



BUSINESS STATISTICS: A FIRST COURSE

商务统计学

Business Administration Classics
工商管理经典译丛

(第7版) (SEVENTH EDITION)

戴维·莱文 (David M. Levine)
[美] 凯瑟琳·赛贝特 (Kathryn A. Szabat) 著
戴维·斯蒂芬 (David F. Stephan)
岳海燕 胡宾海 等 译
胡大源 审校



中国人民大学出版社

BUSINESS STATISTICS: A FIRST COURSE

商务统计学

(第7版) (SEVENTH EDITION)

Business Administration Classics
工商管理经典译丛

戴维·莱文 (David M. Levine)
[美] 凯瑟琳·赛贝特 (Kathryn A. Szabat) 著
戴维·斯蒂芬 (David F. Stephan)
岳海燕 胡宾海 等 译
胡大源 审校

中国人民大学出版社

· 北 京 ·

图书在版编目 (CIP) 数据

商务统计学：第 7 版/ (美) 戴维·莱文等著；岳海燕等译。—北京：中国人民大学出版社，2017.9
(工商管理经典译丛)

ISBN 978-7-300-24940-7

I. ①商… II. ①戴… III. ①商业统计学-高等学校-教材 IV. ①F712.3

中国版本图书馆 CIP 数据核字 (2017) 第 220238 号

工商管理经典译丛

商务统计学 (第 7 版)

戴维·莱文

[美] 凯瑟琳·赛贝特 著

戴维·斯蒂芬

岳海燕 胡宾海 等 译

胡大源 审校

Shangwu Tongjixue

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

电 话 010-62511242 (总编室)

010-82501766 (邮购部)

010-62515195 (发行公司)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com> (人大教研网)

经 销 新华书店

印 刷 涿州市星河印刷有限公司

规 格 185 mm×260 mm 16 开本

印 张 31.25 插页 2

字 数 719 000

邮政编码 100080

010-62511770 (质管部)

010-62514148 (门市部)

010-62515275 (盗版举报)

版 次 2017 年 9 月第 1 版

印 次 2017 年 9 月第 1 次印刷

定 价 68.00 元

工商管理经典译丛

出版说明

随着中国改革开放的深入发展，中国经济高速增长，为中国企业带来了勃勃生机，也为中国管理人才提供了成长和一显身手的广阔天地。时代呼唤能够在国际市场上搏击的中国企业家，时代呼唤谙熟国际市场规则的职业经理人。中国的工商管理教育事业也迎来了快速发展的良机。中国人民大学出版社正是为了适应这样一种时代的需要，从1997年开始就组织策划“工商管理经典译丛”，这是国内第一套与国际管理教育全面接轨的引进版工商管理类丛书，该套丛书凝聚着100多位管理学专家学者的心血，一经推出，立即受到了国内管理学界和企业界读者们的一致好评和普遍欢迎，并持续畅销数年。全国人大常委会副委员长、国家自然科学基金会管理科学部主任成思危先生，以及全国MBA教育指导委员会的专家们，都对这套丛书给予了很高的评价，认为这套译丛为中国工商管理教育事业做了开创性的工作，为国内管理专业教学首次系统地引进了优秀的范本，并为广大管理专业教师提高教材甄选和编写水平发挥了很大的作用。据统计，本丛书现已成为目前国内管理院校和企业培训中采用率最高、影响最大的引进版教材。其中《人力资源管理》（第六版）获第十二届“中国图书奖”；《管理学》（第四版）获全国优秀畅销书奖。

进入21世纪后，随着经济全球化和信息化的发展，国际MBA教育在课程体系上进行了重大的改革，从20世纪80年代以行为科学为基础，注重营销管理、运营管理、财务管理到战略管理等方面的研究，到开始重视沟通、创业、公共关系和商业伦理等人文类内容，并且增加了基于网络的电子商务、技术管理、业务流程重组和统计学等技术类内容。另外，管理教育的国际化趋势也越来越明显，主要表现在师资的国际化、生源的国际化 and 教材的国际化方面。近年来，随着我国MBA和工商管理教育事业的快速发展，国内管理类引进版教材的品种越来越多，出版和更新的周期也在明显加快。为此，我们这套“工商管理经典译丛”也适时更新版本，增加新的内容，同时还将陆续推出新的系列和配套的案例教材、教学参考书，以顺应国际管理教育发展的大趋势。

