

统计学

(原书第5版)

Statistics for Engineers and the Sciences

(Fifth Edition)



(美) William Mendenhall 著
Terry Sincich

梁冯珍 关静 等译 史道济 审校



机械工业出版社
China Machine Press

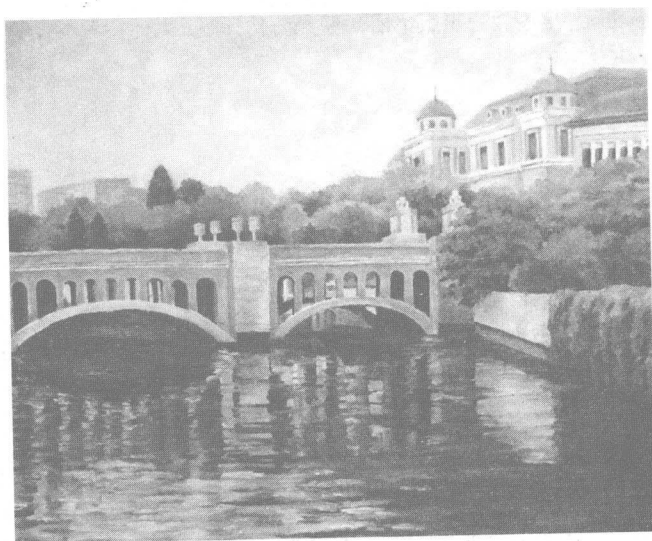
统计学精品译丛

统计学

(原书第5版)

Statistics for Engineers and the Sciences

(Fifth Edition)



(美) William Mendenhall 著
Terry Sincich

梁冯珍 关静 等译 史道济 审校



机械工业出版社
China Machine Press

本书是一本联系实际应用的统计教材. 全书共 17 章, 主要介绍描述性统计、概率、离散随机变量、连续随机变量、二元概率分布及抽样分布、置信区间估计、假设检验、分类数据分析、简单线性回归、多重回归分析、模型构造、试验设计的原则、试验设计的方差分析、非参数统计、统计过程和质量控制、产品和系统的可靠性. 此外, 本书的附录部分还介绍了一些统计软件的使用方法.

本书内容丰富, 很少涉及统计学理论的严格数学证明, 绝大部分是与实际应用紧密联系的例子和练习, 适合作为理工科各专业本科生、研究生的统计学教材, 也可作为相关领域研究人员的参考读物.

Simplified Chinese edition copyright © 2008 by Pearson Education Asia Limited and China Machine Press.

Original English language title: *Statistics for Engineering and the Sciences, Fifth Edition* (ISBN 0-13-187706-2) by William Mendenhall, Terry Sincich, Copyright © 2007.

All rights reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Sun Microsystems, Inc. .

本书封面贴有 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: 01-2006-5286

图书在版编目(CIP)数据

统计学(原书第5版)/(美)门登霍尔(William, M.), (美)辛塞奇(Terry, S.)著; 梁冯珍等译. —北京: 机械工业出版社, 2009. 8

(统计学精品译丛)

书名原文: *Statistics for Engineers and the Sciences, Fifth Edition*

ISBN 978-7-111-26437-8

I. 统… II. ①门… ②辛… ③梁… III. 统计学 IV. C8

中国版本图书馆 CIP 数据核字 (2009) 第 025667 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: 王春华

北京瑞德印刷有限公司印刷

2009 年 10 月第 1 版第 1 次印刷

186mm × 240mm · 51.75 印张

标准书号: ISBN 978-7-111-26437-8

定价: 128.00 元

凡购本书, 如有倒页、脱页、缺页, 由本社发行部调换
本社购书热线: (010) 68326294

前 言

本书概要

本书是为工程专业和自然科学专业的学生设计的、供两个学期使用的统计学课程教材。一旦这些学生毕业，并且找到了工作，他们就将涉及数据的收集和分析，并且需要批判性地思考结果。这就要求他们了解数据描述及统计推断的基本概念，并且熟悉工作中需要用到的统计方法。

教学法

第1~6章介绍学习统计学的目的，说明如何描述数据集，并且给出一些概率论的基本概念。第7章和第8章介绍关于总体参数的两种推断方法：估计和假设检验。这些概念在其余几章中被扩展为在分析工程和科学数据时有用的主题，包括分类数据分析（第9章）、回归分析以及模型构造（第10~12章）、试验设计的方差分析（第13~14章）、非参数统计（第15章）、统计质量控制（第16章）以及产品和系统的可靠性（第17章）。

我们假定读者学习了两个学期的微积分学，即读者对微积分学原理有一个基本了解。要求读者具有运用一元微分法和积分法的能力，但书中也有一些有关多元微积分的例题，作为选学内容。因此，我们在一元范围内只是概括地介绍理论概念，但是教师利用选学内容、例题以及练习可以很容易地将其深入到统计学理论和数学概念之中。

本书特色

本书的主要特色如下：

1. **理论和应用相结合**。将数理统计的基本理论概念整合为一门统计方法的课程，供两个学期讲授。因此，教师可以选择将其作为以基本概念和应用统计为重点的一门课程，也可以作为偏向应用又介绍基本统计推断理论方法的一门课程。

