

统计学

(原书第6版)

Statistics for Engineering and the Sciences

(Sixth Edition)



[美] 威廉·M. 门登霍尔 特里·L. 辛西奇 著
(William M. Mendenhall) (Terry L. Sincich)

关静 等译



机械工业出版社
China Machine Press

统计学精品译丛

统计学

(原书第6版)

Statistics for Engineering and the Sciences

(Sixth Edition)



[美] 威廉·M. 门登霍尔 特里·L. 辛西奇 著
(William M. Mendenhall) (Terry L. Sincich)

关静 等译



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

统计学 (原书第 6 版) / (美) 威廉·M. 门登霍尔 (William M. Mendenhall), (美) 特里·L. 辛西奇 (Terry L. Sincich) 著; 关静等译. —北京: 机械工业出版社, 2018.6
(统计学精品译丛)

书名原文: Statistics for Engineering and the Sciences, Sixth Edition

ISBN 978-7-111-60365-8

I. 统… II. ①威… ②特… ③关… III. 统计学—高等学校—教材 IV. C8

中国版本图书馆 CIP 数据核字 (2018) 第 138191 号

本书版权登记号: 图字 01-2017-0479

Statistics for Engineering and the Sciences, Sixth Edition by William M. Mendenhall and Terry L. Sincich (ISBN 978-1-4987-2885-0).

Copyright © 2016 by Taylor & Francis Group, LLC

Authorized translation from the English language edition published by CRC Press, part of Taylor & Francis Group LLC. All rights reserved.

China Machine Press is authorized to publish and distribute exclusively the Chinese (Simplified Characters) language edition. This edition is authorized for sale in the People's Republic of China only (excluding Hong Kong, Macao SAR and Taiwan). No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

Copies of this book sold without a Taylor & Francis sticker on the cover are unauthorized and illegal.

本书原版由 Taylor & Francis 出版集团旗下 CRC 出版公司出版, 并经授权翻译出版。版权所有, 侵权必究。

本书中文简体字翻译版授权由机械工业出版社独家出版并仅限在中华人民共和国境内 (不包括香港、澳门特别行政区及台湾地区) 销售。未经出版者书面许可, 不得以任何方式复制或抄袭本书的任何内容。

本书封面贴有 Taylor & Francis 公司防伪标签, 无标签者不得销售。

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。

出版发行: 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: 迟振春

印刷: 北京瑞德印刷有限公司

开本: 186mm×240mm 1/16

书号: ISBN 978-7-111-60365-8

责任校对: 李秋荣

版次: 2018 年 6 月第 1 版第 1 次印刷

印张: 59

定价: 139.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

购书热线: (010) 68326294 88379649 68995259

投稿热线: (010) 88379604

读者信箱: hzjsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光/邹晓东

译者序

William M. Mendenhall 与 Terry L. Sincich 编写的《Statistics for Engineering and the Sciences》第 5 版, 自 2007 年出版以来得到了很多同行的肯定, 被称赞为“是一本经典的统计教材”。2016 年, 他们在第 5 版的基础上对书中的内容进行更新、调整和修正, 形成了第 6 版。

本书作为理工科各专业本科生或研究生的统计学教材, 有两大特点: 一是理论与实际相结合, 重点突出实际应用; 二是强调统计软件的使用, 便于对数据进行分析。随着大数据的快速发展, 有效地分析数据并解决实际问题是我们应该具备的基本能力。而本书就是以来自各个领域的实例为出发点, 讲述统计方法, 并应用这些方法来分析实际问题, 从而便于读者理解并掌握这些方法。例如, 在每章的开始以“活动中的统计学”提出一个具体问题, 然后介绍理论方法, 最后以“活动中的统计学回顾”给出解决方法, 即利用该章介绍的统计方法结合统计软件对提出的问题进行分析和解答。