本译丛入选的书目，都是世界著名的权威出版机构畅销全球的工商管理教材，被各国和地区的著名大学商学院和管理学院所普遍选用，是国际工商管理教育界最具影响力的教科书。本丛书的作者，皆为管理学界享有盛誉的著名教授，他们的这些教材，经过了美国和世界各地数千所大学和管理学院教学实践的检验，被证明是论述精辟、视野开阔、资料丰富、通俗易懂，又具有生动性、启发性和可操作性的经典之作。本译丛的译者，大多是国内各著名大学的优秀中青年学术骨干，他们大都曾留学欧美，在长期的教学研究和实践中积累了丰富的经验，具有较高的翻译水平。

本丛书的引进和运作过程，从市场调研与选题策划、每本书的推荐与论证、对译者翻译水平的考察与甄选、翻译规程与交稿要求的制定、对译者质量的严格把关和控制，到版式、

封面和插图的设计等各方面,都坚持高水平和高标准的原则,力图奉献给读者一套译文准确、文字流畅、从内容到形式都保持原著风格的工商管理精品图书。

本丛书参考了国际上通行的MBA和工商管理专业核心课程的设置,并充分兼顾了我国管理各专业现行通开课与专业课程设置,以及企业管理培训的要求,故适应面较广,既可用于管理各专业不同层次的教学,又可供各类管理人员培训和自学使用。

为了本丛书的出版,我们成立了由中国人民大学、北京大学、中国社会科学院等单位专家学者组成的编辑委员会,德高望重的袁宝华同志、黄达教授和中国人民大学校长纪宝成教授,都给了我们强有力的支持,使本丛书得以在管理学界和企业界产生较大的影响。许多我国留美学者和国内管理学界著名专家教授,参与了原著的推荐、论证和翻译工作,原我社编辑闻洁女士在这套书的总体策划中付出了很多心血。在此,谨向他们致以崇高的敬意并表示衷心的感谢。

愿这套丛书为我国MBA和工商管理教育事业的发展,为中国企业管理水平的不断提升继续做出应有的贡献。

中国人民大学出版社

网络时代的商务统计学

商务统计学可以解释为在准确调查研究基础上,对可靠的数据进行分析与概括,整理出一套有助于经营管理者在不确定条件下作出正确决策的原理和方法。这一解释注重准确的观察、科学的分析和正确的决策。

多年来我在北京大学为有丰富工作经验的EMBA学员讲授商务统计课程,之所以选用莱文等作者的这本《商务统计学》作为教材,最主要的原因就是这本书的实用性强,并且有大量切合实际的应用案例和习题。2010年,我们在翻译本书第5版时就感到,本书作者一定花了不少时间紧密关注各类媒体报道中的统计调查信息,然后精心概括整理成生动实用的案例或者习题。作者在第4章讲解贝叶斯定理时,详细介绍了垃圾邮件过滤器的工作原理,我们在这几年的教学和研究中不断发现这个案例的精妙之处,深感贝叶斯统计思路对解决网络时代的许多新问题都有重要的借鉴价值。

例如,近年来随着互联网金融的迅速崛起,相关金融风险不断增大。贝叶斯定理可以帮助我们通过大数据分析来识别违约者。首先,我们需要准确可靠的观察记录,在此基础上得到违约者占比,这就是先验概率;然后,通过深入分析观察记录找出违约者与正常商家的不同之处,贝叶斯定理告诉我们如何根据这些不同之处来提高识别违约者的概率,又称为后验概率。通常讲解贝叶斯定理的应用到此为止。然而,本书中关于垃圾邮件过滤器工作原理的案例讨论让我们打开了思路。在现实生活中,违约者不会束手就擒,总会想方设法逃避识别。然而,正是这些试图逃避识别的做法有可能成为贝叶斯统计方法提高识别的精准程度的信息来源。目前贝叶斯统计方法已经越来越多地用于提高机器学习的能力。