2. **统计软件应用指导**。老师和学生可以选用统计软件进行统计计算。本书介绍了三个流行的统计软件包（SAS、SPSS 以及 MINITAB）的输出结果以及 Microsoft Excel 的输出结果。附录 C、附录 D 以及附录 E 介绍了菜单屏幕和对话框的使用，是为初学者设计的，这些辅导材料不需要预先使用这些软件的经验。

3. **主题和应用的结合范围**。为了满足未来工程师和科学家的种种需要，本书提供了覆盖范围广泛的数据分析主题。本书对多元回归以及模型构造（第11章和第12章）、试验设计的原理（第13章）、质量控制（第16章）以及可靠性（第17章）等内容的安排与通常的初等统计学教程不同。虽然这些题材通常涉及理论概念，但是讲述是面向应用的。

4. **基于大量实际数据的练习**。本书包含了大量的应用练习，主要是为了激发学生的学习兴趣 and 启发学生利用所学方法解决实际问题。几乎每一个练习都是基于摘自专业期刊或者从工程和自然科学团体得到的数据或试验结果。练习放在每章重点节的末尾以及各章的末尾。

译者序

国内已经出版了许多统计学方面的教材,包括国内专家学者编写的以及翻译的,都各有特点.本书是佛罗里达大学 William Mendenhall 教授与南佛罗里达大学 Terry Sincich 副教授编写的《Statistics for Engineering and the Sciences》的第5版,自2007年出版以来,得到了许多同行的肯定评价,被称赞为“是一本真正经典的统计方面的教材”,“容易读”,“在利用实际数据方面做出了非常出色的工作”,“向学生们极其清楚地讲述了统计学/概率论的基本概念”.

本书作为理工科各专业本科生的统计学考材,并不十分强调统计学理论的严格数学证明.在这样一本大部头著作中,证明只占很少一部分,绝大部分是与实际应用紧密联系的例子与练习.本书的最大特点:联系实际.全书给出了近250个例题、1200道练习题,每章最后还有一节“STATISTICS IN ACTION”,我们译成“活动中的统计学”,所有这些都是由活跃在各个领域的工程师、科学家以及统计学家提供的有据可查(表明了出处)的真实例子.这些例子涉及数、理、化、生、地、天文等自然科学,以及几乎所有工程技术领域.除了少量的理论练习,每个应用练习都有一个标题.有些问题的标题可能多处出现,如果将它们串在一起,就构成这个问题由浅入深的研究过程.

本书的第二个特点是介绍了统计软件的应用.统计学是数据的科学,一个真实的统计问题离不开统计数据的大量计算.书中的练习当然只能列出其中一小部分数据,附录中有SAS、MINITAB、SPSS等统计软件的用法简介.有兴趣利用统计软件解决实际问题的读者,可从华章网站(www.hzbook.com)下载包含全部数据的文件,按相应练习中的要求,解释运行结果,这是十分有益的.

我们翻译的是原著的第5版,它在前几版的基础上增加了许多新内容.全书共17章,包括了统计学的基本内容,其中有不少是国内高校“统计学”课程中没有或很少提及的,如试验设计、统计质量控制、可靠性等.每章末有快速回顾,帮助读者更好、更快地掌握内容.

翻译这样一本大作对我们的水平和耐心是一次考验.虽然在我们共同努力下,克服困难,完成了任务,但其中肯定存在不少问题,特别是某些专业术语的翻译,借助于互联网,可以找到许多相关的意义,而我们以前又从来没有接触过这个领域,因此有时只能望文生义选择我们认为比较合适的译名.参与本书翻译的有:梁冯珍、关静、吴新荣、韩月丽、李秀敏、徐付霞、刘晶、李璠、郭慧、辛凌雯、蔡霞、贺广婷.其中关静负责前9章,梁冯珍负责后8章,最后由史道济审校定稿.由于我们水平有限,翻译不当之处在所难免,恳请广大专家及读者指出,不胜感激.

译者

2009年5月

5. **“活动中的统计学”案例分析.** 现在本书在每章末尾新增了一个当代的科学研究(“活动中的统计学”)以及附带的数据与分析. 我们的目的是向学生展示评估发现和思考涉及的统计问题, 以及应用正确的统计方法的重要性.

6. **在线提供数据集.** 与例题、练习以及案例有关的所有数据集都可以从网上在线获得. 每个数据集在本书中都有一个 CD 图标和文件名. 数据文件以 5 种不同的格式(MINITAB、SAS、SPSS、Excel 和 ASC II)保存(以便易于导入其他统计软件包中). 利用统计软件分析这些数据, 将计算减少到最低限度, 使学生可以集中精力解释结果.

第 5 版的更新之处

虽然目的和范围与以前版本是一样的, 但是本书的第 5 版包含若干重要的改动、增补和加强:

1. **每章末尾的“活动中的统计学”案例.** 正如上面提到的, 现在本书在每章末尾增加了一个实际的最新科学研究的讨论. 详细地讨论数据和分析, 从而展示本章的统计方法是怎样用来回答此领域里的相关问题的.

2. **最新的统计软件输出结果.** 在本书中, 我们增加了大量输出结果, 包括最新版本的 SPSS 和 Excel 输出结果以及最新的 SAS 和 MINITAB 输出结果. 输出结果及相关的每种统计方法使教师能够把重点放在对统计结果的解释上, 而不是把重点放在为了得到结果而需要的计算上.

3. **章末的总结材料.** 在每一章的末尾, 我们通过快速回顾、符号汇集以及每章总结提示, 提供一个主题概要. 这些可以帮助学生总结和提炼本章的重点, 是有用的学习工具.

