

统计学经典译丛

PROBABILITY AND STATISTICS FOR ENGINEERS AND SCIENTISTS

概率与统计

(理工类·第9版)

罗纳德·沃波尔 (Ronald E. Walpole) 雷蒙德·迈尔斯 (Raymond H. Myers)
沙伦·迈尔斯 (Sharon L. Myers) 叶可英 (Keying Ye) 著
袁东学 龙少波 译 贾俊平 主审

STATISTICS CLASSICS

 Pearson

 中国人民大学出版社



STATISTICS CLASSICS 统计学经典译丛



PROBABILITY AND STATISTICS FOR ENGINEERS AND SCIENTISTS

概率与统计

(理工类·第9版)

罗纳德·沃波尔 (Ronald E. Walpole)
雷蒙德·迈尔斯 (Raymond H. Myers)
沙伦·迈尔斯 (Sharon L. Myers)
叶可英 (Keying Ye)
袁东学 龙少波
贾俊平

著

译

主审

中国人民大学出版社

· 北京 ·

图书在版编目 (CIP) 数据

概率与统计: 理工类: 第9版/罗纳德·沃波尔等著; 袁东学、龙少波译. —北京: 中国人民大学出版社, 2016. 11
(统计学经典教材)
ISBN 978-7-300-23343-7

I. ①概… II. ①罗… ②袁… III. ①概率论 ②数理统计 IV. ①O21

中国版本图书馆 CIP 数据核字 (2016) 第 212972 号

统计学经典教材

概率与统计 (理工类·第9版)

罗纳德·沃波尔
雷蒙德·迈尔斯 著
沙伦·迈尔斯
叶可英
袁东学 龙少波 译
贾俊平 主审
Gailü yu Tongji

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 三河市汇鑫印务有限公司

规 格 185 mm × 260 mm 16 开本

印 张 32 插页 1

字 数 888 000

邮政编码 100080

010-62511770 (质管部)

010-62514148 (门市部)

010-62515275 (盗版举报)

版 次 2016 年 11 月第 1 版

印 次 2016 年 11 月第 1 次印刷

定 价 69.00 元

版权所有 侵权必究

印装差错 负责调换

译者序

本书是由罗纳德·沃波尔等著的一部优秀的统计学教材，其内容系统，结构层次分明，案例丰富。这样的一部教材十分难得，英文版本已更新至第9版。从内容来看，本书深入浅出，注重统计思想，而不偏重理论推导，对数学和概率论基础要求少，更多地强调实用性。此外，本书最为显著和耀眼的特点在于，其大部分章节都包含旨在建立起与本书其他章节之间联系的相关讨论。如此一来，读者更容易从宏观上认识统计的全貌，这正是本版教材的用心之处，也是其显著有别于本书历史版本之处。因此，本书适合统计学和相关专业的本科和研究生，尤其适合需要获得统计学的全景的读者。

从发展的角度来看，“统计是动态的历史，历史是静态的统计”。这说明统计理论本身是不断发展和完善的，是一个持续的过程，而且其产生和发展还与生产和社会历史进步密切联系，依赖于特定的社会历史条件（或实践），受到它的影响和制约（或激励）。例如，在农业和工业试验中，人们对试验设计产生了需求，于是催生了这个分支的发展，进一步在工业生产中又对质量控制和可靠性提出了要求。眼下大数据（big data）这个概念风行于世，姑且不论其内涵和外延，试问大数据是现在才有的吗？古圣先贤就没想过要“变废为宝”吗？技术进步，尤其是计算机技术的进步，使得这种想法成为可能，而且人们也认识到了大数据的重要性，于是从应用层面对处理大数据的方法提出了要求。我们可以看到，在生产和社会历史的实践过程中，普遍的规律是，人们常常需要通过个体（样本）来认识一般现象（总体），正是人类的现实需求和认识论意义上的这种天性催生了统计（思维）的发展，于是以归纳总结范式来认识世界的认知逻辑构成了统计学的基础逻辑，可见，统计学并不是在思维层面上具有纯粹性的数学演绎。但在理论层面，统计学将基础建立在大数定律和中心极限定理之上，确保了从个体（样本）到一般（总体）在一定基础（概率基础）上的正确性。因此，围绕大数定律和中心极限定理进行演绎的概率论便构成统计研究的基础，本书第2~6章以及第7章的部分内容就属于概率论的范畴。

我们开展统计研究来认识一般现象（总体）的过程中，主要面临两大统计问题：一个是参数估计问题（第8章），另一个是假设检验问题（第9章）。本书在内容上对这两部分的内容进行了充分的介绍，并且在结构安排上充分考虑了这两个问题之间的关联。读者在学习这两部分时需要掌握点估计、无偏估计、区间估计、置信水平、预测区间、容忍限、零假设（或原假设）、备择假设、犯第一类错误的概率 α 、犯第二类错误的概率 β 、 P 值（或显著性）、自由度、检验的势等基础性概念的思想和内涵，前后联系来理解参数估计和假设检验两者之间的异同，从而对统计思想、概念和方法有更深刻、更准确的认识。例如，置信区间、预测区间和容忍区间这三者之间的差异及其产生的原因、所基于的假设以及在何种类型的科学研究或科学问题中应当采用何种区间形式。又如，零假设和备择假设之间，何者更为重要？当我们拒绝了备择假设时，是否可以说“接受零假设”这样的话，为什么？这些内容需要读者根据相应的统计思想去体会。