此外, 本书的例题近 250 个、习题超过 1000 道, 涉及数、理、化、天文、地理、生物等自然科学以及几乎所有的工程技术领域。除少量的理论练习外, 都是一些真实的问题。特别是, 书中关于 SAS、MINITAB、SPSS 等统计软件的介绍, 便于读者选用统计软件进行统计计算, 从而加深对统计方法的理解和掌握。

第 6 版是在第 5 版译文的基础上翻译而成的, 因此, 这里要感谢第 5 版译者史道济、梁冯珍等做出的贡献。第 6 版由关静全面负责翻译, 杨香云、闫一冰、陈永沛、魏伟等也参与了本书的部分翻译和校正工作。本书涉及的领域非常广, 特别是涉及很多专业术语, 我们通过查阅许多资料, 尽量采用较为贴切的翻译。由于水平有限, 翻译不当之处在所难免, 恳请广大读者及专家批评指正。

译者

2018 年 4 月

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:用置信区间估计和假设检验。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建(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)。

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,教师可以选择将本书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有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建(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,但讲述是面向应用的。

4. **基于大量实际数据的示例和练习**. 本书包含了大量的示例和练习,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。几乎每一个练习和示例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。应用练习放在每章重点节的末尾以及各章的末尾。

5. **“活动中的统计学”案例研究**. 每章都以一个当代的实际科学问题(“活动中的统计学”)讨论开始,并附带数据。由案例研究得到的分析和推断是每章的重点(“活动中的统计学回顾”)。我们的目的是向学生展示在评价结果和思考涉及的统计问题时,应用正确统计方法的重要性。

6. **章末的总结材料**. 在每一章的末尾,我们通过“快速回顾”——“重要公式”“符号汇集”以及

“本章总结提示”，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。

7. **随机变量的标准数学符号。**在有关随机变量的各章中，我们使用标准的数学符号来表示随机变量。大写字母表示随机变量，小写字母表示随机变量的可能取值。

8. **自助法和贝叶斯法。**在选学章节给出了科学研究中更为流行的两种估计方法(7.12节)以及假设检验方法(8.13节)——自助法和贝叶斯法。

9. **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站 www.crcpress.com/product/isbn/9781498728850 在线获得。每个数据集在本书中都有一个  图标和文件名。数据文件以4种不同的格式 MINITAB、SAS、SPSS 和 Excel 保存。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6版的更新之处

虽然目的和范围与以前版本是一样的，但是本书第6版包含若干重要的改动、增补和加强：

1. **超过1000个练习，修改和更新达到30%。**增加了许多新的和更新的练习，这些练习都是基于当代工程和自然科学相关的研究，并附带实际数据。其中很多练习选自科学期刊，可以培养和提升批判性思维能力。

2. **更新技术。**整本书增加了很多统计软件的输出。所有的输出来自统计软件(SAS、SPSS 和 MINITAB)，相应的使用说明已经修订在最新版本的软件中。

3. **活动中的统计学回顾。**对于第6版，每章开始介绍“活动中的统计学”案例(见上)。讲完所需要的方法之后，在章末“活动中的统计学回顾”中给出解决方法(数据分析和推断)和讨论。

4. **第1章：收集数据/抽样。**所有关于抽样概念的内容(如随机抽样和抽样调查设计)已经简化并移到1.4节，目的是让学生早点了解重要的抽样问题。

5. **第7章：配对和独立样本。**增加了一个例子(例7.12)，直接比较用配对法和独立样本 t 检验法分析相同的数据。

6. **第8章：假设检验/ p 值。**假设检验中 p 值一节(8.5节)转为强调它在工程与科学相关问题中应用的重要性。本书其余部分，由假设检验得到的结论都是基于 p 值。

7. **第10章和第11章：回归残差。**增加了一节新的内容(10.8节)，它是关于用回归残差来检验简单线性回归分析中所需的假设。多元回归一章中的类似节(11.10节)已经更改，重点介绍回归残差的不同使用，包括假设证明和检测异常值及有影响的观测值。

