

PUTONG TONGJIXUE JIANMING JIAOCHENG

普通统计学简明教程

—— 附TI电脑指令与程序

谢衷洁 编著



北京大学出版社
PEKING UNIVERSITY PRESS

PUTONG TONGJISHU JIANKANG JIAOCHENG

普通统计学简明教程

——附T1电脑指令与程序

周永清 编著



● 中国人民大学出版社
RENMIN DAIXUE CHUBANSHE

普通统计学简明教程

(附 TI 电脑指令与程序)

谢表洁 编著



北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目(CIP)数据

普通统计学简明教程/谢衷洁编著. —北京:北京大学出版社, 2004. 9
ISBN 7-301-07797-1

I. 普… II. 谢… III. 统计学-高等学校-教材 IV. C8

中国版本图书馆 CIP 数据核字(2004)第 090192 号

书 名: 普通统计学简明教程(附 TI 电脑指令与程序)

著作责任者: 谢衷洁 编著

责任编辑: 刘 勇

标准书号: ISBN 7-301-07797-1/O · 0609

出版发行: 北京大学出版社

地 址: 北京市海淀区中关村 北京大学校内 100871

网 址: <http://cbs.pku.edu.cn> 电子信箱: zpup@pup.pku.edu.cn

电 话: 邮购部 62752015 发行部 62750672 理科编辑部 62752021

排版者: 北京高新特打字服务社 51736661

印刷者: 北京大学印刷厂

经 销 者: 新华书店

730mm×980mm 16 开本 22 印张 399 千字 10 页插图

2004 年 9 月第 1 版 2004 年 9 月第 1 次印刷

印 数: 0001—4000 册

定 价: 25.00 元

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,翻版必究

统计学是关于收集和分析数据的 科学和艺术

——《不列颠百科全书》

作者简介

谢表洁 北京大学数学科学学院教授、博士生导师，1959年毕业于北京大学数学力学系，长期从事数理统计的教学和科研工作，取得了丰硕的成果。研究方向是时间序列分析及其应用，尤其是谱分析、建模及小波方法。其应用领域涉及天文学中天王星光环的发现，通讯卫星的交调分析，内分泌激素、脑诱发电位研究，海洋与大陆地球物理及宏观经济与汇率研究等多学科领域。至今已在国内外发表了八十多篇论文，并有六部著作在国内外出版。历任北京大学统计实验室主任、数理统计研究所副所长；中国概率统计学会理事、副秘书长，时间序列专委会主任；是全国政协第八、九、十届委员，北京市政协第七、八届委员、常委。曾获国家自然科学奖三等奖、国家教委科技进步奖一等奖、北京市科技进步二等奖、中国人民解放军军级科技进步奖，并在国外获得第八届、第十四届维也纳控制论与系统研究大会的“最优论文奖”。

内 容 简 介

本书是一部为大学生、中学生讲述统计学基础知识的入门性教科书。全书共分七章,内容包括:开篇引言,绪论,描述性统计学,概率论基础知识,随机变量的分布与数字表征,假设检验,相关分析与回归分析等;各章并配有适量的习题供教师和学生选用。本书作者长期从事数理统计的教学和科研工作,具有丰富的教学经验并取得了丰硕的研究成果。本书以社会上关心的热点和典型问题为开篇引言,以生动、精彩的实例为引线,深入浅出地讲述统计学的基本概念、思想和方法,教给读者如何去收集、分析随机数据,如何运用统计学的方法和技巧,并且基于统计分析的结果给出客观、公正和科学的结论。本书侧重于介绍统计学有关的概念和方法,强调统计学的分析与应用,因此本书中多选用典型实例和真实的数据(它们来自二十多个科学领域)作为统计方法的讨论对象,因而对统计学的教学和科研人员具有很好的参考价值。

本书可作为综合大学、高等师范院校、医学院校、理工科大学及财经类非概率统计各专业大学生的统计课的入门教材,也可供从事数理统计工作的科技工作者、管理者阅读。由于本书深入浅出,不要求读者以学过高等数学为先修课程,并且配备有 TI 电脑指令,因而也适合于广大中学教师作为中学统计学的教学参考书,同时也可作为中学生的课外参考书。



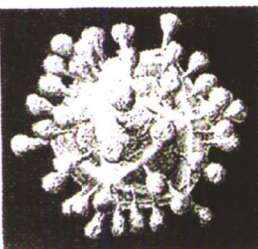
彩图 1 王蒙先生在他的书中，叙述了他在北戴河海滨遇到的一个骗人的游戏，回家后用扑克牌做实验证实了他的观察(有关此游戏的中奖概率的计算，请参阅第三章)