对于人工智能或机器学习,应用成效如何往往并不仅仅取决于数据的大小,更重要的是算法是否符合实际。研究社会经济问题与研究自然科学问题的一个不同之处就是要考虑到人的行为变化。从本书关于垃圾邮件过滤器工作原理的案例分析可以得到一个启示:卓有成效的人工智能离不开人们对新情况的观察和分析,在此基础上不断改进算法方能取得成功。这种交互式动态贝叶斯统计方法表明,成功的人工智能背后必定有一群勇于探索并且善于学习的人。

理解商务统计的关键在于结合现实问题进行思考和分析。要学好商务统计,熟悉并透彻理解一些典型案例不失为一条事半功倍的学习途径。案例越生动,越令人信服,就越容易给人留下长久的印象。对这本教材感兴趣的读者,学习商务统计并不局限于对纯理论的兴趣,而是为了系统地了解在不确定条件下作出正确决策的分析思路与方法,以便更好地解决现实

世界中遇到的实际问题。善于触类旁通的决策者可以根据掌握的各种信息与丰富的个人经历迅速进行类比,分析异同,作出判断,制定对策。这正是百年哈佛商学院案例教学法的精妙之处。

与众多统计学著作相比,本书更适合那些注重学以致用读者的需求。根据本书第7版前言中作者阐述的教学准则,可将本书的特色概况如下:

(1) 帮助读者了解统计学对改进经营管理和不断创新的重要意义。为此,作者花了大量时间参加统计应用学术活动,密切关注传统媒体和新媒体中的有关报道,广泛收集当今现实生活中真实生动的案例,这些案例涉及会计、金融、管理、营销等公司经营管理的各个领域。这本商务统计教材中的数学推导很少,作者热切希望书中大量课堂教学实例和各章后面的案例及习题能够帮助读者理解和掌握如何运用统计学原理和方法来解决现实世界的经营管理问题。

(2) 对于大多数读者,与了解计算过程相比,更重要的是对统计结果的理解。教师在课堂教学过程中讲述计算过程的目的之一就是帮助学生了解在统计应用过程中怎样根据实际情况来解释统计计算结果。统计学家乔治·伯克斯的名言“所有模型都是错误的,但有些是有用的”,讲的就是数学模型都是在特定假设下建立的。而现实问题通常难以完全满足模型的假设条件,这就需要研究人员根据实际情况权衡主次、作出判断。正确的结论往往并不只是根据统计分析结果得出,而是要在正确调查研究的基础上深入探索数据背后的规律。统计研究得出的规律还要经得起实践的检验和长时间的验证。

(3) 为了便于读者分析实例,本书在相应网站上提供了大量电子表格形式的数据集,在讲解教学案例时提供了一些简便易学的统计软件的计算结果(如 Excel 和 Minitab)。本书在过去十年中数次修订的另一个主要原因就是为了跟上计算机软件和互联网技术发展的时代步伐。书中不断更新的教学案例在帮助读者理解统计学原理的同时,还有助于提高统计应用软件的使用能力。

随着信息化管理专业程度的不断提高,在管理决策过程中广泛存在一种倾向,即过于依赖软件公司各类频繁升级的管理信息系统,而忽视经营管理者对于数据收集和数据分析能力的训练和培养,结果往往使企业花费巨资请咨询公司帮助建立的管理信息系统难以取得预期的成效。管理决策者常常没有意识到,经营管理中面临的许多实际问题用 Excel 这类简单易学的统计软件就可以分析解决。在充满不确定性的网络时代,管理者的调查研究能力和基于数据分析的决策能力远比管理信息系统的升级重要。

初次阅读本书的读者可能会觉得这本教材的篇幅太大。对于有教师指导的学生,这也许算不上什么缺点,任课教师自会选择要讲授的章节,其他部分充其量只是作为参考,供有兴趣的读者翻看。这种编写教材的方式在国内外的大学和研究生教材中已司空见惯。因此,对于自学为主的读者,可以把这本书作为常用商务统计方法的工具书,先大致了解各章的主要内容,然后针对自己需要解决的问题来深入学习。事实上,作者由浅入深、循序渐进的表述方法是很适合自学的。此外,书中大量应用案例不但增强了这本教材的可读性,而且有助于读者从不同的角度更好地了解当今国际商务环境的新变化。