4. **更多的实际数据练习.** 本书添加了许多新的“现实”的科学练习. 它们全都来自报刊、科学杂志以及专业学报.

5. **第 1 章: 统计学和批判性思考.** 增加了统计学在评估由数据作出推断的可信性与价值中的作用一节, 即统计学在批判性思考中的作用(1.4 节).

6. **第 2 章: 描述性统计歪曲事实真相.** 在新的 2.8 节中包含了一些例题, 用来说明图形和数字描述性度量如何歪曲对数据的观察, 以及当这种事情发生时, 怎样认识它.

7. **第 4~6 章: 随机变量的标准数学符号.** 在有关随机变量的各章中, 我们都使用标准的数学符号来表示一个随机变量. 大写字母用于表示随机变量, 小写字母表示随机变量的可能取值.

8. **第 6 章: 蒙特卡罗 (Monte Carlo) 模拟.** 我们通过 6.8 节中的例题强调用蒙特卡罗模拟逼近抽样分布.

9. **第 7 章和第 8 章: 自助法和贝叶斯方法.** 现在新的可选小节给出另外两种估计方法(7.12 节)以及假设检验方法(8.13 节)——自助法和贝叶斯方法.

10. **第 9 章: 列联表的精确检验.** 新的可选小节(9.6 节)给出了 2×2 列联表独立性的费希尔 (Fisher) 精确检验.

11. **第 11 章: 多元回归模型.** 在第 11 章的不同小节中给出了几个不同的多元回归模型: 一阶模型(11.7 节)、交互作用模型(11.8 节)以及二阶模型(11.9 节).

12. **第 12 章: 外部模型确认.** 在第 12 章中增加了外部确认多元回归模型有效性的可选小节(12.9 节), 讨论了交叉确认法和刀切法.

13. **第 13 章和第 14 章: 试验设计和方差分析.** 以往关于方差分析的题材现在分成两章: “试验设计的原理”(第 13 章)和“试验设计的方差分析”(第 14 章).

14. **第 16 章：能力分析**. 在有关统计过程和质量控制的一章中现在包括了有关能力分析重要性的新的可选小节(16.9 节).

全书还有许多细节上不是很明显的变化, 这是按照本书当前的读者和审阅人的建议作出的修改.

辅助读物

本书以下面的补充性读物为配套资料:

1. **学生解答指南**(Mark Dummeldinger)(ISBN:0-13-187708-9). 本指南包括本书中全部奇数号练习的完全解答.

2. **教师解答指南**(Mark Dummeldinger)(ISBN:0-13-187707-0). 本指南包括本书中全部偶数号练习的完全解答.

致谢

本书是许多人多年共同努力的结果. 首先, 我们感谢下列各位教授, 他们对本版以及前几版的评审意见已经改入到第 5 版中. 本书第 5 版的审阅人:

Dharam Chopra(Wichita University)

Edward Danial(Morgan State University)

Frank Guess(University of Tennessee)

Carol O'Connor Holloman(University of Louisville)

Nancy Matthews(University of Oklahoma)

Alan Rabideau(University of Buffalo)

Charles Reilly(University of Central Florida)

Larry Ringer(Texas A&M University)

Arnold Sweet(Purdue University)

Paul Switzer(Stanford University)

以前几版的审阅人:

Carl Bodenschatz(United States Air Force Academy)

George C. Derringer(Battelle Columbus, Ohio, Division)

Danny Dyer(University of Texas-Arlington)

Herberg Eisenberg(West Virginia College of Graduate Studies)

Christopher Ennis(Normandale Community College)

Nasrollah Etemadi(University of Illinois-Chicago)

Linda Gans(California State Polytechnic University)

Carol Gattis(University of Arkansas)

K. G. Janardan(Eastern Michigan University)

H. Lennon(Coventry Polytechnic, Coventry, England)

Jeffery Maxey(University of Central Florida)

Curtis McKnight(University of Oklahoma)

Chand Midha(University of Akron)
Balgobin Nandram(Worcester Polytechnic Institute)
Paul Nelson(Kansas State University)
Norbert Oppenheim(City College of New York)
Giovanni Parmigiani(Duke University)
David Powers(Clarkson University)
David Robinson(St. Cloud State University)
Shiva Saksena(University of North Carolina-Wilmington)
Dennis Wackerly(University of Florida)
Donald Woods(Texas A&M University)

其他贡献者

特别感谢我们的助手 Mark Dummeldinger, Ed Danial 和 Weiwei Chen 博士, 他们非常认真地检查了第 5 版, 帮助我们有了一本高度精确和清晰的教材. 最后感谢培生教育出版集团的工作人员 Sally Yagan、Jennifer Urban 以及 Joanne Wendelken, 他们在本书编写和出版的各个环节给予了我们很大的帮助.