彩图 2 位于几个小区之间的北京大学某超市生意十分兴隆(有关顾客购物额度的规律性描述，请参阅第四章)

科学家如何破译“非典”病毒

【香港《南华早报》4月17日文萃】
 科学家如何破译致命病毒的遗传代码(作者:亚历克斯·骆)
 一种冠状病毒的样本(取自一位死于多伦多的患者的肺部,重量只有百万分之一)在全球对抗非典的战斗中起到关键作用。
 温哥华不列颠哥伦比亚研究局基因组学

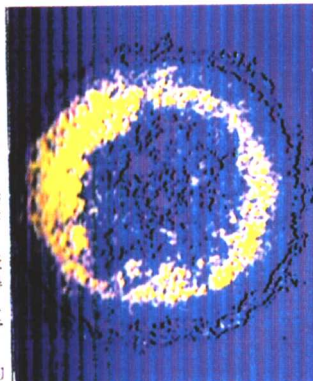


冠状病毒样品

由于基因测序必须使用DNA,这种病毒的RNA首先要转换成DNA。

不列颠哥伦比亚研究局的分子生物学家维克托·林博士说:“RNA的稳定性要差许多,处理起来也难得多。把它转换成DNA花去了我们的大部分时间,差不多四天。”

最初的病毒样本的



彩图3 上右图为艾滋病病毒,左图中间为SARS冠状病毒。科研人员发现冠状病毒感染人体的机制与艾滋病病毒相似^①(有关“非典”病人死亡率的估计问题,请参阅第五章及练习题)



The Santa Cruz Predatory Bird Research Group

DDT and the Peregrine Falcon

Understanding "Bio-accumulation" of Toxins in the Environment

Even the youngest students that we talk to in the third and fourth grade seem to understand what it means to be endangered and that the peregrine falcon and bald eagle were added to the endangered species list largely because of the impact of the pesticide DDT.

However the issue becomes a bit murkier when we try to explain to old HDW it was that DDT caused a population crash among peregrine falcons and other species. To understand students must have a working knowledge of the food chain or food web, which are best illustrated by a pyramid. Young falcons can visit our friend, Quaker, on Stoltz at www.wakinyi.com for games, songs and puzzles that teach young people about the web and predator-prey relationships and the web of life. The peregrine falcon exists at the peak of the pyramid. The broad base of this pyramid represents the many birds that a peregrine falcon feeds upon in a year and lower trophic levels that supply food for the birds—insects and other invertebrates. DDT takes decades to degrade in the soil, stream and river bottoms and in coastal sediments, and so continues to be a reproductive problem for a few birds.

Chemical Structure of DDT



When DDT was made and common use, daily doses of the chemical accumulated in the fatty tissues of the peregrine. The use of DDT was diminished and for some uses banned in 1972. However, residue DDT in the environment today continues to contaminate peregrine falcons. Through a biological coincidence, the stored chemicals acted to "boost" the movement of

彩图4 Santa Cruz加州大学食肉性鸟类研究组关于DDT危害环境和猎鹰的告示(有关DDT中毒的对照实验的结果分析,请参阅第五章)

① 医学家何大一等科研人员发现:冠状病毒感染人体的机制与艾滋病病毒相似。(见埃非社香港2003年5月11日电)



彩图5 许多金发女郎都属高加索人种，她(他)们的眼睛许多是蓝色的(有关人的头发的颜色是否与眼睛虹膜的颜色相关的统计学检验，请参阅第六章)