这本《商务统计学》(第7版)是著名国际教育出版商培生公司2016年出版的最新版本。作者在第1章前增加了一个全新的部分——“首先要学的重要内容”,概括介绍网络时代数据的新特征以及统计思维方式和分析方法的新变化,以帮助读者更有针对性、更有效率地学习。

本书“首先要学的重要内容”和第1~3章由岳海燕翻译,第4~8章由胡宾海翻译,第

9~11章由冯婧翻译，第12~13章和附录及练习题答案*由沈威宇翻译。我对全文进行了审校和统一调整。对原文理解不当和翻译错误之处主要应由我负责，欢迎读者批评指正。

这本教材的翻译也是一次默契的团队合作过程，在此我们要特别感谢中国人民大学出版社的编辑，他们在翻译过程中为我们提供了许多帮助，并为本书的出版做了大量细致具体的工作。

胡大源

* 练习题答案请到中国人民大学出版社工商管理分社的网站 (www.rdjg.com.cn) 下载。

首先要学的重要内容 // 001

- 1 统计学是一种思维方式 // 001
- 2 数据：应该如何定义 // 002
- 3 统计学正在改变面貌 // 004
- 4 统计学：商学教育中的重要组成部分 // 005
- 参考文献 // 006 关键词 // 006

第 1 章 数据定义与收集 // 007

- 1.1 定义变量 // 008
- 1.2 收集数据 // 010
- 1.3 抽样方法的类型 // 013
- 1.4 调查误差的类型 // 017

小 结 // 021

- 参考文献 // 021 关键词 // 021
- 检查你对本章的理解 // 022 本章复习题 // 022
- 第 1 章案例 // 024

第 2 章 整理并使变量可视化 // 027

- 2.1 整理属性变量 // 028
- 2.2 整理数值数据 // 033
- 2.3 属性变量的可视化 // 041
- 2.4 数值变量的可视化 // 049
- 2.5 两个数值变量的可视化 // 058
- 2.6 一组变量的整理和可视化 // 062
- 2.7 变量整理与可视化中的挑战 // 066

小 结 // 073

- 参考文献 // 074 关键词 // 074
- 检查你对本章的理解 // 075 本章复习题 // 075 报告写作练习 // 082
- 第 2 章案例 // 082

第 3 章 数值描述度量 // 085

- 3.1 集中趋势 // 086

- 3.2 变异程度和分布形状 // 089
- 3.3 探索数值数据 // 101
- 3.4 总体的数值描述度量 // 109
- 3.5 协方差与相关系数 // 113
- 3.6 描述性统计量: 缺陷和道德问题 // 119