目 录

译者序

前言

第 1 章 绪论 1

- 1.1 统计学:数据的科学 1
- 1.2 统计学的基本要素 1
- 1.3 数据类型 4
- 1.4 统计学在批判性思考中的作用 5
- 1.5 本书介绍的统计方法导引 6

第 2 章 描述性统计 9

- 2.1 描述定性数据的图形法和数值法 9
- 2.2 描述定量数据的图形法 13
- 2.3 描述定量数据的数值法 20
- 2.4 中心趋势的度量 20
- 2.5 变异性的度量 24
- 2.6 相对位置的度量 28
- 2.7 检测异常值的方法 31
- 2.8 描述性统计歪曲事实真相 34

第 3 章 概率 46

- 3.1 概率在统计学中的作用 46
- 3.2 事件、样本空间和概率 46
- 3.3 复合事件 54
- 3.4 补事件 57
- 3.5 条件概率 60
- 3.6 并和交的概率法则 63
- *3.7 贝叶斯法则 71
- 3.8 计数法则 73
- 3.9 概率和统计的示例 82
- 3.10 随机抽样 83

第 4 章 离散随机变量 94

- 4.1 离散随机变量的定义 94
- 4.2 离散随机变量的概率分布 94
- 4.3 随机变量的期望值 98

4.4 一些有用的期望值定理 101

4.5 伯努利试验 103

4.6 二项概率分布 104

4.7 多项概率分布 109

4.8 负二项概率分布和几何概率分布 114

4.9 超几何概率分布 117

4.10 泊松概率分布 121

*4.11 矩和矩母函数 126

第 5 章 连续随机变量 136

5.1 连续随机变量 136

5.2 连续随机变量的密度函数 137

5.3 连续随机变量的期望值 139

5.4 均匀概率分布 143

5.5 正态概率分布 145

5.6 判定正态性的描述性方法 149

5.7 Γ 型概率分布 153

5.8 威布尔概率分布 157

5.9 β 型概率分布 159

*5.10 矩和矩母函数 162

第 6 章 二元概率分布及抽样分布 170

6.1 二元离散随机变量的概率分布 170

6.2 二元连续随机变量的概率分布 174

6.3 两个随机变量的函数的期望值 177

6.4 独立性 178

6.5 两个随机变量的协方差和相关性 180

*6.6 随机变量函数的概率分布和期望值 182

6.7 抽样分布 189

6.8 用蒙特卡罗模拟逼近抽样分布 189

6.9 均值与和的抽样分布 191

6.10 二项分布的正态逼近 196

6.11 与正态分布有关的抽样分布 199

第 7 章 用置信区间估计 210

7.1 点估计及其性质	210	10.3 估计 β_0 和 β_1 : 最小二乘法	352
7.2 求点估计: 经典估计方法	213	10.4 最小二乘估计的性质	361
7.3 求区间估计: 枢轴法	219	10.5 σ^2 的估计量	363
7.4 总体均值的估计	224	10.6 评价模型的效用: 进行关于斜率 β_1 的推断	366
7.5 两个总体均值差的估计: 独立样本	229	10.7 相关系数	370
7.6 两个总体均值差的估计: 配对	234	10.8 决定系数	373
7.7 总体比率的估计	239	10.9 利用模型估计和预测	377
7.8 两个总体比率差的估计	241	10.10 一个完整的例子	383
7.9 总体方差的估计	244	10.11 简单线性回归步骤的小结	385
7.10 两个总体方差比的估计	246	第 11 章 多重回归分析	396
7.11 选择样本容量	250	11.1 多重回归模型的一般形式	396
*7.12 其他区间估计方法: 自助法和 贝叶斯法	253	11.2 模型假定	397
第 8 章 假设检验	266	11.3 拟合模型: 最小二乘法	397
8.1 假设统计检验与置信区间的关系	266	11.4 用矩阵代数计算: 关于单个 β 参数的估计和推断	398
8.2 统计检验的要素与性质	266	11.5 评价整体模型的恰当性	404
8.3 求检验统计量: 经典方法	270	11.6 $E(y)$ 的置信区间和未来值 y 的 预测区间	407
8.4 选择原假设和备择假设	274	11.7 定量预测量的一阶模型	414
8.5 检验总体均值	275	11.8 定量预测量的交互作用模型	423
8.6 检验的观测显著性水平	280	11.9 定量预测量的二阶(二次)模型	427
8.7 检验两个总体均值的差: 独立样本	283	11.10 检查假定: 残差分析	433
8.8 检验两个总体均值的差: 配对	288	11.11 某些陷阱: 可估性、多重共线性 和外推	448
8.9 检验总体比率	293	11.12 多重回归分析的步骤总结	454
8.10 检验两个总体比率的差	295	第 12 章 模型构建	467
8.11 检验总体方差	298	12.1 引言: 为什么模型构建是重要的	467
8.12 检验两个总体方差的比	301	12.2 自变量的两种类型: 定量的和 定性的	468
*8.13 其他检验方法: 自助法和 贝叶斯法	304	12.3 一元定量自变量模型	469
第 9 章 分类数据分析	318	12.4 二元定量自变量模型	475
9.1 分类数据和多项概率	318	*12.5 编码定量自变量	481
9.2 估计单向表中的类型概率	318	12.6 一元定性自变量模型	485
9.3 检验单向表中的类型概率	322	12.7 定量和定性自变量模型	490
9.4 关于双向表(列联表)中类型 概率的推断	324	12.8 比较嵌套模型的检验	498
9.5 固定边缘和的列联表	331	*12.9 外部模型确认	505
*9.6 列联表分析中独立性的精确检验	335	12.10 逐步回归	506
第 10 章 简单线性回归	350	第 13 章 试验设计的原理	522
10.1 回归模型	350		
10.2 模型假定	350		