8. **第13章：试验设计。**增加了两个新的例子(例13.6和例13.7)，它们是关于试验设计中样本大小的选择。

9. **第14章：方差分析。**增加了两个新的例子(例14.8和例14.10)，它们是有定量因子的双因子试验。第一个例子采用传统的ANOVA模型，第二个采用有高阶项的回归模型。

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做出的修改。

辅助读物

1. **学生解答指南。**本指南包括本书中全部奇数号练习的完整解答。

2. **教师解答指南。**本指南给出本书中全部偶数号练习的解答。本指南仔细认真，以保证所有的

解答方法和符号与全书的核心内容保持一致。

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到第6版中：

第6版的审阅人

Shyamaia Nagaraj (University of Michigan)

Stacie Pisano (University of Virginia)

Vishnu Nanduri (University of Wisconsin-Milwaukee)

Shuchi Jain (Virginia Commonwealth University)

David Lovell (University of Maryland)

Raj Mutharasan (Drexel University)

Gary Wasserman (Wayne State University)

Nasser Fard (Northeastern University)

以前几版的审阅人

Carl Bodenschatz (United States Air Force Academy)

Dharam Chopra (Wichita University)

Edward Danial (Morgan State University)

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)

Frank Guess (University of Tennessee)

Carol O'Connor Holloman (University of Louisville)

K. G. Janardan (Eastern Michigan University)

H. Lennon (Coventry Polytechnic, Coventry,

England)

Nancy Matthews (University of Oklahoma)

Jeffery Maxey (University of Central Florida)

Curtis McKnight (University of Oklahoma)

Chand Midha (University of Akron)

Bal gobin Nandram (Worcester Polytechnic Institute)

Paul Nelson (Kansas State University)

Norbert Oppenheim (City College of New York)

Giovanni Parmigiani (Duke University)

David Powers (Clarkson University)

Alan Rabideau (University of Buffalo)

Charles Reilly (University of Central Florida)

Larry Ringer (Texas A&M University)

David Robinson (St. Cloud State University)

Shiva Saksena (University of North Carolina-Wilmington)

Arnold Sweet (Purdue University)

Paul Switzer (Stanford University)

Dennis Wackerly (University of Florida)

Donald Woods (Texas A&M University)

其他贡献者

特别感谢包括 Nancy Boudreau 在内的助手，他们中很多人已经和我们共事了很多年。最后，Taylor & Francis 集团出版人员 David Grubbs、Jessica Vakili 和 Suzanne Lassandro 在本书编写、出版和发行的各个环节给予了我们很大的帮助。

目 录

译者序	1.8
前言	5.8
第1章 绪论	1
活动中的统计学：田纳西河中鱼的 DDT 污染	1
1.1 统计学：数据的科学	1
1.2 统计学的基本要素	2
1.3 数据类型	4
1.4 收集数据：抽样	6
1.5 统计学在批判性思考中的作用	11
1.6 本书介绍的统计方法导引	12
活动中的统计学回顾：田纳西河中鱼的 DDT 污染——确定数据收集的方法、总体、样本和数据类型	13
第2章 描述性统计	15
活动中的统计学：亚拉巴马州田纳西河中污染鱼的特征	15
2.1 描述定性数据的图形法和数值法	15
2.2 描述定量数据的图形法	20
2.3 描述定量数据的数值法	27
2.4 中心趋势的度量	28
2.5 变异性的度量	33
2.6 相对位置的度量	38
2.7 检测异常值的方法	40
2.8 描述性统计歪曲事实真相	44
活动中的统计学回顾：亚拉巴马州田纳西河中污染鱼的特征	48
第3章 概率	58
活动中的统计学：NASA 太空船仪表码中的软件缺陷评估预测器	58
3.1 概率在统计学中的作用	58
3.2 事件、样本空间和概率	59