小结 // 119

参考文献 // 120 关键词 // 120

检查你对本章的理解 // 121 本章复习题 // 121 报告写作练习 // 125

第3章案例 // 125

第4章 概率论基础 // 127

4.1 概率论基本概念 // 128

4.2 条件概率 // 138

4.3 贝叶斯定理 // 147

4.4 计数规则 // 151

4.5 道德问题与概率 // 155

小结 // 156

参考文献 // 156 关键词 // 156

检查你对本章的理解 // 157 本章复习题 // 157

第4章案例 // 159

第5章 离散概率分布 // 161

5.1 离散变量的概率分布 // 161

5.2 二项分布 // 166

5.3 泊松分布 // 173

小结 // 178

参考文献 // 178 关键词 // 178

检查你对本章的理解 // 178 本章复习题 // 178

第5章案例 // 181

第6章 正态分布 // 183

6.1 连续概率分布 // 183

6.2 正态分布 // 184

6.3 评估正态性 // 198

小结 // 204

参考文献 // 204 关键词 // 205

检查你对本章的理解 // 205 本章复习题 // 205

第6章案例 // 207

第7章 抽样分布 // 209

7.1 抽样分布 // 209

7.2	均值的抽样分布	//	210	
7.3	比率的抽样分布	//	222	
	小 结	//	225	
	参考文献	//	226	关键词 // 226
	检查你对本章的理解	//	226	本章复习题 // 226
	第7章案例	//	228	
第8章	置信区间估计	//	230	
8.1	对总体均值的置信区间估计 (σ 已知)	//	231	
8.2	对总体均值的置信区间估计 (σ 未知)	//	237	
8.3	比率的置信区间估计	//	247	
8.4	确定样本容量	//	250	
8.5	置信区间估计与道德问题	//	256	
8.6	自助抽样法	//	256	
	小 结	//	257	
	参考文献	//	257	关键词 // 257
	检查你对本章的理解	//	258	本章复习题 // 258 报告写作练习 // 262
	第8章案例	//	262	
第9章	假设检验基础：单样本检验	//	265	
9.1	基本假设检验方法	//	266	
9.2	对均值的 t 检验 (σ 未知)	//	278	
9.3	单侧检验	//	285	
9.4	对比率假设的 Z 检验	//	290	
9.5	假设检验潜在缺陷与职业道德事项	//	294	
	小 结	//	296	
	参考文献	//	296	关键词 // 296
	检查你对本章的理解	//	297	本章复习题 // 297 报告写作练习 // 300
	第9章案例	//	300	
第10章	两样本检验与单因素方差分析	//	301	
10.1	两个独立总体的均值比较	//	301	
10.2	两个相关样本的均值比较	//	311	
10.3	两个独立总体比率差异的比较	//	319	
10.4	两个总体方差差异的 F 检验	//	325	
10.5	单因素方差分析	//	330	
10.6	影响大小	//	345	
	小 结	//	345	
	参考文献	//	347	关键词 // 347
	检查你对本章的理解	//	347	本章复习题 // 348 报告写作练习 // 350
	第10章案例	//	350	

第 11 章 χ^2 检验 // 354

11.1 两个比率差异的 χ^2 检验 // 354

11.2 多个比率差异的 χ^2 检验 // 362

11.3 独立性的 χ^2 检验 // 367

小 结 // 374

参考文献 // 374 关键词 // 375

检查你对本章的理解 // 375 本章复习题 // 375

第 11 章案例 // 377

第 12 章 简单线性回归 // 380

12.1 回归模型的类型 // 381

12.2 建立简单线性回归方程 // 383

12.3 离差的度量 // 391

12.4 假设条件 // 396

12.5 残差分析 // 396

12.6 度量自相关性: Durbin-Watson 检验 // 401

12.7 对斜率的统计推断与相关系数 // 406

12.8 均值估计和单个数值预测 // 413

12.9 回归分析的潜在缺陷 // 417

小 结 // 420

参考文献 // 420 关键词 // 420

检查你对本章的理解 // 422 本章复习题 // 422 报告写作练习 // 426

第 12 章案例 // 427

第 13 章 多元回归 // 428

13.1 建立多元回归模型 // 428

13.2 调整的可决系数与整体显著性检验 // 434

13.3 多元回归模型的残差分析 // 438

13.4 总体回归系数的统计推断 // 441

13.5 虚拟变量和交互项在回归模型中的应用 // 445

小 结 // 451

参考文献 // 453 关键词 // 453

检查你对本章的理解 // 453 本章复习题 // 453

第 13 章案例 // 457

附 录 // 459

首先要学的重要内容

■ Important Things to Learn First

学习目标

- 无处不在的数据使学习统计学变得至关重要
- 统计学是一种有助于改善决策的思维方式
- 怎样使用 DCOVA 框架来解决商务问题
- 商务分析学的重要性
- 商务分析学给学生带来的机会

统计应用

“你不能逃避数据”

每天你都听到数据这个词，你可能知道数据就是关于这个世界的事实。你可能以为数据就是数字，比如民意测验结果显示 45% 的被访者认为经济会在明年有所改善。但数据不仅仅是数值化的事实，例如，你每次访问网络搜索引擎、发送或收取电子邮件或手机短信或者在某个社交网站上发帖的时候，都在生成和使用数据。

在这个更为宽泛的数据的概念下，你会注意到很多故事中的人物之所以能够揭发阴谋、预告灾难或者抓获罪犯，往往就是由于他们预先收集了“大量的数据”。你可能听说过对于政府机构通过收集数据实行暗中“监控”的担忧。你可能还听说过一些商业机构如何通过“挖掘”数据来获取利润。你或许已经意识到了在当今世界，你不能逃避数据。

