



全国统计教材编审委员会“十二五”规划教材

# 非参数统计

第四版

吴喜之 赵博娟 编著

 中国统计出版社  
China Statistics Press



全国统计教材编审委员会“十二五”规划教材

# 非参数统计

第四版

吴喜之 赵博娟 编著

 中国统计出版社  
China Statistics Press

## 图书在版编目(CIP)数据

非参数统计 / 吴喜之, 赵博娟编著. —4 版. —北京: 中国统计出版社, 2013. 10

全国统计教材编审委员会“十二五”规划教材

ISBN 978-7-5037-6977-1

I. ①非… II. ①吴… ②赵… III. ①非参数统计—高等学校—教材 IV. ①0212.7

中国版本图书馆 CIP 数据核字(2013)第 226793 号

## 非参数统计(第四版)

作 者/吴喜之 赵博娟

责任编辑/杨映霜 姜 洋

封面设计/上智博文

出版发行/中国统计出版社

通信地址/北京市丰台区西三环南路甲 6 号 邮政编码/100073

电 话/邮购(010)63376909 书店(010)68783171

网 址/<http://csp.stats.gov.cn/>

印 刷/河北天普润印刷厂

经 销/新华书店

开 本/710×1000mm 1/16

字 数/230 千字

印 张/14

印 数/1—3000 册

版 别/2013 年 10 月第 4 版

版 次/2013 年 10 月第 1 次印刷

定 价/28.00 元

版权所有。未经许可, 本书的任何部分不得以任何方式在世界任何地区以任何文字翻印、拷贝、仿制或转载。

中国统计版图书, 如有印装错误, 本社发行部负责调换。

# 国家统计局

## 全国统计教材编审委员会

---

顾问 罗 兰 袁 卫 冯士雍 吴喜之  
方积乾 王吉利 庞 皓 李子奈

主任 徐一帆

副主任 严建辉 田鲁生 邱 东 施建军  
耿 直 徐勇勇

委员(按姓氏笔划排序)

|     |     |     |     |     |
|-----|-----|-----|-----|-----|
| 丁立宏 | 万崇华 | 马 骏 | 毛有丰 | 王兆军 |
| 王佐仁 | 王振龙 | 王惠文 | 丘京南 | 史代敏 |
| 龙 玲 | 刘建平 | 刘俊昌 | 向书坚 | 孙秋碧 |
| 朱 胜 | 朱仲义 | 许 鹏 | 余华银 | 张小斐 |
| 张仲梁 | 张忠占 | 李 康 | 李兴绪 | 李宝瑜 |
| 李金昌 | 李朝鲜 | 杨 虎 | 杨汭华 | 杨映霜 |
| 汪荣明 | 肖红叶 | 苏为华 | 陈 峰 | 陈相成 |
| 房祥忠 | 林金官 | 罗良清 | 郑 明 | 柯惠新 |
| 柳 青 | 胡太忠 | 贺 佳 | 赵彦云 | 赵耐青 |
| 凌 亢 | 唐年胜 | 徐天和 | 徐国祥 | 郭建华 |
| 崔恒建 | 傅德印 | 景学安 | 曾五一 | 程维虎 |
| 蒋 萍 | 潘 璠 | 颜 虹 |     |     |

## 出版说明

全国统计教材编审委员会是国家统计局领导下的、全国统计教材建设工作的最高指导机构和咨询机构,自1988年成立以来,分别组织编写和出版了“七五”至“十一五”全国统计规划教材。

“十二五”时期,是我国全面实施素质教育,全面提高高等教育质量,深化教育体制改革,推动教育事业科学发展,提高教育现代化水平的时期。“十二五”伊始,统计学迎来了历史性的重大变革和飞跃。2011年2月,在国务院学位委员会第28次会议通过的新的《学位授予和人才培养学科目录(2011)》(以下简称“学科目录”)中,统计学从数学和经济学中独立出来,成为一级学科。这一变革和飞跃将对中国统计教育事业产生巨大而深远的影响,中国统计教育事业将在“十二五”时期发生积极变化。

正是在这一背景下,全国统计教材编审委员会制定了《“十二五”全国统计教材建设规划》(以下简称“规划”)。根据“学科目录”在统计学下设有数理统计学,社会经济统计学,生物卫生统计学,金融统计、风险管理与精算学,应用统计5个二级学科的构架,“规划”对“十二五”全国统计规划教材建设作了全面部署,具有以下特点:

第一,打破以往统计规划教材出版学科单一的格局。全面发展数理统计学,社会经济统计学,生物卫生统计学,金融统计、风险管理与精算学,应用统计5个二级学科规划教材的出版,使“十二五”全国统计规划教材涵盖5个二级学科,形成学科全面并平衡发展的出版局面。