3.3 复合事件	67
3.4 补事件	69
3.5 条件概率	72
3.6 并和交的概率法则	76
*3.7 贝叶斯法则	84
3.8 计数法则	87
3.9 概率和统计的示例	96
活动中的统计学回顾：NASA 太空船仪表码中的软件缺陷评估预测器	97
第4章 离散随机变量	104
活动中的统计学：“一次性”装置的可靠性	104
4.1 离散随机变量的定义	104
4.2 离散随机变量的概率分布	105
4.3 随机变量的期望值	109
4.4 一些有用的期望值定理	112
4.5 伯努利试验	113
4.6 二项概率分布	114
4.7 多项概率分布	120
4.8 负二项概率分布和几何概率分布	124
4.9 超几何概率分布	128
4.10 泊松概率分布	131
*4.11 矩和矩母函数	137
活动中的统计学回顾：“一次性”装置的可靠性	139
第5章 连续随机变量	146
活动中的统计学：超级武器的开发——优化命中率	146
5.1 连续随机变量的定义	146
5.2 连续随机变量的密度函数	148
5.3 连续随机变量的期望值	150
5.4 均匀概率分布	154
5.5 正态概率分布	156

5.6 判定正态性的描述性方法	161	活动中的统计学回顾: PET 饮料瓶的破裂 强度	281
5.7 Γ 型概率分布	166	第 8 章 假设检验	292
5.8 威布尔概率分布	170	活动中的统计学: 比较溶解药片方法—— 溶解方法等效性检验	292
5.9 β 型概率分布	173	8.1 假设统计检验与置信区间的关系	293
5.10 矩和矩母函数	176	8.2 统计检验的要素与性质	293
活动中的统计学回顾: 超级武器的开发—— 优化命中率	177	8.3 求统计检验: 经典方法	298
第 6 章 二元概率分布及抽样分布	185	8.4 选择原假设和备择假设	301
活动中的统计学: Up/Down 维修系统的 可用性	185	8.5 检验的观测显著性水平	302
6.1 离散随机变量的二元概率分布	185	8.6 检验总体均值	304
6.2 连续随机变量的二元概率分布	190	8.7 检验两个总体均值的差: 独立样本	311
6.3 两个随机变量函数的期望值	193	8.8 检验两个总体均值的差: 配对	319
6.4 独立性	194	8.9 检验总体比率	324
6.5 两个随机变量的协方差和相关性	196	8.10 检验两个总体比率的差	327
6.6 随机变量函数的概率分布和期望值	199	8.11 检验总体方差	332
6.7 抽样分布	205	8.12 检验两个总体方差的比	335
6.8 用蒙特卡罗模拟逼近抽样分布	205	8.13 其他检验方法: 自助法和贝叶斯法	340
6.9 均值与和的抽样分布	208	活动中的统计学回顾: 比较溶解药片方法—— 溶解方法等效性检验	344
6.10 二项分布的正态逼近	212	第 9 章 分类数据分析	354
6.11 与正态分布有关的抽样分布	215	活动中的统计学: 残忍的组织移植案例—— 谁应该为损害赔偿负责	354
活动中的统计学回顾: Up/Down 维修系统的 可用性	219	9.1 分类数据和多项概率	355
第 7 章 用置信区间估计	227	9.2 估计单向表中的类型概率	355
活动中的统计学: PET 饮料瓶的破裂强度	227	9.3 检验单向表中的类型概率	359
7.1 点估计及其性质	227	9.4 关于双向表(列联表)中类型概率的 推断	363
7.2 求点估计: 经典估计方法	231	9.5 固定边缘和的列联表	369
7.3 求区间估计: 枢轴法	236	9.6 列联表分析中独立性的精确检验	373
7.4 总体均值的估计	242	活动中的统计学回顾: 残忍的组织移植案例—— 谁应该为损害赔偿负责	379
7.5 两个总体均值差的估计: 独立样本	247	第 10 章 简单线性回归	387
7.6 两个总体均值差的估计: 配对	253	活动中的统计学: 探矿魔杖真的能发现水吗	387
7.7 总体比率的估计	259	10.1 回归模型	388
7.8 两个总体比率差的估计	262	10.2 模型假定	389
7.9 总体方差的估计	265	10.3 估计 β_0 和 β_1 : 最小二乘法	391
7.10 两个总体方差比的估计	269	10.4 最小二乘估计的性质	401
7.11 选择样本容量	273	10.5 σ^2 的估计量	403
7.12 其他区间估计方法: 自助法和 贝叶斯法	276		