尽管你不能逃避数据，却有可能选择躲开数据。如果你躲开数据，就可能由于盲目地接受其他人对数据的概括总结而上当受骗。（回忆一下那些声称能获得高额回报其实却是完全虚假的金融骗局。）如果你躲开数据，就可能完全依赖“跟着感觉走”来做决策，与在商务管理课程中所学的理性决策过程相比，效率大为降低。当你意识到躲开数据并不可行时，你就会理解有效处理数据是一项重要技能。在验证这项技能的过程中，你就会发现你不能逃避学习统计学，统计学方法可以使你有效地处理数据。

1 统计学是一种思维方式

统计学是关于有效处理数据的方法，这些方法代表了一种可以帮助你更好地做出决策的思维方式。如果你曾经通过制作图表来概括数据，或者通过诸如平均值之类的计算来概括数据，那你就已经运用过统计学。然而，统计学远远不止这些通常讲授的技巧，浏览一下目录你就会有所了解。

以往在初级阶段讲授的统计学很可能强调进行数学运算。相比之下，当今在商务管理中依靠统计软件运算比手算更快、更精准。而且，用软件计算也仅仅构成众多统

计应用任务中的一小部分。想要最好地理解统计学是一种思维方式, 你需要一个框架把统计学的各项任务组织起来。**DCOVA 框架** (DCOVA framework) 就是这样的思维框架。

DCOVA 框架

DCOVA 框架包括以下任务:

- **定义** (define) 为解决某个问题或者实现某个目标而要研究的数据。
- 从适当的来源**收集** (collect) 数据。
- 通过创建表格对收集到的数据进行**整理** (organize)。
- 通过创建图形使收集到的数据更加**可视化** (visualize)。
- **分析** (analyze) 收集到的数据以便得出结论并演示结果。

定义、收集、整理、可视化与分析这几项任务可以帮助你应用统计学做出商务决策。通常需要首先完成前两项任务从而获取有意义的成果, 其他三项任务在实践中的顺序可以改变, 有时甚至可以同时进行。例如, 某些数据可视化方法可以帮助人们在整理数据的过程中同时做出初步的分析判断。

借助 DCOVA 框架有利于在商务活动的以下四个领域中应用统计学方法:

- 概括商务数据并使其可视化;
- 从数据分析中得出结论;
- 对商务活动做出可靠的预测;
- 改进商务管理的运营过程。

在本书中, 尤其是在每章开始的“统计应用”部分, 你会看到 DCOVA 框架帮助你运用统计学的具体例子。例如, 在某一章中, 你会学习如何证明某个营销方案是否增加了产品的销量, 而在另一章里, 你又会学习某个电视台怎样降低不必要的人员费用支出。

2 数据: 应该如何定义

如果像本章开篇那样把数据简单定义为“关于这个世界的事实”或许会令人费解, 因为有的事实具有单一性, 如某些事物的某一数值, 而另一些事实则具有集合性, 如某些事物的一系列数值。进一步举例来说, 本书的作者之一“戴维·莱文”就是单一的事实, 而“戴维、凯瑟琳与 (另一个) 戴维”则是本书的作者的集合。此外, 如果每一件事都是数据, 那么你又该如何区分“戴维·莱文”和“商务统计学” (合著者与书名) 这两个关于本书的完全不同的事实呢? 统计学家为了避免这类困扰, 使用了一个更为确切的数据定义, 这就是变量。

在统计学中, **数据** (data) 是“有助于辨认事物发生的某个特质或者属性的值”。例如, “戴维·莱文”与“凯瑟琳·赛贝特”都是数据, 因为它们都可以帮助读者辨识本书的作者。在本书中, 数据总是用复数形式, 以提醒你数据是数值的整体或集合。虽然数据可以具有单一性, 如“戴维·莱文”就是单一性数据 (datum)。通常表示单一数值时还会用到数据点 (data point)、观测值 (observation)、反应 (response) 以及单一数值 (single data value) 等术语。

变量 (variable) 用来表示与数据数值相关的事物特质或属性。假如你要对本书

的数据集进行定义，就可以把“合著者”和“书名”定义为变量。

本书用特征 (characteristic) 一词取代“特质或属性” (trait or property)，用“某个物体或个人”代替模糊的“某个事物”来定义变量和数据。