上面提到，称为“统计方法”的所有方法的主题在很大程度上几乎都是围绕估计和假

设检验问题而展开的，只不过第9章以后的内容在很大程度上要依赖于我们的统计建模。在许多科研和工程领域，人们需要对很多不同类型的实际问题去进行建模，从而挖掘变量或参数之间的关系，统计上将这种关系定义为相关关系。不过，统计模型的框架通常用处不大，除非可以获取用于估计模型参数的数据。第10章和第11章介绍回归模型时这一点就会非常清楚。而且，与第8章有关的一些概念和理论还要进一步推广。第9章所讨论到的假设检验的框架、 P 值、检验的势以及对样本容量的选取这些内容，在一起还将发挥非常重要的作用。在第10章和第11章，读者需要根据简单线性回归中的回归线和拟合回归线等基本概念（尤其是相关性的概念）以及回归分析所基于的基础假设来学习，因此也就可以理解为何要采用最小二乘法来估计模型的参数。估计出参数之后我们是否就有信心将该模型用于加深我们对工业生产的理解或者预测？不然，原始模型表达式通常还需要通过拟合优度检验（如拟合不足检验等）、残差检验（如异方差）等方法来进行统计诊断，以判别是否存在违背基础假设的情况，并对模型质量进行评估，借此再对模型进行修正。

现实情形中我们所需估计和检验的总体参数通常不止两个，此时，第8章和第9章中的方法就显得窘迫了，我们需要采用方差分析技术或开展因子试验来解决这一类问题。可见，第12章是全书极为关键的一章，它是试验设计和方差分析这类重要问题的入门基础。试验设计中，随机指派消除了由于系统性指派造成的偏差，并将试验因子的不一致性造成的风险均匀地分配给每一个因子水平，从而最好地模拟出模型的假设条件。在第12章，本书还通过在试验中进行区组化，将试验单元分成一个个称为区组的相似对，因子水平或处理则随机地指派给每个区组，从而减少实验误差所造成的影响。在本章中，本书将相似对自然地推广到规模更大的区组上，而方差分析就是我们进行分析的主要工具。

第13章与第12章所讨论的是一样的问题，它只是推广到了包含一个以上因子的情形，不过，对这类问题的分析难度会因为我们对各因子之间所存在交互的解释而加大。而且在有些时候，科学实验中的交互效应相对于主因子（主效应）而言所起的作用更加重要。交互效应的存在使得我们要更加重视对图形方法的利用。在第13章和第14章，由于因子组合的方式繁多，因此本书就随机化方法进行了更详细的讨论。在阅读试验设计这三章的内容时，读者需要找到其中的关联，并认识到随机化的重要性以及交互效应对结果解释所起的作用等重要问题。虽然第14章有两方面的内容基本上就是第12章和第13章这两章的基本原则的一个推广，但在因子试验中我们是采用回归分析的方法来解决问题的，因为假设绝大多数因子都是定量的，而且都是在连续统上进行测度的。预测方程是根据设计性试验的数据得到的，于是我们可以将其用于工艺改进，甚至是工艺过程优化。此外，本书还提出了部分因子试验，即只对全因子试验中的部分因子进行试验，因为开展全因子试验往往存在过高的成本。

本译本有三处删节需要向读者说明。第一，本书的英文原版为读者提供了丰富的习题，出版社从教材本身的体量考虑，在习题部分只保留了编号为奇数的习题，不过即便是这样，本版教材的练习题在规模上仍然是非常庞大的。第二，本书将原版教材中不影响全书完整性的可选章节也进行了删减。第三，考虑到计算机在教学中的普遍应用，包括Excel在内的多款软件都可以直接计算出均值等重要统计信息，而且使用软件是现代统计学习中不可或缺的一项技能，因此本书对附录A中的内容并未收录，而是选择放在网上（www.rdjg.com），读者可自行下载。此外，为了方便我们使用软件来解决统计问题，英文原版的作者还在他们的网站上提供了原书中所有例子和习题的数据集，请读者自行到指定网站下载。

限于译者水平，译文中定有不当之处，欢迎专家和读者批评指正。

袁东学

前言

一般方法与教学水平要求

我们编纂第9版更多地不在于增加新的内容，而是要让本书清晰易懂。通过各章末新增内容所建立起的章节之间的关联关系，我们部分地达成了这一目标。我们将章末的这些评述称作 Pot Holes。因为这些内容非常有益于读者构建起本书的大框架，找到各章内容在这个框架中的位置，并让读者认识到其中的局限及对方法的误用所可能落入的陷阱。此外，通过我们在某几个章节中加入的课堂作业，读者可以更深入地了解在现实世界中是如何使用统计方法的。读者可以独自或者分组去采集试验数据并进行推断。在某些例题中，对问题的解答将极好地阐明某个概念的内涵，或给出对某个重要的统计结果的经验解释。我们不仅对已有的一些例子进行了推广，还以案例研究的方式引入了一些新的例子，我们在其中所做的评述都是为了使读者清晰地理解现实情境中的统计概念。

在本版中，我们会继续在理论和应用两者之间选择一个平衡点。本版所用到的微积分以及其他数学基础（例如，线性代数）与上一版处在大致相同的水平上。在我们将探讨的对象锁定为概率中的法则与概念时，统计学中分析工具的适用范围随着对微积分的应用就扩大了。第2~9章我们会着重讲解概率分布和统计推断。第10~14章会稍微用到线性代数和矩阵的知识来主要讲解线性回归和方差分析。因此，学习这部分内容的读者应当具备一个学期的微分学和积分学的知识储备。线性代数的知识非常有用，却并非必需的，如果教师不把第11章中用到矩阵代数知识的多元线性回归这一小节的内容放入教学大纲的话。和之前的几个版本一样，我们给出了大量的习题，涉及科研和工程的现实问题，需要读者前去挑战。习题中所涉及的许多数据集读者都可以到如下网址下载：<http://www.pearsonhighered.com/datasets>。

