: 总体方差未知	215	13.6	预测	347
9.3	两个总体比例差异的置信区间 (大样本)	219	13.7	非线性模型的转化	349
9.4	正态分布总体方差的置信区间	220	13.8	回归模型的虚拟变量	355
9.5	样本容量的确定	223	13.9	多元回归分析的应用方法	361
第 10 章	假设检验	230	第 14 章	回归分析的进一步讨论	378
10.1	假设检验的概念	230	14.1	建立模型的方法论	378
10.2	正态分布均值的检验: 总体 方差已知	234	14.2	虚拟变量和试验设计	380
10.3	正态分布均值的检验: 总体 方差未知	242	14.3	因变量的滞后值作为回归量	387
10.4	总体比例的检验 (大样本)	245	14.4	设定的偏差	390
10.5	评价检验的效力	248	14.5	多重共线性	393
第 11 章	假设检验 II	256	14.6	异方差性	395
11.1	两个总体均值差异的检验	257	14.7	误差的自相关	398
11.2	两总体比例差异的检验 (大样本)	266	第 15 章	非参数估计	411
11.3	一个正态分布方差的检验	268	15.1	符号检验和置信区间	411
11.4	两正态分布总体方差相等的 检验	271	15.2	威尔科克森符号秩检验	417
11.5	假设检验的评论	274	15.3	曼-惠特尼 U 检验	421
第 12 章	简单线性回归	281	15.4	威尔科克森秩和检验	424
12.1	相关性分析	281	15.5	斯皮尔曼秩相关	426
12.2	线性回归模型	284	第 16 章	拟合优度检验和列联表	430
12.3	用最小二乘法求系数估计值	288	16.1	确定概率的拟合优度检验	430
12.4	线性回归方程的解释效力	292	16.2	总体参数未知的拟合优度 检验	434
12.5	统计推断: 假设检验和 置信区间	297	16.3	列联表	437
12.6	预测	304	第 17 章	方差分析	446
12.7	图形分析	308	17.1	几个总体均值的比较	446
第 13 章	多元回归	317	17.2	单因素方差分析	447
13.1	多元回归模型	317	17.3	克鲁斯卡尔-沃利斯检验	455
13.2	系数的估计	323	17.4	双因素方差分析: 每个单元格 包含一个观测值, 随机分组	457
13.3	多元回归方程的解释效力	328	17.5	双因素方差分析: 每个单元格 包含一个以上观测值	464
13.4	单个回归系数的置信区间和 假设检验	333	第 18 章	质量简介	477
13.5	回归系数的检验	342	18.1	质量的重要性	477
			18.2	均值和标准差的控制图	481
			18.3	生产过程的能力	488
			18.4	比例的控制图	491
			18.5	发生次数的控制图	494

第 19 章 时间序列分析与预测	501	20.5 确定样本容量	552
19.1 指数	502	20.6 其他抽样方法	555
19.2 随机性的非参数检验	508		
19.3 时间序列成分	511	第 21 章 统计决策理论	564
19.4 移动平均	513	21.1 不确定性条件下的决策	564
19.5 指数平滑	520	21.2 不包含确定概率的解决方案： 最大最小准则、最小最大遗憾 准则	567
19.6 自回归模型	528	21.3 期望货币化价值、决策树	570
19.7 自回归综合移动平均模型	532	21.4 样本信息：贝叶斯分析与评价	578
第 20 章 抽样的进一步讨论	535	21.5 考虑风险：效用分析	588
20.1 抽样研究的基本步骤	535		
20.2 抽样和非抽样误差	539	附录 A 表格	595
20.3 简单随机抽样	540	附录 B 偶数题答案 [⊖]	635
20.4 分层抽样	544		