变量

就是物体或个人的特征。

数据

就是与变量相关的各个值的集合。

考虑一下把个人从人类总体中辨别出来的特征。姓名、身高、体重、眼睛颜色、婚姻状况、调整后的总收入以及住址都是个人的特征。所有这些特质都有可能用作描述个人的变量。

把本书作者的姓和名定义为变量，很明显有效的值则是“戴维·莱文”“凯瑟琳·赛贝特”与“戴维·斯蒂芬”，而不是“莱文”“赛贝特”与“斯蒂芬”。在此需要注意定义中的文化或其他方面的假定，例如，姓名的后半部分是像北美洲常见的那样指家庭的姓氏，还是很多亚洲国家那样是个人自己特有的名字？

统计学

有了数据的定义，就可以把本书的主题，**统计学** (statistics)，定义为将数据转化为对决策者有用的信息的方法。

统计学可以帮助你确定你的数据能否用来改善决策，因此，统计学有助于确定数值差异是否有显著的实际意义，或者这种差异仅仅是偶然所致。为了形象地说明这一点，请考虑以下有关不同数据的新闻报道：

- “观看免费节目前可接受的广告长度” (USA Today, February 16, 2012, p. 1B)。对 1 179 名 18 岁及以上的成年人的调查显示，54% 的人在观看免费节目前可以接受 15 秒的在线广告时间。

- “关于 Facebook 的六个新事实” (Pew Research Center, bit.ly/IkENZcA, February 3, 2014)。一项调查显示女性比男性更可能把下列选项作为使用 Facebook 的原因：观看图片与视频、同时与许多人分享、看娱乐性与好笑的帖子、学习帮助他人的方法、从朋友圈中获得支持。

- “追随 Tweets” (H. Rui, A. Whinston, and E. Winkler, *The Wall Street Journal*, November 30, 2009, p. R4)。在该研究中，作者发现在推特社交信息服务 (Twitter social messaging service) 评论中某一特定产品被提到的次数可以用来对该产品的销售趋势做出准确的预测。

没有统计学，你就不能确定这些报道中的“数字”是否代表有用的信息。没有统计学，你就不能证明某些判断，例如上面提到的推文 (tweets) 的数量可以用来预测某些产品的销售数量。没有统计学，你就不能看出大量数据有时候显示出来的某些模式。

在统计学中，**统计描述** (descriptive statistics) 主要用来概括和展示数据。在幼儿园的课堂上计数物品的数量可能是你第一次使用统计描述。**统计推断** (inferential statistics) 则利用从小群体收集的数据来得出有关大群体的结论。如果你在低年级的时候正式学习过一些统计学知识，你很可能已经学过统计描述的方法，这是本书前几

章的重点；而你可能并不熟悉本书后面章节讨论的统计推断的方法。

3 统计学正在改变面貌

此前在“统计应用”部分提到的你不能“逃避”数据已经激励人们越来越多地使用过去不存在、难以应用或者鲜为人知的统计方法。你可能在其他课程中学到过这些信息与沟通领域的新方法，它们推动了统计学在商务活动中的广泛应用，并使统计学知识对于取得商务活动成功越来越显示出其至关重要的作用。这就是统计学正在改变的面貌。

商务分析学

将当今使得统计学卓有成效和至关重要的所有变化合在一起就形成了**商务分析学**（business analytics），这是一套最能体现统计学的变化的分析方法。商务分析学将传统的统计方法与管理科学和信息科学方法结合在一起，形成了一套跨学科的分析工具，用来支持以事实为依据的管理决策。商务分析学能够帮助你：

- 应用统计方法分析和探讨数据，找出此前人们无法预料的事物间的关联关系。
- 应用管理科学的方法开发优化模型，改进从战略制定到各个层面的日常管理。
- 使用信息系统的方法来收集和处理不同容量的数据集，包括那些原本难以开展有效研究的容量巨大的数据集。

即便你从未听说过商务分析学这个术语，上述方法的实际应用对你来说也并不陌生。关于政府机构挖掘个人数据来遏制犯罪和恐怖主义的头条新闻，关于大公司如何了解你的秘密，以及那个令人难忘但多少有些夸张的例子——“塔吉特公司（Target）如何知道顾客怀孕了”的传闻，甚至那些关于社交网络或流媒体公司向用户推荐选项或出售定向广告的讨论，都显示了统计学的面貌正在改变。