第9版变化概览

- 在几个章节中加入了课堂作业，读者可以更深入地了解在现实世界中是如何使用统计方法的。我们要求读者去得到或采集他们自己的试验数据，并据此进行推断。

- 在本版中我们加入了更多的案例研究以及其他一些扩展的例题，帮助读者去理解现实情境中用到的统计方法。例如，现实情境中对置信限、预测限、容忍限的解释。

- 我们在某些章节末尾添加了 Pot Holes，在某些章节对其进行了扩展。我们做出这些评论的目的在于让读者找到各个章节在本书大框架中的位置，并理解各章之间的关联。此外，在这部分内容中我们还会指出一些可能误用本章所展示统计方法的地方，并提请注意。

- 第1章增加了更多有关单值统计量及图形方法的内容，而且对于抽样和试验设计也给出了

新的基础材料。

- 第7章添加的有关抽样分布的例子,旨在说明 P 值和假设检验。这样有助于读者做好充分的准备来接受这些在第9章讲述到的、具有难度的主题。
- 第11章进一步提出了模型中单个回归变量与其他变量之间存在严重共线性的问题。
- 第14章介绍了响应面法这一重要的主题。其中使用的噪声变量有助于我们以图形方式来呈现均值和方差(对偶响应面)建模。
- 第14章介绍了中心组合(合成)设计。

内容和教学计划

本书适合在一个学期或两个学期学完。一个学期的教学计划包括第1~9章的内容会更合理。这样一来,这门课的课程计划中就可以包含估计和假设检验这些基础性内容。如果教师期望你能够学到简单线性回归的内容,则可能会要求你继续学习第10章中的部分内容。如果教师期望你能够学到方差分析而非回归,那么一个学期的教学计划中就会包括第12章,而不是第10~11章。第12章主要讲述单因素方差分析。另一个选择是,删除第5~6章的部分内容。如果这样选择的话,需要删除第5~6章中的一个或多个离散型或连续型分布,包括负二项分布、几何分布、伽马分布、韦布尔分布、贝塔分布、对数正态分布。在一个学期的课程计划中,还可以考虑删除极大似然估计、预测、第8章中的容忍限。我们可以看到,一个学期的课程安排具有内在的灵活性,关键取决于教师对回归、方差分析、试验设计、响应面法中的哪些内容更感兴趣。不过,有几个离散型分布和连续型分布(第5~6章)在工程和科研领域中有着广泛的应用。

第10~16章中的大量内容都可以放在有两个学期课程安排的第二学期。第10~11章中介绍的是简单线性回归和多元线性回归的内容。不过,第11章中的内容具有很大的灵活性。多元线性回归还包括分类变量或示性变量、诸如进行模型选择的逐步回归这样的序贯方法、研究残差以检测违背假设的问题、交叉验证和对统计量 $PRESS$ 和 C_p 的应用、逻辑回归这样的“特殊主题”。可以着重介绍对正交回归因子——这是对第14章所介绍的试验设计的初级形式——的应用。第12~13章中的大量内容都是关于具有固定效应模型、随机效应模型、混合效应模型的方差分析的。第14章着重介绍全因子试验和部分因子试验中对两水平设计的应用,其中说明了筛选设计这个特殊的问题。第14章还有一个特点,在新的一节中使用响应面法来说明如何应用试验设计去求解最优的过程条件(或工艺条件),并讨论了如何使用中心组合(合成)设计来拟合二阶模型。响应面法因此也被扩展为包含对稳健参数设计这类问题的分析。噪声变量则被用来支撑对偶响应面模型。第15~16章适度地对非参数统计、质量控制的内容进行了介绍。

第1章从简单的数学层面对统计推断进行了概括性的介绍。这一章的内容从第8版扩展到更全面的包含单值统计量和图形方法这些内容。本章的宗旨是要为读者初步构建起基本概念,以助力读者对后续内容能有更深入的理解。其中包括抽样、数据采集、试验设计的基本概念,也介绍了图形工具的一些基本方面,使读者获得对数据集中所蕴藏信息的一个认识。此外,还增加了茎叶图和箱线图。图形进行了系统化,并做了标记。对科研问题中的不确定性和变差我们进行了全面深入的探讨和说明。其中我们提供了如何从所研究的过程或科研问题中整理出重要特征的一些例子,在诸如生产过程、生物医学研究、生物学研究和其他科研问题这些实际情境中对这些思想进行了说明,而且对离散型数据和连续型数据的应用进行了对比。本章的重点在于介绍如何使用模型,以及可以通过图形工具获取的关于统计模型的信息。

第2~4章介绍的是基础概率、离散型随机变量和连续型随机变量的内容。第5~6章着重介绍

了某些特殊的离散型分布和连续型分布,以及它们之间的关联。在这些章节中我们同样重点列举了现实科研和工程研究中所使用分布的例题。其中给出了例子、案例研究、大量的习题来启发读者如何去使用这些分布,而且通过课堂作业让读者体会到如何将这些分布应用到现实生活中。第7章包含对图形工具的介绍,不过这些都是对我们在第1章所呈现和阐述的基本图形工具的扩展。而且我们采用例题对概率图进行了探讨与说明,并深入介绍了抽样分布这个重要的概念,还阐述了中心极限定理以及独立正态抽样情境中样本方差的分布。其中也介绍了 t 分布和 F 分布,以引出我们在后续章节对它们的应用。在第7章中新增的内容有助于读者认识到假设检验的重要性,以引出 P 值的概念。