⊖ 本书附录 B 偶数题答案在随书所附光盘中。

第1章

为什么要学习统计学

引言

在信息时代，世界充满了数据。报纸文章和电视新闻报道里经常包含诸如“今天道琼斯指数跌了6点”，“上个月消费者价格指数上涨了0.8%”，“上一次民意调查显示总统的支持率在63%左右”，“临床研究表明一种乳腺癌新药对98%的患者没有明显副作用”的陈述。为了更好地把握事件的发展状态，我们需要获取并理解大量的数据。政府、企业和科研机构花费数十亿美元来收集数据。联邦政府通过自己努力和要求企业发布信息等手段来推动这项活动的开展，私人部门也在这项活动中发挥着自己的作用。著名的盖洛普民意调查和尼尔森收视率调查仅仅是市场调研研究冰山的一角。在过去的几年中，数据收集的数量以显著的比例增长。

我们必须搞清所有数据的意思。计算机的发明，为快速处理、总结、分析数据和产生并存储更多的数据提供了条件。计算机将证券报价等数据带到我们的指尖。我们必须恰当地分析和解释这些数据。

1.1 不确定环境下的决策

决策常常是根据不完全信息做出的。例如，尽管可能没有明确的职业目标，但是大学新生在入学之前就得选择专业；在一种新药对患者的副作用、使用后的存活率、复发率等信息未知的情况下，癌症患者可能被要求参加该药的临床试验（见参考文献1）。类似地，决策者在做出商务决策时，他们也不能确定那些从不同方面影响最终结果的因素将来的表现。

制造商在对一份合同进行投标时，并不能完全确定将来的生产总成本和竞标者的标底。尽管如此，制造商必须进行投标。投资者并不能完全确定将来资本市场是繁荣、稳定还是萧条，但是他必须在将来市场走势不确定的情况下决定如何在股票、债券和货币市场工具之间平衡投资组合。

考虑下列陈述：

- “在将来的6个月里，IBM的股票价格将高于现在。”
- “如果联邦预算赤字像预计的那么高，则在今年剩余的时间里利率将继续走高。”
- “一个大学毕业生的年收入要高于没有接受过大学教育的人。”

在这些陈述里都包含了暗示肯定性的语句，但是我们并不能保证这些确定性主张的真实性。尽管分析师预测在接下来的几个月里IBM的股票价格可能会上涨，但是他也不能完全肯定。因此，需要对这些陈述稍做修改，表述如下：

- “在将来的 6 个月里, IBM 的股票价格很可能会高于现在。”
- “如果联邦预算赤字像预计的那么高, 则在今年剩余的时间里利率很可能会继续走高。”
- “一个大学毕业生的年收入很可能高于没有接受过大学教育的人。”

谨慎的措辞非常重要, 用含糊不清的陈述来替代没有根据的精确陈述是不合适的。“可能”、“很可能”到底表示什么? 尤其是在处理可能性或不确定性时, 应该用贴切的语句来表述自己的意见。

练习

基础练习

1. 修改下列陈述来反映不确定性:
 - a. 提高这种产品市场份额的最好机会是针对 18~25 岁的消费群体做广告。
 - b. 如果标底定在这个水平, 将低于竞争对手的标底, 可以安全获得合同。
 - c. 在两个月内, 美国的汽油成本将会上升。
2. 给出一个必须在不确定条件下做出市场决策的例子。
3. 给出一个必须在不确定条件下做出财务决策的例子。

1.2 抽样

在向市场推出一种新产品之前, 制造商会通过市场调查了解该产品的市场需求水平。实际上, 制造商对于所有的潜在购买者(总体)都感兴趣。但是, 总体通常很大, 不便于分析, 同时收集总体的所有信息也不现实或成本很高。甚至在资源充分的情况下, 由于时间的限制, 我们也只能从总体中抽取一个子集(样本)进行研究。

总体和样本

总体(population)是调查者所关心的所有项目的全集。总体的规模 N 可以很大, 甚至可以无穷大。样本(sample)是总体的一个子集, 样本容量为 n 。

总体的例子包括:

- 美国所有的登记选民。
- 你们学校所有的学生。
- 居住在佛罗里达州棕榈滩的所有家庭。
- 在纽约证券交易所交易的所有股票。
- 某保险公司在特定年份所收到的全部医疗保险索赔。
- 某公司所有的应收账款。