13.1 引言	522	第 16 章 统计过程和质量控制	647
13.2 试验设计术语	522	16.1 全面质量管理	647
13.3 控制试验中的信息	523	16.2 计量控制图	647
13.4 减少噪声的设计	524	16.3 均值控制图: $\bar{x}$ 图	651
13.5 增加容量设计	529	16.4 过程变异控制图: R 图	659
13.6 选择样本容量	533	16.5 发现控制图中的趋势: 游程分析	663
13.7 随机化的重要性	534	16.6 不合格品百分率控制图: p 图	664
第 14 章 试验设计的方差分析	539	16.7 每个个体缺陷数控制图: c 图	668
14.1 引言	539	16.8 容许限	671
14.2 方差分析中的逻辑	539	* 16.9 能力分析	674
14.3 单因子完全随机化设计	540	16.10 不合格品的抽样验收	679
14.4 随机化区组设计	550	* 16.11 其他抽样计划	682
14.5 双因子析因试验	560	* 16.12 调优操作	683
* 14.6 更复杂的析因设计	573	第 17 章 产品和系统的可靠性	693
* 14.7 套式抽样设计	580	17.1 引言	693
14.8 处理均值的多重比较	589	17.2 失效时间分布	693
14.9 检查 ANOVA 假定	595	17.3 危险率	694
第 15 章 非参数统计	609	17.4 寿命试验: 删失抽样	697
15.1 引言: 分布自由检验	609	17.5 估计指数失效时间分布的参数	697
15.2 检验单个总体的位置	610	17.6 估计威布尔失效时间分布的参数	701
15.3 比较两个总体: 独立随机样本	614	17.7 系统可靠性	705
15.4 比较两个总体: 配对设计	621	附录 A 矩阵代数	713
15.5 比较三个或更多总体: 完全随机化设计	626	附录 B 有用的统计表	725
15.6 比较三个或更多总体: 随机化区组设计	629	附录 C SAS 的视窗指导	759
15.7 非参数回归	633	附录 D MINITAB 视窗指导	772
		附录 E SPSS 视窗指导	787
		习题简答	801

第1章 绪论

目标 认识统计学在分析工程技术和自然科学数据中的作用。

1.1 统计学：数据的科学

根据《兰登书屋大学字典》(*The Random House College Dictionary*)，统计学是“对用数字表示的事实或数据进行收集、分类、分析以及解释的科学”。简而言之，**统计学就是数据的科学**。

定义 1.1 统计学是数据的科学。它包括数据的收集、分类、概括、整理、分析以及解释。

统计学通常应用于两种类型的问题：

1. 概括、描述以及探索数据。
2. 利用样本数据推断被选取样本的数据集的性质。

作为一个描述统计应用的例子，考虑美国的人口普查，它涉及数据集的收集，目的在于反映生活在美国的大约 29 500 万人口的社会经济特征。对于计算机软件工程师来说，管理这个巨大的数据库，其中一个问题就是利用统计方法描述数据。类似地，一位环境工程师利用统计学描述一个数据集，这个数据集记录了工厂在过去一年中氧化硫的日排放量。致力于这些应用的统计学分支称为**描述性统计**。

定义 1.2 致力于数据集的整理、概括以及描述的统计学分支称作**描述性统计**。

有时数据集刻画的是一种感兴趣的现象，这样的数据集在自然状态下是无法得到的、代价昂贵或者耗费时间才能获得的。在这种情况下，我们得到数据的一个子集(称为样本)，利用这个样本信息来推断它的性质。为了说明概念，假定感兴趣的现象是在一个有人居住但偏僻的太平洋岛上饮用水的质量。你可能期望水的质量依赖于水的温度、近期降雨量大小等因素。实际上，如果在同一个地区相同时刻重复测量水的质量，即使在水温相同的情况下，质量测量值也会变化。因此，“饮用水的质量”现象由一个很大的数据集来刻画，而这个数据集由许多个(事实上是无限多个)水的质量测量值组成——一个只是概念上的数据集。为了确定这个数据集的性质，我们从中抽样，即我们记录在特殊的时间和地点收集到的 n 份水样本的质量，然后利用这 n 个质量测量值的样本推断感兴趣的观念上很大的数据集的性质。解决这类问题的统计学分支称为**推断统计学**。

定义 1.3 利用样本数据对一个很大的数据集作出推断的统计学分支称作**推断统计学**。

1.2 统计学的基本要素

在统计学术语中，我们把想要描述的数据集或刻画了我们感兴趣现象的数据集称为**总体**。这样，我们可以将**样本**定义为取自总体的数据的子集。有时，总体和样本这两个词用来表示对象集和对它进行的测量(即试验单位集)。在一项研究中，这些术语的意义从它们的上下文中可以清楚地看出。

定义 1.4 统计**总体**是我们感兴趣的¹目标的数据集(通常很大，有时是概念上的)。

定义 1.5 **样本**是取自目标总体的数据子集。

定义 1.6 测量值采集的对象(例如，人、事物、交易、标本或者事件)称作**试验单位**。

注：一个总体由对许多试验单位采集的数据组成。

在研究总体和样本的过程中，我们专注于总体试验单位的一个或多个特征或性质，统计学称这

些特征为**变量**. 例如, 在饮用水质量的研究中, 工程师感兴趣的两个变量是在 100 mL 的水样本中, 氯的残留量(按每百万的含量测量)以及排泄物大肠杆菌的数量.

定义 1.7 变量是单个试验单位的特征或性质.