第8章包含单样本和两样本的点估计和区间估计问题,而且我们通过例证对不同类型的区间——置信区间、预测区间和容忍区间——之间的区别进行了深入讨论。我们还通过生产情境中的一个案例研究对这三类统计区间进行了阐述。在该案例研究中,我们着重说明了各区间之间的差异,其产生的原因、所基于的假设,以及在何种类型的科学研究或科学问题中应当采用何种区间形式。此外,本章还新增了一个关于比例推断的近似方法。在第9章开篇,我们通过一个简单的例子就假设检验的实际内涵进行了说明,并重点介绍了诸如零假设和备择假设、概率的作用与 P 值、检验的势这些基本概念。随后给出了标准情形中单样本和两样本检验的例证。当然,其中也介绍了配对观测的两样本 t 检验。案例研究有助于读者形成对因子之间交互效应的真实含义的清晰认识,以及在处理和试验单元之间存在交互效应的情形中,清晰认识到其所可能出现的危险。第9章最后一节是非常重要的一节,它将第8章和第9章(估计与假设检验)与主要讲述统计模型的第10~15章联系了起来。因此,对于读者而言,清晰地认识到其中的强关联非常重要。

第10~11章是简单线性回归和多元线性回归的内容。本书用了更多的笔墨对回归变量之间多重共线性带来的影响进行讨论。就单个回归变量与模型中的因变量之间存在很大程度上的相关性的情形,我们在书中说明了它的影响。针对(多重共线性)这个概念,书中重新回顾了序贯模型筛选法(前向法、后向法、逐步法等),并给出了每种方法所使用 P 值的合理数值。第11章主要介绍非线性模型,其中特别介绍了逻辑回归,它在工程学和生物学中具有很多应用。可见,书中有极其多的内容都在介绍多元回归,这为教授这门课程的教师留下了极大的灵活度。第11章最后一节是将其与第13~14章联系起来的评价值内容。本书在内容上突出了几个新的特点,有助于从总体上对此有更好的理解。比如,大部分章的最后一节都会提醒读者可能会遇到的所需注意的地方和难点。值得一提的是,实践中存在的几种类型的响应变量(如比例响应变量、计数响应变量、其他几种响应变量)是不能使用标准的最小二乘估计的,因为标准的假设已不再成立,而违背假设则可能引起严重的误差。在某些情形中,对响应变量所做的数据变换可以缓解这一困境。第12~13章的主题是方差分析,这两章对于教师来说也具有相应的灵活度。第12章的内容是完全随机化设计中的单因素方差分析,并以对方差的检验与多重比较相补充。其中以在区组中对处理进行比较为重点,并讨论随机化的全区组。图形方法也拓展到了方差分析中,以形象化的推断形式辅助读者进行正式的统计推断,这样可以基于读者在本书中所学内容来极大地助力其科学和工程研究。在这里我们也为读者新增了课堂作业,读者需要采用恰当的随机化方式混合到每个方案/计划中,并应用图形技术和 P 值来作答。第13章将第12章中的内容扩展成因子结构中包含3个或多个因子在内的情形。第13章中的方差分析既包括随机效应模型,又包含固定效应模型。第14章的内容为 2^k 部分因子设计;本书以例题和案例研究的形式为读者呈现了如何应用筛选设计和 2^k 某些特别的高阶的部分因子设计。本章新增的两项内容是响应面法和稳健参数设计。本书将上述主题通过一个案例研究连接了起来,该案例研究表达和说明的是一个对偶响应面设计与分析,其特点是要使用到过程均值和方差响应面。

■ 电脑软件

本书自第7章的案例研究开始使用 SAS 和 MINITAB 两者所输出的分析结果及其所产生的图形素材。本书将电脑软件的内容包含其中，源于我们坚信读者是具备电脑使用经验的，且可以读懂电脑软件的输出结果与图形，即使你的老师所使用的软件不是本书所采用的软件。对于读者而言，多接触几种软件可以拓展自身的经验基础。而且不要期望在本门课程的学习之中使用到的软件，会是毕业之后实务操作中所要求你们应用的那款软件。本书在适当的地方为其中的例子和案例研究都配有不同类型的残差图、分位数图、正态概率图，以及其他种类的图形。第10~14章就大量使用了这些种类的图形。

■ 补充材料

《教师参考答案手册》：该手册中给出了所有习题的答案，可从培生教育教师资源中心（Pearson Education's Instructor Resource Center）获取。

《学生参考答案手册》：该手册给出了选中习题的完整答案，它是读者学习和解决难题的重要工具书。

PPT 课件：该课件中包含本书中绝大多数图形和表格，可从培生教育教师资源中心获取。

StatCrunch 电子文本：该电子文本是交互式的在线系统，其中包含 StatCrunch，它是基于网络的一款强大的统计软件。内嵌的 StatCrunch 按钮有助于用户一点击该按钮就能打开本书中的所有数据集和表格，并立即使用 StatCrunch 进行数据分析。

StatCrunch™：StatCrunch 是基于网络的统计软件，它有助于用户进行复杂的分析、分享数据集、根据他们的数据生产出令人信服的报告。用户可以上传自己的数据到 StatCrunch，也可以在由超过 12 000 个公之于众的共享数据集构成的数据库中搜寻数据，该数据库几乎涵盖每一个领域。交互式的图形输出结果有助于用户深入理解统计概念，并且通过获取到的交互式图形输出以可视化数据的形式丰富我们的报告。此外，它还有其他一些特点：

- 大量的数量方法和图形工具，为用户分析并深入了解每个数据集提供了途径。
- 报告选项，为用户生产众多可视化的数据表现形式提供了便利。
- 在线调查工具，为用户通过网络形式快速进行调查并管理这些调查提供了路径。