我们的最终目的是利用有效的样本数据来描述整个总体。我们需要一个能够代表总体的样本。如何取得该样本呢? 在抽取样本的过程中, 我们必须遵循的一个重要原则就是随机抽取。

随机抽样

简单随机抽样(simple random sampling)是从总体中抽取一个包含 n 个个体的样本的过程。通过这种方法, 能够保证总体中每个个体都是随机抽取的, 每个个体被选中的可能性是相同的, 同时每个容量为 n 的可能样本被抽取的机会也是相同的。这种方法是如此普通以至于人们通常把形容词“简单”去掉, 将抽取的样本称为随机样本(random sample)。

抽样广泛应用于商务领域以及其他学科。人们通过产品样本来判断某生产过程的运行是否正常。应收账款的审计通常也是基于一个样本。在总统竞选年度, 可以通过登记选民的样本来估计选民的选举偏好, 或许可以用某投票点的投票结果来预测哪一位候选人将会获得该州的选

民选票。但是,抽取样本仅仅是达到目的的手段,我们学习统计学并不是为了对样本做出描述,而是为了得出关于总体的结论。统计学就是研究如何利用样本信息对总体的决策做出推断。当然,在推断过程中总会存在一些不确定性。

假设我们想知道美国登记选民的平均年龄。显然,总体的规模太大,我们只能抽取一个随机样本,也许该样本包含 500 个登记选民,然后计算他们的平均年龄。因为该平均年龄是根据样本数据求得的,所以被称为统计量。如果我们能够计算出整个总体的平均年龄,则该结果被称为参数。在本书中,我们将学习利用统计量确定参数的方法。我们必须认识到只要精确的总体参数未知,就总会存在一些不确定因素。

参数和统计量

参数 (parameter) 是总体特有的特征。统计量 (statistic) 是样本特有的特征。

练习

基础练习

4. 给出下列总体的一个参数:
 - a. 居住在佛罗里达州棕榈滩所有家庭的收入。
 - b. 纽约证券交易所所有股票的年收益。
 - c. 某保险公司在特定年份所收到的全部医疗保险索赔的成本。
 - d. 某公司所有的应收账款额。
5. 你们学校通过调查来确定学生每周上网冲浪的平均时间。根据容量为 174 名学生的样本计算出来的平均时间为 6.1 小时。
 - a. 该调查的总体是什么?
 - b. 该调查的样本是什么?
 - c. 该调查的统计量是什么?
 - d. 6.1 小时是参数还是统计量?
6. 某航空公司声称从奥兰多国际机场启程的某条航线的航班晚点概率小于 1%。通过调查 200 个航班的随机样本,发现其中 1.5% 的航班晚点。
 - a. 该问题的总体是什么?
 - b. 该问题的样本是什么?
 - c. 该问题的统计量是什么?
 - d. 1.5% 是参数还是统计量?

1.3 描述统计学和推断统计学

统计思考过程包含从问题定义到做出决策的所有步骤。当确定并定义了一个问题后,我们根据设计的方法收集不同过程所产生的数据,并采用一种或多种统计方法分析这些数据,通过该分析我们可以获取信息。利用特殊的经验、理论、文献和其他统计方法,将信息依次转化成知识。描述统计学和推断统计学都可以将信息转化为用来指导制定更好决策的知识。

描述统计学和推断统计学

描述统计学 (descriptive statistics) 包括用来概述、处理数据并把数据转化为信息的图形和数值方法。推断统计学 (inferential statistics) 为预报、预测、估计等提供了基础,利用这些方法可以将信息转化为知识。

1.3.1 描述数据

在例 1-1 中,我们可以看到某谷物食品生产车间的日常生产模式图。

例 1-1 谷物食品生产 (描述统计学)

小麦谷物食品公司的一名生产主管将员工组成一个小组研究谷物食品的生产过程。在研究的第一个阶段,随机抽取一些袋装食品进行称重,并测量产品的密度。然后,该生产主管想了解与日常产出模式有关的数据,抽取 10 天的生产率数据。绘图描述这些数据,并评论你的发现。

解: 通过图 1-1, 生产主管能够确定产量较低的天数,同时也可以确定产量增加的天数。