10.6 评价模型的效用: 进行关于斜率 β_1 的推断	406	12.7 定量和定性自变量模型	549
10.7 相关系数和决定系数	411	12.8 比较嵌套模型的检验	558
10.8 利用模型估计和预测	418	12.9 外部模型确认	564
10.9 检验假定: 残差分析	425	12.10 逐步回归	566
10.10 一个完整的例子	434	活动中的统计学回顾: 取消州内货车运输业管制	572
10.11 简单线性回归步骤小结	437	第13章 试验设计的原理	586
活动中的统计学回顾: 探矿魔杖真的能发现水吗	437	活动中的统计学: 加锌环氧涂层的防腐行为	586
第11章 多重回归分析	447	13.1 引言	586
活动中的统计学: 高速公路建设中的串通投标	447	13.2 试验设计术语	587
11.1 多重回归模型的一般形式	447	13.3 控制试验中的信息	588
11.2 模型假定	448	13.4 减少噪声的设计	589
11.3 拟合模型: 最小二乘法	449	13.5 增加容量设计	594
11.4 用矩阵代数计算: 关于单个 β 参数的估计和推断	450	13.6 选择样本容量	598
11.5 评价整体模型的恰当性	457	13.7 随机化的重要性	601
11.6 $E(y)$ 的置信区间和未来值 y 的预测区间	460	活动中的统计学回顾: 加锌环氧涂层的防腐行为	601
11.7 定量预测量的一阶模型	468	第14章 试验设计的方差分析	606
11.8 定量预测量的交互作用模型	478	活动中的统计学: 房地产开发中的污染物——一个小样本情况下处理不当的案例	606
11.9 定量预测量的二阶(二次)模型	482	14.1 引言	606
11.10 回归残差和异常值	489	14.2 方差分析中的逻辑	606
11.11 某些陷阱: 可估性、多重共线性和外推	500	14.3 单因子完全随机化设计	608
11.12 多重回归分析步骤小结	507	14.4 随机化区组设计	618
活动中的统计学回顾: 高速公路建设中的串通投标	507	14.5 双因子析因试验	629
第12章 模型构建	522	*14.6 更复杂的析因设计	645
活动中的统计学: 取消州内货车运输业管制	522	*14.7 套式抽样设计	653
12.1 引言: 为什么模型构建很重要	522	14.8 处理均值的多重比较	662
12.2 自变量的两种类型: 定量的和定性的	523	14.9 检查 ANOVA 假定	668
12.3 一元定量自变量模型	525	活动中的统计学回顾: 房地产开发中的污染物——一个小样本情况下处理不当的案例	671
12.4 二元或多元定量自变量模型	531	第15章 非参数统计	685
*12.5 编码定量自变量	539	活动中的统计学: 新罕布什尔州的地下井水污染如此脆弱	685
12.6 一元定性自变量模型	543	15.1 引言: 分布自由检验	685
		15.2 检验单个总体的位置	686