取得资格的用户可以获取 StatCrunch 的接口。更多详情，请访问网站 www.statcrunch.com。

致 谢

在此我们要向参与审阅本书之前各个版本，并为本版提供良多有益建议的同行致意。他们是迈阿密大学的 David Groggel、拉里坦谷社区学院的 Lance Hemlow、得克萨斯大学圣安东尼奥分校的 Ying Ji、北艾奥瓦大学的 Thomas Kline、罗格斯大学的 Sheila Lawrence、布鲁姆社区学院的 Luis Moreno、科罗拉多大学波德分校的 Donald Waldman、斯伯汀大学的 Marlene Will。我们还要感谢为本书提供帮助的米勒斯维尔大学的 Delray Schulz、后京学院的 Roxane Burrows 和 Frank Chmely。

感谢培生为本书的出版而工作的编辑和印刷人员，尤其是总编辑 Deirdre Lynch、策划编辑 Christopher Cummings、执行编辑 Christine O'Brien、制作编辑 Tracy Patruno、文字编辑 Sally Lifland。非常感谢为本书提供颇多有益意见和建议的校对员 Gail Magin。最后，我们要感谢弗吉尼亚理工大学统计咨询中心，该中心为本书提供了有关现实生活的丰富数据实例。

目 录

第1章 统计与数据分析导言	1	5.4 负二项分布和几何分布	109
1.1 概述：统计推断、样本、总体及概率的 作用	1	5.5 泊松分布和泊松过程	111
1.2 抽样过程与数据采集	4	5.6 可能的错误观点及危害；与其他章节的 联系	117
1.3 位置测度：样本均值和中位数	7	第6章 几个连续型概率分布	118
1.4 变异的测度	9	6.1 连续型均匀分布	118
1.5 离散型和连续型数据	11	6.2 正态分布	119
1.6 统计建模、科学检验以及图形诊断	11	6.3 正态曲线下的面积	121
1.7 统计研究的一般类型：设计性试验、 观测性研究以及回溯性研究	18	6.4 正态分布的应用	126
第2章 概 率	22	6.5 二项分布的正态近似	131
2.1 样本空间	22	6.6 伽马分布和指数分布	136
2.2 事件	24	6.7 卡方分布	141
2.3 样本点计数	28	6.8 贝塔分布	141
2.4 事件的概率	33	6.9 对数正态分布	142
2.5 加法法则	35	6.10 韦布尔分布	143
2.6 条件概率、独立性及乘积法则	39	6.11 可能的错误观点及危害；与其他 章节的联系	147
2.7 贝叶斯准则	45	第7章 基本的抽样分布和描述性数据 分析	149
2.8 可能的错误观点及危害；与其他章节的 联系	51	7.1 随机抽样	149
第3章 随机变量和概率分布	53	7.2 一些重要的统计量	150
3.1 随机变量的概念	53	7.3 抽样分布	153
3.2 离散型概率分布	55	7.4 均值的抽样分布和中心极限定理	154
3.3 连续型概率分布	58	7.5 S^2 的抽样分布	160
3.4 联合概率分布	62	7.6 t 分布	163
3.5 可能的错误观点及危害；与其他章节的 联系	73	7.7 F 分布	166
第4章 数学期望	75	7.8 分位图和概率图	169
4.1 随机变量的均值	75	7.9 可能的错误观点及危害；与其他章节的 联系	174
4.2 随机变量的方差和协方差	80	第8章 单样本和两样本的估计问题	175
4.3 随机变量的线性组合的均值和方差	87	8.1 引言	175
4.4 切贝雪夫定理	93	8.2 统计推断	175
4.5 可能的错误观点及危害；与其他章节的 联系	97	8.3 经典估计方法	175
第5章 几个离散型概率分布	98	8.4 单样本：均值估计	178
5.1 引言与动机	98	8.5 点估计的标准误	183
5.2 二项分布和多项式分布	98	8.6 预测区间	184
5.3 超几何分布	104	8.7 容忍限	186
		8.8 两样本：两个均值之差的估计	188

8.9 配对观测	193	11.1 引言	303
8.10 单样本：比例的估计	196	11.2 系数的估计	303
8.11 两样本：两个比例之差的估计	199	11.3 矩阵形式的线性回归模型	307
8.12 单样本：方差估计	201	11.4 最小二乘估计量的性质	311
8.13 两样本：两个方差之比的估计	203	11.5 多元线性回归中的推断	312
8.14 可能的错误观点及危害；与其他 章节的联系	207	11.6 通过假设检验来选取拟合模型	317
第9章 单样本和两样本假设检验	209	11.7 分类变量或示性变量	321
9.1 统计假设：基本概念	209	11.8 模型选择的序贯方法	324
9.2 统计假设的检验	210	11.9 残差问题及违背假设的情况 (模型检验)	329
9.3 利用 P 值在假设检验中进行决策	218	11.10 交叉验证、 C_p 以及模型选择的其他 准则	333
9.4 单样本：与单个均值相关的检验	220	11.11 非理想条件下特殊的非线性模型	340
9.5 两样本：两个均值的检验问题	226	11.12 可能的错误观点及危害；与其他 章节的联系	346
9.6 均值检验问题中样本容量的选取	232	第12章 单因子试验：一般性问题	347
9.7 均值比较的图形方法	235	12.1 方差分析技术	347
9.8 单样本：单个比例的检验	238	12.2 试验设计策略	348
9.9 两样本：两个比例的检验问题	240	12.3 单边方差分析：完全随机化设计	348
9.10 方差的单样本和两样本检验问题	242	12.4 方差齐性检验	354
9.11 拟合优度检验	245	12.5 自由度为1的对照	357
9.12 独立性检验（分类数据）	248	12.6 多重比较	360
9.13 齐性检验	250	12.7 对区组中的处理进行比较的问题	366
9.14 两样本的案例研究	252	12.8 完全随机区组化设计	366
9.15 可能的错误观点及危害；与其他 章节的联系	257	12.9 图形方法与模型诊断	372
第10章 简单线性回归和相关性	259	12.10 方差分析中的数据变换	375
10.1 线性回归简介	259	12.11 随机效应模型	377
10.2 简单线性回归模型	260	12.12 案例研究	380
10.3 最小二乘和拟合模型	263	12.13 可能的错误观点及危害；与其他 章节的联系	386
10.4 最小二乘估计的性质	267	第13章 因子试验（两因子或多因子）	387
10.5 回归系数的推断	269	13.1 引言	387
10.6 预测	274	13.2 两因子试验中的交互效应	388
10.7 对回归模型的选取	277	13.3 两因子方差分析	389
10.8 方差分析	278	13.4 三因子试验	400
10.9 回归问题中线性性的检验：具有 重复观测的数据	280	13.5 随机效应和混合效应模型的因子 试验	407
10.10 数据的图形和变换	285	13.6 可能的错误观点及危害；与其他 章节的联系	414
10.11 简单线性回归的案例研究	289	第14章 2^k 因子试验和部分因子试验	415
10.12 相关性	292	14.1 引言	415
10.13 可能的错误观点及危害；与其他 章节的联系	301	14.2 2^k 因子试验：效应的估计和方差	
第11章 多元线性回归和一些非线性回归 模型	303		