例 1.1 肯塔基大学运输研究项目的工程师们收集了发生在肯塔基列克星敦交叉路口的事故数据. 研究目的之一是估计在没有只许左转弯车道的交叉路口发生左转弯事故的比例. 这个估计值将用来在所有的列克星敦主要交叉路口设置左转弯车道提供数值理由(或准则). 工程师们用一年多的时间收集了 50 个没有只许左转弯车道的主干道上交叉路口的事故数据. 在每个交叉路口, 他们监测通行车辆并记录发生事故的左转弯汽车的总数.

- a. 确定这项研究的变量以及试验单位.
- b. 描述目标总体以及样本.
- c. 运输工程师想要作出什么推断?

解 a. 由于工程师收集了 50 个交叉路口的数据, 试验单位是没有左转弯车道的交叉路口. 测量的变量是发生事故的左转弯汽车总数.

b. 研究的目的是对于所有列克星敦的主要交叉路口设置左转弯车道提出一些准则; 因此, 目标总体由这个城市所有主要交叉路口组成. 样本由工程师监测的 50 个交叉路口的子集组成.

c. 工程师将利用样本数据估计发生在列克星敦的所有主要交叉路口的左转弯事故的比例. (在第 7 章中我们将知道这个估计值是样本中左转弯事故的个数除以样本中左转弯的汽车总数.)

前面的定义和例题给出了推断统计问题 5 个要素中的 4 个: 总体、一个或多个感兴趣的变量、样本以及推断. 第 5 个要素是关于知道这个推断如何有效, 即推断的**可靠性**. 推断的可靠性度量将统计学和算命术区分开来. 与统计学家一样, 看手相的可能会检查一个样本(你的手)并对总体(你将来的寿命)作出推断. 但是和统计推断不同, 看手相的推断不包括推断有多大可能是真实的度量.

为说明问题, 考虑例 1.1 中运输工程师在肯塔基列克星敦交叉路口处左转弯交通事故比例的估计. 工程师们感兴趣的是估计的误差(即样本事故发生率和目标总体事故发生率之间的差异). 利用统计方法, 我们可以给出一个估计误差的界. 这个界是一个数(如 10%), 我们的估计误差不大可能超过它. 在后面的章节中, 我们将会学习用这个界来帮助度量推断的“置信度”. 全书都会讨论统计推断的可靠性. 现在, 已经简单地认识到没有可靠性来度量的推断是不完整的.

定义 1.8 可靠性度量是关于统计推断不确定程度的一个陈述(通常是定量的).

下面给出了描述性统计和推断性统计问题所包含的要素.

描述性统计问题的 4 要素

1. 感兴趣的总体或样本.
2. 被研究的一个或多个变量(总体或样本单位的特征).
3. 表格、图形或者数字概括工具.
4. 确定数据类型.

推断性统计问题的 5 要素

1. 感兴趣的总体.
2. 被研究的一个或多个变量(试验单位的特征).
3. 试验单位的样本.

4. 基于包含在样本中的信息对总体的推断。
5. 推断的可靠性度量。

应用练习

- 1.1 钢的防腐研究. 希腊雅典国家技术大学材料科学与工程系的研究者检测了钢表层的不同环氧涂层的防腐性能. (*Pigment & Resin Technology*, 32 卷, 2003) 从一卷薄钢板上切下几条钢片, 涂上 4 种不同类型环氧层(S1、S2、S3 以及 S4)中的一种. 将这些钢片放在水中一天后, 确定每个钢片的腐蚀率(毫微安/平方厘米).
 - a. 这个研究的试验单位是什么?
 - b. 假设你仅对描述涂上 S1 类型的环氧层的钢板的腐蚀率感兴趣, 定义目标总体以及相关的样本.
- 1.2 软件重新使用的成功或失败. 由渥太华大学主持的软件工程知识库 PROMISE 软件是一个公开的通用数据集, 为研究者建立预测软件模型提供服务. 有关软件重新使用的 PROMISE 数据集保存在 SWREUSE 文件中, 对于 24 个新软件开发项目样本中的每个项目, 它提供了利用以前开发的软件是成功或者失败的信息. (数据来源: *IEEE Transactions on Software Engineering*, 28 卷, 2002.) 在 24 个项目中, 9 个失败, 15 个成功运行.
 - a. 确定这个研究的试验单位.
 - b. 描述样本所来自的总体.
 - c. 利用样本信息对总体作出推断.
- 1.3 由杀虫剂引起的果树污染. 将杀虫剂用于一大批成熟的庄稼无疑会导致周围空气的污染. *Environmental Science & Technology* (1993 年 10 月) 刊登了在加利福尼亚州圣华金河谷的休眠果园使用杀虫剂毒死蜱后硫的残留量. 收集了一个果树林在一段喷洒较密集的时期(总共 13 天)每天的周围空气样本, 测量了每天的硫含量水平(ng/m^3).
 - a. 确定研究者感兴趣的总体.
 - b. 确定样本.
- 1.4 地震的地表运动. 在 *Journal of Earthquake Engineering* (2004 年 11 月), 一个由土木工程师和环境工程师组成的团队研究了发生在 1940 年至 1995 年之间, 世界范围内的 15 次地震的地表运动特征. 每次地震测量的三个(多个)变量是地表运动的类型(短、长或者向前)、地震的震级(里氏震级)以及最大地表加速度(ft/s). 研究目的之一是估计任意地表运动周期的非弹性谱.
 - a. 确定此研究的试验单位.
 - b. 15 次地震的数据代表一个总体还是一个样本? 说明原因.
- 1.5 计算机电源系统负载电流. 电子工程师发现计算机电源系统中有高的中性线电流是一个潜在的问题. 为了确定问题的范围, 他们对 146 个美国站点的计算机电源系统的负载电流进行了调查(*IEEE Transactions on Industry Applications*, 1990 年 7/8 月). 调查显示, 有高的中性线电流的站点对满载电流站点的比例不到 10%.
 - a. 确定感兴趣的总体.
 - b. 确定样本.
 - c. 利用样本信息作出关于总体的推断.
- 1.6 蔬菜的预冷. 为了给市场准备佛罗里达蔬菜, 研究者开发了一种新的预冷方法. 系统利用设计的空气和水的混合, 迫使比常规水冷法更低温的水流来有效地冷却. 为了比较两种系统的有效性, 将 20 批新鲜的西红柿分成两部分: 一部分用新方法冷却, 另一部分用常规方法冷却. 记录了为有效冷却每批西红柿所需的水量(加仑, 1 英加仑 = 4.546 升, 1 美加仑 = 3.785 升).
 - a. 确定这个问题的总体、样本以及所做的统计推断类型.
 - b. 怎样利用样本数据比较两种系统的冷却效率?
- 1.7 每周一氧化碳数据. 温室效应气体的世界数据中心收集并保存了温室以及空气中相关气体的数据. 其中一个数据集列出了位于阿拉斯加科尔德湾气象站的空气中每周一氧化碳的含量(十亿分之一). 2000 年至 2002 年的每周数据保存在 COGAS 文件中.
 - a. 确定被测量的变量以及相应的试验单位.
 - b. 如果你仅对描述 2000 年至 2002 年中科尔德湾站的每周一氧化碳的值感兴趣, 那么数据代表的是总体还是样本? 说明原因.
- 1.8 检测次品. 检查一条生产线上所有产品中的次品是一个花费大且耗时的过程. 检查次品的一个有效且经济的方法是由质量控制工程师选择并检查一部分产品. 计算在检查产品中次品的百分比, 然后用它估计这条生产线上所有产品中有次品的百分比. 对这个问题确定总体、样本以及统计推断类型.