“大数据”

近年来你不可能逃避的数据又出现了一些新形式，包括人们所熟知的“大数据”。**大数据**（big data）是用传统方法不易浏览或分析的数据集合。

大数据并没有一个准确实用的定义，然而使用这个术语意味着收集到的数据量很大，数据产生的速度很快（通常接近实时），而且数据形式多种多样，并不限于数据处理记录、文件和表格等传统结构化数据。“大量、快速和多样”（volume, velocity, variety）的特征（见参考文献4）有助于把大数据与看似“大量”但只是包含许多重复记录或同样文件结构的数据集区别开来。

大数据为企业获得新的管理理念或者从数据资源中获取价值提供了机会（见参考文献7）。企业可以从统计学尤其是商务分析学的新方法应用中获取这些新理念或价值。

软件在统计学中的整合作用

在第1节中提到在商务管理中用统计软件运算比手算更快、更精准。同样的道理，本书更注重如何解释应用软件得出的统计结果，而不是这些结果的计算过程。本

书使用微软的 Excel 与 Minitab 来得出计算结果，并且展示这些软件是怎样整合在应用统计方法进行商务决策的全过程中的。

Excel 和 Minitab 都使用工作簿存储需要分析的数据。工作簿 (worksheets) 以表格形式排列数据，行和列的交叉点形成**单元格** (cells)，用来输入内容。在 Minitab 和 Excel 中，通常用列来输入变量的数值，每一列为一个变量。在使用上述软件进行统计分析时，可以选择一列或多列数据（一个或多个变量）来执行适当的计算程序。这意味着本书包含的例子和题目均为传统意义上的结构化数据而不是被称为大数据的（非结构化）数据集。无须担心，学习处理结构化数据可以让你掌握统计学原理，日后当然可以应用于分析大数据。

4 统计学：商学教育中的重要组成部分

统计学的面貌正在发生变化，这就意味着统计学已经成为工商管理教育中非常重要的一部分。在当前数据导向的商务环境中，你需要具备常用的分析技能，以便正确处理数据、解释分析结论并把结果融合到各领域的决策过程中，如会计、金融、人力资源管理、市场营销、战略与计划以及供应链管理等商务管理领域。

你的决策过程将越来越依靠数据而不是基于个人经历的直觉或灵感。数据导向已被实践证明是成功的，研究表明那些常用统计分析学帮助决策的公司在生产率、创新和竞争力方面都有所提升。应用数据与数据分析进行商务决策不容小视。均衡地具备多项技能——例如统计学分析方法、建立模型的技巧以及信息技术方面的基本技能，与管理技能——例如把握商务机会的能力、问题解决能力以及沟通技能——将让你为当前或将来步入职场做好充分的准备（见参考文献 1）。

商学院的学生以往仅把统计学当作一门内容与所学专业并不相关的必修课。如果你打开这本书并且有类似的想法，你就忽视了统计学的发展变化。在你学习使用 DCOVA 框架以在第 1 节列出的四类商务活动中应用统计方法的同时，让这本书来帮助你更好地理解这些变化的深刻意义。

使用本书的最好方式

本书使用 DCOVA 框架来组织和展示统计学的内容。要使本书充分发挥作用，首先要确保理解 DCOVA 框架。在理解 DCOVA 框架的基础上，你可以把本书的章节划分成以下几部分：

- 第 1 章：定义 (define) 与收集 (collect) 任务，这是应用统计方法必不可少的初始任务。
- 第 2 章与第 3 章：整理 (organize) 与可视化 (visualize) 任务，用来概括与展示商务数据。
- 第 3 章与第 4~11 章：分析 (analyze) 任务的方法，有助于使用样本数据得出关于总体的结论。
- 第 12 章与第 13 章：分析 (analyze) 任务的方法，有助于做出可靠的预测。

为了更好地学习各章节的内容，请首先阅读各章开头的“统计应用”中的各种情境。每章的情境都描述一种商务情形，其中的问题可以用该章讨论的方法来解决。这些情境是章节中用到的许多讨论统计方法的例子的来源。在每章结尾，“回顾”部分