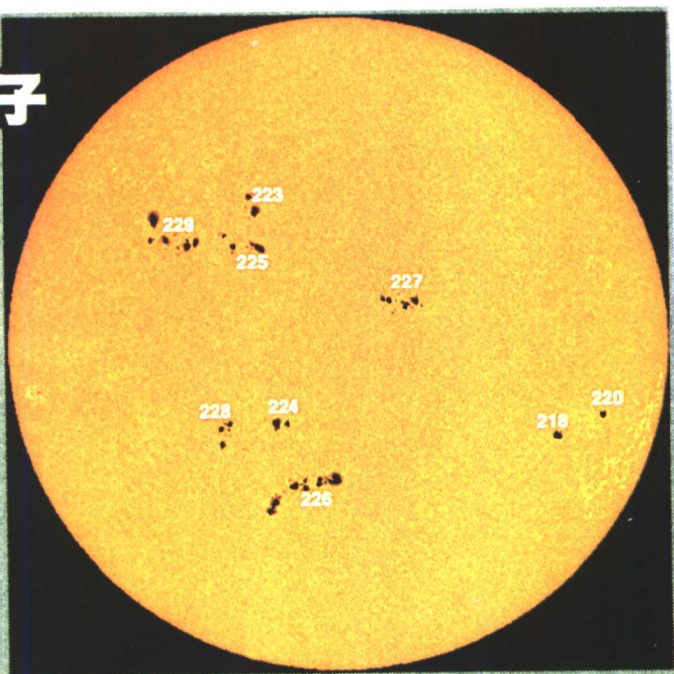
彩图6 图片显示1986年挑战者号航天飞机升空72秒发生爆炸瞬间的场景，7名宇航员全部遇难(有关此次灾难的事故原因分析，请参阅第七章)



彩图7 比萨斜塔位于意大利的比萨市，斜塔由白色大理石砌成，为圆柱形建筑(有关比萨斜塔倾斜量的分析与预报，请参阅第七章)

太陽黑子

美國太空總署發布周二拍攝到的太陽影像，顯示多處太陽黑子，其中編號225、226、227和229這4處會影響太陽表面噴出的日暈，若日暈朝着地球而來，會嚴重干擾地球通訊。



彩图8 所谓太阳黑子，它是由暗黑的本影和在其周围的半影组成，最大的黑子直径比几个地球还大(有关太阳黑子活动周期统计学的检测方法，请参阅第七章)

前 言

20 世纪 90 年代中期的一次北京大学教授会上,当时任副校长的王义道教授指出:目前我校可供文科学生选修的理科课程太少,因为大部分理科课程都是要求学生学过高等数学为前提,而许多文科学生是没学过高等数学的,希望有科研和教学经验的老教授设法开一些不要求高等数学为先修课程,文科学生能接受的高质量的理科课程。在概率统计系同仁们的热情鼓励和帮助下,一门取名为“普通统计学”的全校通选课由我来执教。未曾想到开设之后受到全校文、理科学生的热烈欢迎,最多时 150 人的教室有 300 多人听课,学生来自 20 多个系。一个最大的难题就是听课学生中有数学学得很多的物理系学生,也有只具备中学初等数学的英语、俄语、历史系等数学基础较弱的学生,因此,如何将这门课的讲授,从教材到教学手段都能够吸引住数学背景有很大差异的众多学生就成为一个非常具有挑战性的课题。几年来的实践证明我们的教学是成功的,本课程始终受到文、理科学生的欢迎,上课的积极性和出勤率很高,那怕是在 SARS 期间学校通知学生可以不来课堂上课的情况下仍有 2/3 以上的学生来听课。

我们对本课程教学的指导思想是:

1. 侧重点是统计学的思想和方法,而不是严谨的数学论述和推导。但是书中介绍的统计学的思想和方法却是有扎实的理论基础的。

2. 统计学是和现实生活非常贴近的学科,切忌把本课程变成一门数学课。本书大力收集来自自然科学、工程技术、社会科学、医学、经济与金融等多种学科的实际数据作为提出统计学中不同问题的实例,进而讲解统计学是如何解决这些问题的。由此,学生不仅感受到统计学的实用性,许多学生还会联想到本专业的问题,促进了他们的学习和研究兴趣。

3. 根据文、理科学生的不同需求,在练习和课外阅读材料上适当区别对待。本书有一些带 * 的段落是不属于课堂教学范围的,它们或者是属于数学推导或者属于扩大知识面的内容。有些学生不满足于课堂教学内容的,可在教师指导下阅读这些内容并做一些较难的练习。

4. 鼓励学生应用统计软件和电脑来解题,但前提条件是必须搞清楚有关的统计概念、背景并学会正确的统计方法的使用。当今社会上各种统计软件已相当流行,如 Excel, SAS, SPSS, S-Plus 等以及如 TI 公司出品的 TI-89、TI-92 Plus、V-200 图形计算器等都是学习统计和将统计应用于研究的强有力的工具。我们的学生毕业后,在工作岗位上大部分也可以使用这些 IT 产品。在正确掌握统计的使用方法的基础上,使用这些手段可以大大节约计算时间。学生可以多做题,而且很多图形对学统计的人理解统计的内容和结果是很有帮助的。