1.3 数据类型

数据分为定量和定性两种类型^①。定量数据表示事物的数量或个数，用数值标度量。例如，电源频率(MHz)是一个定量变量，类似于钢管的断裂强度(lb/in²，磅/平方英寸)。与之相反，定性(或分类)数据没有量的解释，它们只能分类。相应于一组 n 个工科毕业生， n 个职位集合是一个定性数据集。防腐环氧涂层中使用的色素种类(锌或云母)也表示为定性数据。

定义 1.9 定量数据表示事物的数量或个数。

定义 1.10 定性数据没有量的解释，即它们只能以类型分类。

例 1.2 *Journal of Performance of Constructed Facilities* (1990 年 2 月) 报告了费城地区配水网络的性能尺度。作为研究的一部分，测量了每个抽样水管截面的如下变量，确定每个变量产生的数据是定量的还是定性的。

- a. 水管直径(in)。
- b. 水管材料(钢铁或者 PVC)。
- c. 水管位置(市中心或者郊区)。
- d. 水管长度(ft)。

解 水管直径(in)以及水管长度(ft)都是有意义的数值标度量度的。因此，这两个变量产生定量数据，水管材料和水管位置只能分类——材料是钢铁或者 PVC；位置是市中心或者郊区。因此，水管材料和水管位置是定性变量。

用于描述、分析数据的合适的统计工具是依赖于数据类型的。因此，区别定量数据和定性数据是很重要的。

应用练习

1.9 饮用水质量研究。Disasters (28 卷, 2004 年) 发表了一篇关于热带旋风对一个偏远的太平洋岛屿上饮用水质量影响的研究。收集了在旋风亚美袭击这个岛屿后大约 4 个星期的水样本(大小为 500 mL)。对每个水样本记录了下面的变量，确定每个变量是定量的还是定性的。

- a. 收集样本的城镇。
- b. 供水的类型(内陆河、溪、地泉)。
- c. 酸度(pH 值, 1~14)。
- d. 混浊程度(比浊法浊度单位 = NTUs)。
- e. 温度(°C)。
- f. 每 100 mL 水里大肠杆菌个数。
- g. 游离氯的残留量(mg/L)。
- h. 是否有硫化氢(有或没有)。

1.10 新西兰灭绝的鸟类。加利福尼亚大学(河滨分校)的环境工程师正在研究新西兰鸟类总体中灭绝的模型。(*Evolutionary Ecology Research*, 2003. 7.) 下面给出了每个在毛利人殖民地(即欧洲人到达之前)时期栖息于新西兰的鸟类的特征。确定每个变量是定量的还是定性的。

- a. 飞行能力(会飞的或不能飞的)。
- b. 栖息地类型(水栖、地面陆栖或者空中陆栖)。
- c. 筑巢地点(地面、地里的窝、树、地上的窝)。
- d. 巢密度(高或低)。
- e. 食物(鱼、脊椎动物或者无脊椎动物)。
- f. 身体质量(g)。
- g. 蛋的长度(mm)。
- h. 灭绝状态(灭绝、从岛上迁出、没有灭绝)。

1.11 计算机电源系统负载电流。参考练习 1.5 中 *IEEE Transactions on Industry Applications* (1990 年 7/8

① 数据类型的更细致的划分是名义、次序、区间和比率数据。名义数据是定性数据，其类型不能进行有意义的排序。次序数据也是定性数据，但是存在一个从高到低的明显的分组等级。区间和比率数据是两种不同类型的定量数据。对于大多数统计应用(以及在此序言中的所有方法)，将数据分为定量的和定性的已经足够了。

月)计算机电源系统负载流量的调查.除了记录中性线电流对满负载电流的比例,研究者也记录了负载的类型(线-线或者线-中性线)以及计算机系统卖主.对记录的每个变量确定数据类型.