分析.....	416	15.2 符号秩检验	457
14.3 无重复 2^k 因子试验	420	15.3 Wilcoxon 秩和检验	460
14.4 回归中的因子试验	425	15.4 Kruskal-Wallis 检验	463
14.5 正交设计	428	15.5 游程检验	465
14.6 部分因子试验	435	15.6 容忍限	467
14.7 对部分因子试验的分析	440	15.7 秩相关系数	467
14.8 高阶部分因子试验和筛选设计.....	443	第 16 章 统计质量控制	471
14.9 以 8, 16, 32 个设计点构造解析度为 III 和 IV 的设计	443	16.1 引言.....	471
14.10 解析度为 III 的其他两水平设计; Plackett-Burman 设计	445	16.2 控制限的性质	472
14.11 稳健参数设计	445	16.3 控制图的用途	472
14.12 可能的错误观点及危害; 与其他 章节的联系	453	16.4 变量型控制图	472
第 15 章 非参数统计.....	454	16.5 属性控制图	484
15.1 非参数检验	454	16.6 累积和控制图	489
		参考文献	493

第1章 统计与数据分析导言

1.1 概述：统计推断、样本、总体及概率的作用

质量提升自20世纪80年代就是，而且到21世纪仍是美国工业大量关注的焦点，但日本自20世纪中叶就创造了被大量谈及并见诸文献的“工业奇迹”。美国及其他国家都未能使生产高质量产品的环境出现，而日本人却成功了。事实上，他们的大部分成功都要归因于管理层对统计方法和统计思维的应用。

1.1.1 科学资料的应用

在制造、食品开发、电脑软件、能源、制药以及其他领域，应用统计方法时都需要进行信息或科学资料的采集。数据收集已经有一千多年的历史，并不是什么新的东西。数据一般会经收集、汇总、报告及存储等程序，以便查阅。然而，数据收集和推断统计之间却有着巨大的差别，后者在近几十年的时间里才受到应有的重视。

推断统计就是统计从业者所使用的统计方法的一个大“工具箱”。这些统计方法都是为了不确定性和变异出现时的科学决策过程而开发的。生产过程中，特定材料的产品密度并非总是相同的。如果该生产过程是分批次非连续进行的，那么生产线上下来的材料密度就不仅存在各批之间的差异（批间差异），而且会出现批内差异。使用统计方法就是为了分析诸如该过程中的数据资料，从而获知更多可用以提升生产过程质量的优化信息。在该生产过程中，质量就可以以对某个目标密度值的接近程度进行严格定义，而且任何时间都要满足该接近程度准则。大气污染过程研究中，工程师可能会关注用以测量一氧化硫的某特定工具（或方法）的情况。如果该工程师怀疑该工具的有效性，那么其中就会有两类待确认的变异来源。第一类是一天内同一地点一氧化硫数值的变异，第二类是大气中一氧化硫的观测值和真实值在特定时间的变异。这两类来源中的任何一类波动过大（根据该工程师设定的某个标准），就需要更换测量工具（或方法）。在降压新药的生物医学研究中，有85%的患者症状得到缓解，现有药物或旧药则能缓解80%慢性高血压患者的症状。不过，该新药却可能会带来一些副作用，并且生产起来更昂贵。那么是否应该批准新药呢？这是制药厂和美国食品药品监督管理局（FDA）经常遇到的问题（问题通常更复杂）。这时还是需要产生变异的因素进行考察。“85%”这个数值是在该研究所选定患者的基础上得到的，一旦以新患者重复该研究，观测到的则可能会是只有75%的患者症状得到缓解。因此，决策过程中，应当考虑到各次研究之间非人为因素的变异。显然，这个变异很重要，因为在（重复性研究）这样的问题中患者的变化是很常见的。

1.1.2 科学资料中的变异性

上面讨论的问题中所用的统计方法涉及对变异的处理，而我们研究的每种情形中的变异都会科学资料中遇到。如果生产过程中，观测到的产品密度始终如一，并且等于目标值，那么就不需要统计方法。如果用以测量一氧化硫的设备也总是有相同的数值，并且还是精确值（即正确的真实值），那么也不需要统计分析。如果不同患者本身对药物的反应没有差异（即药物能缓解每位患者的症状，或者对所有患者的症状都不能缓解），那么对制药厂的科研人员和美国食品药品监督管理局来说事情就简单多了，决策过程中也不需要统计学家。统计学研究为分析诸如上述过程中的数据提供了大量的方法。这反映了统计推断这门学科的真实本质，即让我们有方