本教材就是根据以上的指导思想、在经过前后五年的教学实践的基础上编写成的,为适应不同的教学对象,内容上有适当的扩充。但整体教材的绝大部分内容可在一个学期每周三学时的教学计划中完成。如果教学对象是文科学生,则概率论基础知识部分可以介绍得精练一点。这部分的理论和练习对他们来讲一般都会感到比较吃力。

本教材得到国家自然科学基金研究项目(No. 10171005)及北京市精品教材项目的资助,在此表示谢意。作者还要感谢日本学术振兴会(JSPS)的资助和日本早稻田大学有关方面的大力支持,使我有机会于 2002 年夏在早稻田大学充分利用他们的图书、资料的优势,参阅和收集许多优秀的统计学的多种教科书和数据,并利用其提供的良好条件写出了本书的一些章节,有些内容在国内的图书中是未曾见到的。

最后,作者还要衷心地感谢陈家鼎教授极其细心地阅读了全书,非常准确并富有建设性地指出了原书稿中的一些不足和欠妥之处。没有他的帮助,本书就不会以现在的面貌和读者见面。

然而,限于作者的水平及对陈家鼎教授修改意见理解的不足,本书的缺陷和错误在所难免,恳请同仁们及广大读者给予批评、指正,以便将来再版时给予更正。

谢衷洁

2004 年 3 月于北京大学

关于用 TI 图形计算器进行统计学 教学的实践

1999 年底,我院原副院长郑志明告诉我说美国德州 TI 公司有一产品,据称很适于统计教学,是否有兴趣研究一下如何运用于我院对全校开设的公选课“普通统计学”。我最初拿到的是 TI-89 计算器和一本中文说明书。当时一看该计算器的外观就以为和普通计算器无太大差异。又看了中文说明书,书中绝大部分内容是介绍求函数值、数值积分、微商、连乘及各种函数的画图;统计学部分书中只介绍了线性回归。说明书看完后觉得没多大意思就放在书桌边没理会它。2000 年元旦过后,一放寒假觉得不妨自己“研究研究”这台计算器。没想到仔细分析和运用之后才发现原来这是一台统计功能非常强大,很适宜全校文、理科学生学习基础统计学的辅助工具,该说明书根本没能反映出 TI-89 计算器统计功能的特点。于是我花了一个寒假的时间整理出一份如何用 TI-89 计算器进行教学的教案,并于 2000 年春在全校一百多人的公选课上开始应用,并立即受到文、理科学生的热烈欢迎,学习的积极性非常高。期末学生对本课程的评估分很高,第二年本课程被学校提升为全校通选课。TI 公司还对三名全班学得最好的文、理科学生赠送了三台机器。这一切使本课程成为了全校的一门优秀课程。本课程的成功除了有比较好的教学内容和经验之外还应该归功于 TI 的图形计算器。正是由于该计算器的强大功能和便利、直观的使用指令,使得哪怕从未接触过计算机程序的文科学生,学起来也觉得比较轻松、有趣。和使用本计算器以前的教学相比:内容大约增加了 $1/3$ 。比如描述性统计学中带野值判断的 Box-Whisker 图;小样本条件下正态性的 qq-norm 图;双总体方差不相等的均值检验;类别的相关性检验;回归分析中预报的区间估计问题、非线性回归中的周期检测等等都是一般理科(非数学)初等统计课讲不到的内容。然而我校文科学生(新闻、英语、历史等)却都能学得很好,应该说 TI 的图形计算器是起了很大的作用的。和以往的统计学课堂教学内容很不同的另一特点是:本课程将若干统计实验纳入到课堂教学中,如:大数定律的实验;用 Monte-Carlo 方法计算定积分;独立和的中心极

限定理; Buffon 投针的实验以近似估计 π 值; 等等. 我想一门面向广大文、理科学学生的普通统计学, 这些内容也应该算是精华的统计实验, 如果没有 TI 图形计算器, 这些实验恐怕是引不到课堂中来的.

我们鼓励学生用 TI-89 计算器算题有一前提, 那就是要求学生对统计方法的背景(典型问题)、概念(如检验问题的 H_0, H_1) 和理论结果(如大样本非正态的检验)等都应该是明白的, 虽然数学上没给出理论证明. 重点在于训练学生学习统计方法的正确使用和对结果的分析.

大家知道, 统计学即将正式纳入中学的教学轨道, 由以上对本书的介绍可以看出: 本书不仅可作为大专院校非数学系大学生的教材, 而且也可作为广大中学教师的教学参考书和中学生的课外参考书.