15.3 比较两个总体: 独立随机样本	691	*16.11 其他抽样计划	770
15.4 比较两个总体: 配对设计	697	*16.12 调优操作	771
15.5 比较三个或更多总体: 完全随机化 设计	704	活动中的统计学回顾: 喷气式飞机燃料添加剂 安全性测试	771
15.6 比较三个或更多总体: 随机化区组 设计	708	第17章 产品和系统的可靠性	781
15.7 非参数回归	711	活动中的统计学: 建立钢筋混凝土桥面恶化的 危险率模型	781
活动中的统计学回顾: 新罕布什尔州的地下 井水污染如此脆弱	717	17.1 引言	781
第16章 统计过程和质量控制	729	17.2 失效时间分布	781
活动中的统计学: 喷气式飞机燃料添加剂 安全性测试	729	17.3 危险率	782
16.1 全面质量管理	730	17.4 寿命试验: 删失抽样	785
16.2 变量控制图	730	17.5 估计指数失效时间分布的参数	786
16.3 均值控制图: $\bar{x}$ 图	735	17.6 估计威布尔失效时间分布的参数	789
16.4 过程变异控制图: R 图	742	17.7 系统可靠性	793
16.5 发现控制图中的趋势: 游程分析	748	活动中的统计学回顾: 建立钢筋混凝土桥面 恶化的危险率模型	797
16.6 不合格品百分率控制图: p 图	749	附录A 矩阵代数	802
16.7 每个个体缺陷数控制图: c 图	754	附录B 有用的统计表	814
16.8 容许限	757	附录C SAS的 Windows 指导	848
*16.9 能力分析	760	附录D MINITAB的 Windows 指导	873
16.10 不合格品的抽样验收	767	附录E SPSS的 Windows 指导	893
		参考文献	911
		部分奇数练习答案	919

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.

活动中的统计学: 田纳西河中鱼的 DDT 污染

化学工厂以及制造工厂经常向附近的河流与溪流中排放有毒的废物, 这些有毒物质对于生存在河流中以及河边的动植物有不利的影响. 其中一种类型的污染物是二氯二苯三氯乙烷, 通常称作 DDT. 在美国, DDT 被用作一种有效的农业杀虫剂. 直到 1972 年, 由于 DDT 的可致癌性, 它才被禁止农业使用. 然而, 因为 DDT 通常是某种生产材料的副产品(如石油馏分油、水-可湿性粉剂和气溶胶), 所以目前对环境仍然有危害.

“活动中的统计学”案例是基于一项检测生活在田纳西河(位于阿拉巴马)及其支流中的鱼类的 DDT 含量的研究. 田纳西河以东西方向流经亚拉巴马州北部, 途经惠勒水库——国家野生动物保护区. 生态学家担心污染的鱼从河流入口洄游到水库, 可能威胁到以鱼为食物的其他野生动物. 这个担心并不是空谈. 在流经印第安克里克人居住区的支流旁边曾经有一座工厂, 支流从河库入口 321 mile (英里, 1 英里 = 1 609. 344 米) 的上游流入田纳西河. 虽然这个工厂已经停产多年, 但是有证据表明工厂向支流中排放的有毒物质污染了相邻区域的所有鱼类.

食品和药物管理局规定一条鱼中 DDT 含量的限量为百万分之五(ppm). 鱼中 DDT 含量超过这个限制被认为是受污染的, 也就是对周边环境有潜在的危害. 田纳西河及其支流中的鱼类是否也被污染了呢? 如果是这样, 污染的鱼向上游洄游了多远?

为了回答这些以及其他问题, 美国陆军工兵部队的成员们在田纳西河以及三条支流(Flint 支流(从河库口 309 mile 的上游流入田纳西河)、Limestone 支流(从河库口 310 mile 的上游流入田纳西河)以及 Spring 支流(从河库口 282 mile 的上游流入田纳西河))的不同地方收集了鱼的样本. 在三条支流的每条支流捕捉 6 条鱼, 沿着田纳西河不同位置(上游英里处)捕捉 126 条鱼, 总共 144 条鱼. 确定每条鱼的重量(g)、长度(cm)以及位置和品种, 然后把鱼切成片, 测量鱼片的 DDT 浓度(百万分之一). 144 条鱼的数据保存在 DDT 文件中.

在本章末尾的“活动中的统计学回顾”中, 我们讨论收集的数据类型和数据收集方法. 本文后面我们会分析这些数据, 以刻画田纳西河中鱼的 DDT 浓度水平, 比较河流不同位置的鱼的 DDT 含量, 并确定对于 DDT 含量的长度和重量(如有必要)的关系.