- 1.12 肺癌的 CT 扫描.已经开发了一种新型的肺癌扫描技术——计算机 X 线断层摄影术(CT).医学家认为 CT 扫描在查找小肿瘤方面比常规的 X 射线更敏感.南佛罗里达大学的 H. Lee Moffitt 癌症中心正在进行全国范围的一个 50 000 名吸烟者的临床试验,以比较 CT 扫描和 X 射线对检测肺癌的有效性.(*Today's Tomorrows*, 2002 年秋.)每个参加试验的吸烟者被随机分配到两种拍摄方法——CT 或者 X 射线中的一种,并且随时记录他们的发展状况.除了所用的拍摄方法外,临床学者还记录了每名吸烟者第一次用分配的拍摄方法检测到肿瘤的年龄.
 - a. 确定本次研究的试验单位.
 - b. 确定每个试验单位度量的两个变量.
 - c. 确定度量变量的类型(定量的或者定性的).
 - d. 根据临床试验,最终将得到什么样的推断?
- 1.13 深孔钻探.“深孔”钻探是当孔的深度与孔的直径之比超过 10 时采用的一类钻探作业过程.成功的深

孔钻探依赖于钻屑的良好排出.曾经有人做过试验,研究当发生钻屑充塞时深孔钻探的性能(*Journal of Engineering for Industry*, 1993.5).下面描述了钻探过程中的一些重要变量,确定每个变量的类型.

- a. 钻屑排出速度(每分钟排出的钻屑).
 - b. 钻探深度(mm).
 - c. 油的速度(mm/s).
 - d. 钻探的类型(单边、BTA 或者推顶).
 - e. 孔表面的质量.
- 1.14 美国国家桥梁目录.联邦公路局(FHWA)定期检查美国所有桥梁的结构缺陷.FHWA 调查的数据编入国家桥梁目录(NBI).下面列出的是由 NBI 维护的大约 100 个变量中的几个.确定每个变量是定量的还是定性的.
 - a. 最大跨度(ft).
 - b. 车道数目.
 - c. 桥梁是否征税(是或否).
 - d. 平均日交通量.
 - e. 桥面条件(好、很好或差).
 - f. 绕行公路或便道的长度(mile).
 - g. 公路类型(州际、美国、州、乡村或城市的).

1.4 统计学在批判性思考中的作用

工程与科学中的试验研究一般都牵涉到利用试验数据(样本)推断某个概念性总体的性质,而这个概念性的总体刻画了试验者感兴趣的现象.这个推断过程是科学方法的一个组成部分.基于试验数据的推断首先用来提出有关现象的理论,然后利用另外的样本数据来检验这个理论.

统计科学对这个过程是怎样发挥作用的?为了回答这个问题,我们必须注意基于样本数据的推断几乎都是有误差的,因为样本不能提供总体精确的映像.样本提供的信息特征依赖于选取的特定样本,因此不同样本得到的信息是不一样的.例如,假定你想估计美国石化工厂中,所有应力腐蚀破裂而导致合金钢损坏的比例.你调查了一个 100 个合金钢损坏的样本的原因,发现 47 个是应力腐蚀破裂导致的.这是否意味着在石化工厂中所有合金钢损坏都精确地有 47% 是应力腐蚀破裂造成的呢?显然不是.假定你不知道合金钢损坏应力腐蚀破裂的真实百分比是 44%.100 个损坏的样本可能有 47 个是应力腐蚀破裂导致的,而另一个 100 个损坏的样本中可能只有 42 个.因此,基于样本的推断具有不确定性.

另一方面,假定一个石化工厂曾经有一次合金钢损坏率为 81%.当给定样本损坏率为 47% 时,这是否是一个非常高的损坏率呢?统计理论用概率来度量关于推断的不确定性.这使得工程师和科学家有能力在有关总体的特殊假定下,计算观测到的特殊样本或者数据测量值的概率.这些概率用来评估关于样本推断的不确定性.例如,在给定的样本信息下,我们可以通过计算观测到如此高比例的机会,来确定工厂 81% 的合金钢损坏率是否非常高.

因此,统计学的主要贡献是,使得工程师和科学家能够用已知的可靠性度量作出推断(关于目标总体的估计以及决策).这样工程师可以根据数据作出明智的决策与推断,即统计学帮助工程师批

判地思考他们的结论.

定义 1.11 统计思想包括应用理性思维和统计科学来批判地评估数据及推断.

1.5 本书介绍的统计方法导引

虽然我们给出了一些探索和描述数据集的有用方法(第2章),但是本书及现代统计学主要强调的是在推断统计学的领域.图1.1的流程图给出了本书各章的轮廓,并且可以作为选择统计方法的指南.