法可用,从而超越纯粹的数据报告,得出有关应用科学体系^①的结论(或进行推断)。统计学家都是凭借概率和统计推断的基本定律来得出有关应用科学体系的结论的。数据通常是以**样本**或一系列**观测**的形式进行采集的。抽样过程的内容将在1.2节^②进行阐述,有关这个问题的探讨会贯穿全书。

从**总体**中采集到的样本有可能包含所有个体,也可能只是某一特定类型的个体。有时总体就是一个应用科学体系。比如,电脑主板的制造商希望能消除缺陷,抽样过程可能就需要随机抽取50块电脑主板收集数据。其中,总体就是特定时间内该企业生产的所有电脑主板。假使对电脑主板的生产过程进行工艺改进,然后又抽取了电脑主板的第二个样本,依照样本得到的有关工艺改进效率的所有结论,都可以推广到工艺改进后制造的全部电脑主板的总体。在药物试验中抽取一些患者作为一个样本,让每个人都服用特定的降压药物,这时我们关注的是高血压患者总体服药后的效果。

通常,严格按照计划日程来采集科学资料很重要。有时计划也会不可避免地有很大局限。因此,我们通常只关注总体中这些对象的某些性质和特征。每个特征对试图了解总体的科学家或工程师这样的研究人员来说都具有独特的工程学或生物学上的重要性。比如,在上面的例证中,生产过程的质量与其产出的产品密度有关,那么工程师可能就会去研究生产过程中如温度、湿度、特定配料的量以及其他工艺条件的影响。工程师可以有条理地将这些**因素**调整到满足任一**配方**或**试验设计**要求的任何水平,然而关注某种树木密度的影响因素的森林科学家,却未必能够进行试验设计,此时他需要进行**观测性研究**,即数据可以在现场采集,但**因素水平**不能进行事先选择。这两类研究都要用到统计推断方法。前一种情况下,推断的好坏取决于是否有恰当的试验设计,后一种情况则受制于所能采集到的数据。比如,对研究雨水对作物产量影响的农学家来说,采集到的只有旱季数据是非常令人伤心的。

统计思维对经理人员的重要性和科研人员对统计推断的使用已经形成了广泛的共识。科研工作者从科学资料中获益良多,因为数据增加了他们对科学现象的认知。产品和工艺工程师从线下优化改进工艺的过程中也获益颇丰,而且他们还在常规的生产数据(线上监测)采集过程中提升了洞察力。这些都有助于他们做出必要的修正来保证所需的质量水平。

有时科研工作者仅仅需要样本数据的概要,也就是说无须进行统计推断。而且,单个**统计量**或**描述统计**也是非常有用的。这些统计量能告诉我们有关这组数据中心位置、变异以及该样本中观测的分布特征的信息。尽管没有结合具体的统计方法来进行**统计推断**,但这些统计量中蕴涵着大量的信息。有时,描述统计还会结合图形使用。现代的统计软件包中都有**均值**、**中位数**、**标准差**以及其他单个统计量的计算,而且能生成图形来展示蕴涵样本性质的“踪迹”。对单个统计量与直方图、茎叶图、散点图以及箱线图等图形的定义和说明,放在后续各节。

1.1.3 概率的作用

本书第2~6章阐述了概率的基本概念。对这些概念的充分彻底梳理,有助于读者提升对统计推断的理解。而缺乏概率理论体系的读者则无法对从数据分析到现代统计方法的内容形成真正的理解。因此,在学习统计推断理论之前,掌握概率论的知识很有必要。概率原理有助于我们量化所得结论的说服力或置信度,即概率理论构成了完整的统计方法和资以我们测度统计推断置信度的重要组成部分。因此,概率论这门学科支撑了描述统计向推断统计的发展。而概率基本原理则将结论翻译成科研工作者和工程师所需要的语言。下例有助于读者理解**P**值的内涵,**P**值其实就

^① 根据书中多次出现的场合,应用科学体系/科学(研究)体系具体来说应是一个科学研究的问题或对象。因此,有时我们会直接译作“科研问题/对象”。——译者注

^② 此处原文有误,原文为第2章,从全书的章节安排来看,应为1.2节。——译者注

是解释统计分析结果的“底线”。

例 1.1 假设工程师获得了生产过程的数据，抽取了 100 个样品，有 10 个次品。抽到次品是很正常的，因此抽中的这 100 个样品就构成了我们的样本。然而，从长期运营来看，企业只能忍受生产过程中 5% 的次品率。这时，概率基本原理就有助于工程师来确定关系到生产过程好坏的样本信息到底有多令人信服。显然，该案例中的总体就是生产的全部产品。假设生产过程的工艺水平是可以接受的，即次品率低于 5%，一次随机抽取的 100 个样品中仍然有 0.028 2 的概率包含 10 个及以上次品。这个小概率说明，在以现有工艺水平进行生产的情况下，从长期来看次品率有超过 5% 的可能。换言之，对一个工艺条件可接受的生产过程来说，这样的样本（抽取 100 个样品，有 10 个次品，次品率超过 5%）出现的可能性极小。事实是，刚好发生了这样的事情。显然，如果生产过程的次品率显著超过 5%，那么出现这种情况的概率就会大很多。