1.1 统计学: 数据的科学

一位成功的工程师或科学家精通收集信息、评价信息并从中得到结论, 这要求适当的统计学训练. 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3亿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,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**描述统计学**.

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**描述统计学**.

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概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做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,可以将**样本**定义为取自总体的数据子集.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100mL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 左转弯汽车事故比例 肯塔基大学运输研究项目的工程师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.这个估计值将用来为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).工程师用一年多的时间收集了50个没有只许左转弯车道的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. 运输工程师想要做出什么推断?

解 a. 由于工程师收集了50个交叉路口的数据,试验单位是没有只许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的, 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)做出推断. 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 中运输工程师对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.

② 应用练习

1.1 女性的 STEM 经历. 在过去的几十年里, 美国国家科学基金会(NSF)提高了女性在非正式科学、技术、工程或数学(STEM)项目中的参与性. 非正式 STEM 经历带来的影响是什么? 这是已发表的研究 *Cascading Influences: Long-Term Impacts of Informal STEM Experiences for Girls* (2013 年 3 月) 中令人感兴趣的问题. 最近参与了 STEM 项目的 159 位年轻女性被招募完成一份在线调查. 她们当中仅有 27% 的人感觉参与 STEM 项目提高了她们对科学的兴趣.

a. 确定研究者感兴趣的总体.

b. 确定样本.

c. 用这项研究中的信息对相关总体做出推断.

1.2 埋地钢结构的防腐蚀. 被埋在地下的钢结构(如管道)容易受腐蚀. 工程师设计试验测试这种结构受腐蚀的可能性. 在 *Materials Performance* (2013 年 3 月) 中, 比较了两项关于钢腐蚀的称作“即关”和“即开”可能性的试验. 试验用于位于土耳其一个石油化工厂的埋地管道. 在 19 个不同的随机选择的管道位置测量“即关”和“即开”腐蚀性. 这项研究的一个目的是确定当试验被用于埋地钢管道时, 一个试验是否比另一个试验更令人满意(即更准确地预测腐蚀的可能性).

a. 这项研究的试验单位是什么?

b. 描述样本.

c. 描述总体.

d. 这是描述统计学还是推断统计学的例子?

- 1.3 视觉注意力技能测试. 格里芬大学(澳大利亚)的研究人员做了一项研究来判定是否电子游戏玩家比非电子游戏玩家有更强的视觉注意力技能. (*Journal of Articles in Support of the Null Hypothesis*, 6卷第1期, 2009年.) 样本中的65名男生分为电子游戏玩家和非电子游戏玩家两组. 然后, 这两组学生必须完成一系列的视觉注意力任务, 包括“视野”测试. 研究发现两组学生的表现没有差别. 从这项分析中, 研究人员推断“玩电子游戏对于视觉注意力改变影响有限”. 因此, 应用推断统计学得到了这个结论. 确定这项研究的总体和样本.

SWREUSE

- 1.4 软件重用的成功或失败. 由渥太华大学主持的PROMISE软件工程知识库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年.) 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做出推断.

- 1.5 地震的地表运动. 在 *Journal of Earthquake Engineering* (2004年11月)上, 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.

- 1.6 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?

COGAS

- 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年科尔德湾站的空气中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. 定量数据表示事物的数量或个数, 用数值标度量. 例如, 半导体的电源频率(MHz)是一个定量变量, 类似于钢管的断裂强度(lb/in^2 , 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.^①

① 数据类型的更细致划分是名义、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(以及本书中的所有方法), 将数据分为定量的和定性的已经足够了.