本书的许多内容用其他统计软件, 如 Excel, 也可以实现其计算和画图功能, 但 TI 公司的计算器, 不管是 TI-89 型、TI-92 Plus 型或新的 V-200 运用于教学对非理科学生更直观和更容易掌握.

应该说, 作者本人对该类产品的认识和开发利用还很不充分, 相信读者在学完本书后一定还会有很多新的应用, 尤其是统计实验. 希望读者能把这些新结果寄给本人, 以便将来在本书再版时作进一步的补充.

最后作者借此书出版的机会向美国德州 TI 公司表示真诚的谢意, 感谢他们多年来对本课程的大力支持和资助, 没有他们的有效配合和支持, 本课程是不会如此受到广大学生们的欢迎的.

谢衷洁

于北京大学承泽园

2004 年 8 月 15 日

目 录

第〇章 开篇引言	(1)
◆ 关于运用 TI-89、TI-92 Plus 及 V-200 图形计算器进行统计分析的 几点说明	(8)
第一章 绪论	(11)
1.1 引言	(11)
1.2 关于随机现象及其统计规律性	(12)
1.3 若干统计学中的名词	(16)
◆ 用 TI 图形计算器做掷硬币的实验	(21)
第二章 描述性统计学	(24)
2.1 引言	(24)
2.2 随机抽样及其频数	(24)
2.2.1 简单的随机抽样	(24)
2.2.2 数据的简单整理	(26)
2.3 直方图、频数密度分布图及频数多边形图	(30)
2.4 茎叶图	(35)
2.5 数字特征	(40)
2.5.1 刻画总体代表性的数值	(40)
2.5.2 刻画分散度的数值	(44)
2.5.3 均值、中位数与方差的简单性质	(47)
2.6 分位点和 Box-Whisker 图	(49)
2.6.1 Q_1 与 Q_3 (1/4 与 3/4 分位点)	(49)
2.6.2 Box-Whisker 图	(52)
2.7 本章的回顾与复习要点	(54)
◆ 用 TI 图形计算器画观测数据的直方图	(55)
◆ 图 2.14 对应数据的茎叶图	(57)
◆ 用 TI 图形计算器计算基本的统计量	(59)
◆ 用 TI 图形计算器画 Box Plot	(59)

5.2.2 正态单总体、方差已知时对均值的检验	(144)
5.2.3 大样本条件下,分布未知时对均值的检验	(145)
5.3 均值和概率的区间估计	(147)
5.4 正态双总体的均值检验	(151)
5.4.1 正态双总体、方差已知时对均值的检验	(151)
5.4.2 正态双总体、方差未知而相等时对均值的检验	(155)
5.4.3 正态双总体、方差不等且未知时对均值的检验	(158)
5.5 正态相关两个总体的均值检验	(160)
5.6 关于两类错误 α 和 β	(164)
5.7 附录: 关于 t 分布的概率密度函数及其期望与方差	(165)
5.8 本章的回顾与复习要点	(166)
◆ 用 TI 图形计算器对观测数据正态性的粗略观察	(167)
◆ 用 TI 图形计算器对正态总体的均值检验	(169)
◆ 用 TI 图形计算器对正态总体的均值检验(不等式检验)	(171)
◆ 用 TI 图形计算器做关于正态总体、方差已知条件下对均值的 Z 检验	(172)
◆ 用 TI 图形计算器做关于正态总体、方差未知条件下对均值的 区间估计	(174)
◆ 用 TI 图形计算器对概率的区间估计	(175)
◆ 用 TI 图形计算器做关于正态双总体、方差已知条件下对均值的检验 ..	(176)
◆ 用 TI 图形计算器做关于正态双总体、方差已知条件下对均值的 区间估计	(177)
◆ 用 TI 图形计算器做关于正态双总体、方差未知但相等条件下 对均值的检验	(178)
◆ 用 TI 图形计算器做关于正态双总体、方差未知且不相等条件下 对均值的检验	(179)
第六章 假设检验(II)	(180)
6.1 引言	(180)
6.2 方差的检验和区间估计	(180)
6.2.1 正态单总体的方差检验	(180)
6.2.2 独立双总体的方差检验	(184)
6.3 类别的相关性检验——列联表分析	(185)
6.4 分布的 χ^2 拟合优度检验	(192)
6.5 双总体的秩和检验	(196)