从该例我们可以清楚地看到，概率基本原理有助于我们将样本中蕴涵的信息转变成我们对该应用科学体系形成的结论性内容。事实上，这些带有概率意义的结论对工程师和经理来说是一个警醒。统计方法（第 9 章会详细阐述）给出的 0.028 2 的 P 值说明，生产过程的工艺条件很有可能是无法接受的。有关 P 值的问题将在后续章节详细介绍。下例给出了另一个例证。

例 1.2 通常，科学研究的本质决定了概率和演绎推理在统计推断中所起的作用。弗吉尼亚理工学院和弗吉尼亚州立大学研究了树根和菌类之间的关系。菌类把矿物质传送给树，树又把糖分传送给菌类。共 2 个样本，每个样本都包含种植在温室内的 10 株北部红橡幼苗。其中一个样本中的幼苗以含氮物质进行培养，另一个样本的幼苗则不用含氮物质。其他条件并无差别，而且所有幼苗都有豆包菌，140 天之后，植株的茎重（单位：克）数据如表 1.1 所示。

表 1.1 例 1.2 的数据

不含氮	含氮
0.32	0.26
0.53	0.43
0.28	0.47
0.37	0.49
0.47	0.52
0.43	0.75
0.36	0.79
0.42	0.86
0.38	0.62
0.43	0.46

该例中，这 2 个样本来自两个**独立总体**。该试验的目的是研究含氮物质的使用对根部生长情况的影响，并且采用的是比较性研究，即比较两个总体的某一重要特征。以图 1.1 的一维点图形式来展示数据非常具有启发性，其中以 $\circ$ 表示“含氮”样本的数据，以 $\times$ 表示“不含氮”样本的数据。

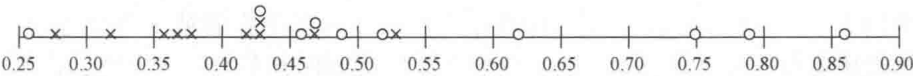


图 1.1 茎重数据的一维点图

我们应该注意到，数据所展现出来的轮廓说明在平均意义上使用含氮物质能够增加茎重。含氮样本中有 4 个观测显著大于不含氮组的每个观测，而含氮样本的大部分观测都低于数据的中

位数。因此，该数据所展示的结果可能说明了氮（对增加根茎生长）的有效性。具体该如何进行量化？我们又该怎样从某种意义上归纳直观上所看到的结果？同前例一样，我们可以应用概率基本原理，即我们可以以概率或 P 值的形式来进行表述。在此我们不会对起概况性作用的某个概率进行统计推断，第9章会对这些方法（包括例1.1中用到的）进行详细介绍。而这里的问题就成了在假设含氮物质（对增加茎重）无效的情况下，也就是说这两个样本都是来自同一个总体的情况下，我们“可以通过样本数据观测到（两个样本来自同一总体）的概率”究竟有多大。如果这个概率非常小，比如是0.03，那么这就成了使用含氮物质确实能影响（显著增加）红橡幼苗平均茎重的有力证据。

1.1.4 概率和统计推断是如何一同起作用的

读者能认清本身就是科学的概率论这门学科和统计推断这门学科之间的区别很重要。正如我们已经说过的那样，使用概率论的概念原理有助于我们对统计推断的结果形成切身的体会。那是因为统计推断使用的就是概率论的概念原理。通过以上两例我们可以看到，研究人员拿到样本信息，运用统计方法和概率基本原理，就能得到有关总体特征的结论（比如，例1.1中的生产工艺过程是不可接受的，而例1.2中含氮物质的确能增加平均茎重）。因此，对一个统计问题而言，在统计推断正确运用概率基本原理的基础上，结合样本我们就能得出有关总体的结论。不过，本质上我们的推理逻辑是归纳法。在进入第2章及以后的章节时，读者要注意，我们不会再像这两个例子那样关注对统计问题的解决，多数例证中间甚至都不包含样本，而是以所有已知特征来对总体进行充分的刻画。我们的重点将放在总体的假定特性上。因此，在已知总体特性的基础上，运用概率基本原理我们就能知道总体假想样本的特征。而这样的推理逻辑在本质上是演绎法。图1.2给出了概率论与统计推断的基本关系。

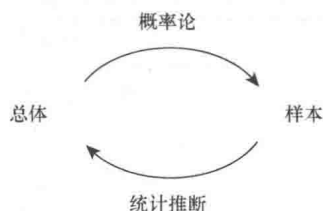


图 1.2 概论率与统计推断的基本关系

从大局着眼，究竟是概率论领域还是统计学领域更重要？显然这两者是相辅相成的，都非常重要。唯一可以确定的是，如果你不满足于仅知如何用，而不涉及太多原理或理论的“使用指南（或手册）”层面的统计学水平，那么一定要先学习概率论。因为除非你知道样本中不确定性产生的基本原理，否则就不能从样本去了解总体。比如例1.1所关注的中心问题是，生产工艺过程所刻画的总体的次品率是否超过5%，即平均意义上100个样品中有5个次品这个假设。而事实是，100个样品中有10个次品。这个现象是支持还是拒绝我们的假设？表面上看起来确实拒绝了该假设，因为100个样品中出现10个次品确实看起来有点儿多了。如果没有概率基本原理，我们又怎么知道这一点？通过后续章节的学习，我们就会知道接受（5%的次品率）该生产工艺过程所应满足的条件。100个样品的样本中出现10个及以上次品的概率为0.0282。

在这两个例子中，我们运用概率基本原理给出了资以科研工作者或工程师进行决策的证据。而数据和结论的桥梁却是建立在统计推断、分布理论以及抽样分布基础之上的，这些内容我们会在后续章节进行阐述。

1.2 抽样过程与数据采集

在1.1节我们简要谈到了抽样和抽样过程的内容。抽样看似一个简单的概念，但所需要解决