定义 1.9 定量数据是按照自然发生的数值度量记录的,即它们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 水管的特性 *Journal of Performance of Constructed Facilities* 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。

- 水管直径(in)。
- 水管材料(钢铁或者 PVC)。
- 水管位置(市中心或者郊区)。
- 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度量的。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 混凝土的性质. *Bulletin of Engineering Geology and the Environment* (69 卷, 2010 年)研究了塞浦路斯自然沙土和混凝沙土的性质。对 20 份土壤样品的每一份观测下列变量。判定每个变量的类型,即定量还是定性。

- 抽样方法(旋转铁心、金属管和塑料管)
- 有效应力水平(牛顿/平方米)
- 阻尼水平(百分比)

1.10 卫星数据库. 美国忧思科学家联盟(Union for Concerned Scientists, UCS)维护卫星数据库,记录了超过 1 000 个目前正在绕地球运转的卫星的数据。其中一些变量存储在数据库中,包括使用或所属国家、基本使用(民用、商用、政府或军事)、轨道类型(近地轨道、中轨道或同步轨道)、经度位置(度)、最高点(即到地球中心的最远高度)、发射质量(千克)、可用电力(瓦)和期望寿命(年)。观测的哪些变量是定性的,哪些是定量的?

1.11 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 500mL)。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。

- 收集样本的城镇。
- 供水的类型(内陆河、溪、地泉)。
- 酸度(pH 值, 1~14)。

d. 混浊程度(比浊法浊度单位, NTU)。

e. 温度(℃)。

f. 每 100mL 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)。

1.12 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 年 7 月。)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。

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

1.13 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扫描方法——CT

或者X射线中的一种,并且随时记录他们的发展状况。除了所用的扫描方法外,临床学者还记录了每名吸烟者第一次用分配的扫描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?

1.14 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 收集数据:抽样

一旦确定了适合研究问题的数据类型——定量或定性,便需要去收集数据。通常,有三种不同的途径获得数据:

1. 数据来自发表的资料
2. 数据来自试验设计
3. 数据来自观察研究(如问卷调查)

有时,感兴趣的数据集可以在已发表的资料中找到,例如书、期刊、报纸或网站。例如,若交通工程师想调查和汇总美国50个州的交通事故死亡率,则可以通过美国政府每年发布的 *Statistical Abstract of the United States* 获得数据集(和其他很多数据集一样)。现在互联网(World Wide Web)提供了媒介,通过网络很容易获得已发表资料中的数据。

通常,在工程和自然科学中另一种收集数据的方法是试验设计,通过试验设计,研究人员可以在研究的试验单位(人、目标或事件)上施加严格的控制。例如,一项经常被引用的医学研究是阿司匹林预防心脏病的可能性。志愿医生被分成两组——治疗组 and 对照组。在治疗组,每位医生每天吃一片阿司匹林,服用一年。对照组的医生服用一片不含阿司匹林的安慰药片(不含药物),形状和阿司匹林药片一样。研究人员(不是指被研究的医生)控制谁服用阿司匹林(治疗组),谁服用安慰药片。在第13章将学习恰当的试验设计比不进行控制可以从数据中获得更多的信息。

最后,观察研究可以用于收集数据。在一项观察研究中,研究人员观察自然条件下的试验单位,并记录感兴趣的变量。例如,工业工程师观测并记录流水线上工人的生产力水平。不同于试验设计,观察研究是研究人员不控制任何试验单位的特征。问卷调查是一种典型的观察研究形式。研究人员抽取一组人,向每个人问一个或多个问题,并记录他们的回答。

定义 1.11 试验设计是指研究人员在抽样的试验单位特征上施加全部控制条件的一种数据收集方法。这些试验的试验单位被分作治疗组和非治疗组(或对照组)。

定义 1.12 观察研究是指在自然条件下观测抽样的试验单位的一种数据收集方法。抽样的试验单位特征不被控制。(例如民意测验和调查。)

无论使用哪种数据收集方法,数据就是来自总体的一个样本。如果我们想使用推断性统计,就必须有一个代表性样本。

定义 1.13 代表性样本是指能够体现感兴趣的总体所具有的基本特征的样本。

例如,考虑做一个民意测验,估计在所有美国公民中相信全球变暖的公民的百分比。民意测验专家若是基于一组属于绿色和平组织公民的问卷调查数据进行估计,则是不明智的。这样的估计几