第二,打破以往统计规划教材出版层次单一的格局。在编写出版好各学科本科生教材的基础上,对研究生教材出版进行深入研究,出版一批高水平高层次的研究生教材,为我国研究生教育、尤其是应用统计研究生教育提供教学服务。同时,积极重视统计专科教材出版,联合各专科院校,组织编写和出版适应统计专科教学和学习的优秀教材。

第三,打破以往统计规划教材出版品种单一的格局。鼓励内容创新,联系统计实践,具有教学内容和教学方法特色的、各高校自编的相同内容选题的精品教材出版,促进统计教学向创新性、创造性和多样性发展。

第四,重视非统计专业的统计教材出版。探讨对非统计专业学生的统计教学问题,为非统计专业学生组织编写和出版概念准确、叙述简练、深入浅出、表达方式活泼、练习题贴近社会生活的统计教材,使统计思想和统计理念深入非统计专业学生,以达到统计教学的最大效果。

第五,重视配合教师教学使用的电子课件和辅助学生学习使用的电子产品的配套出版,促进高校统计教学电子化建设,以期最后能形成系统,提高统计教育现代化水平。

第六,重视对已经出版的统计规划教材的培育和提高,本着去粗存精、去旧加新、与时俱进的原则,继续优化已经出版的统计教材的内容和写作,强化配套课件和习题解答,使它们成为精品,最后锤炼成为经典。

“十二五”期间,编审委员会将本着“重质量,求创新,出精品,育经典”的宗旨,组织我国统计教育界专家学者,编写和编辑出版好本轮教材。本轮教材出版后,将能够形成学科齐全、层次分明、品种多样、配套系统的高质量立体式结构,使我国统计规划教材建设再上新台阶,这将对推动我国统计教育和统计教材改革,推动我国统计教育事业科学发展,提高我国统计教育现代化水平产生积极意义。

让教师的教学和学生的学习事半功倍,并使学生在毕业之后能够学以致用统计教材,是本轮教材的追求。编审委员会将努力使本轮教材好教、好学、好用,尽力使它们在内容上和形式上都向国外先进统计教材看齐。限于水平和经验,在教材的编写和编辑出版过程中仍会有不足,恳请广大师生和社会读者提出批评和建议,我们将虚心接受,并诚挚感谢!

国家统计局  
全国统计教材编审委员会  
2012年7月

## 第四版前言

此书自出版至今得到了广大师生的大力支持,在许多学校被选为“非参数统计”课程的教科书或参考书。通过使用,不少师生对本书提出了宝贵的意见和建议。针对书中错误和不妥之处,本书第二版对许多内容进行了调整和重写,还对例题和习题都做了一些修订和增减。第三版增加了作者,完善了书中的 R 程序和书后附录表,纠正了在第二版中发现的错误和不妥之处。第三版还去掉了第二版中与非参数统计关系不大的第二章,以减少教学负担。

为了更加系统地介绍非参数统计方法,本书第四版对第三版中的内容进行了补充。在第四版的第三章添加了 McNemar 和 Cohen's Kappa 检验,第四章添加了 Cochran 精确检验的 R 程序,第六章添加了 Kendall's  $\tau_b$ , Kendall's  $\tau_c$  (也称 Stuart's  $\tau_c$ ), Goodman-Kruskal's  $\gamma$ , Somers'  $d(C|R)$ , Somers'  $d(R|C)$  和 Somers'  $d$  等度量两个有序变量相关性或关联性 (association) 的方法。我们还更新了第四章的一些程序和有关精确检验附录表,在第八章介绍了胜算比 (Odds ratio), 相对风险 (Relative Risk) 和 Cochran-Mantel-Haenszel 估计等在生物统计中比较常用的列联表分析方法。

延续前几版的做法,第四版增添的各节内容都从分析相关例子入手,根据具体的数据结构,引进将要介绍的检验方法,给出分析结论。在每节末都有软件使用注解,给出了如何分别用 R、SPSS 和 SAS 等软件对例子数据进行分析的具体步骤。每章末都有相关的练习题,以便读者练习使用有关检验方法。对于对第三版中存在的错误和不妥之处,也进行了修正。

第四版保留了对第三版某些章节所加的星号 (\*), 包括: § 1.6、§ 1.7、§ 2.3、§ 3.3、§ 4.2、第五章、第九章及第十章。教师可以根据实际教学需要,介绍性地或选择地讲,也可以完全不讲。

最后,借此书再版之际,我们再次对使用本教材的广大师生表示感谢。感谢他们提出的宝贵意见和建议,他们的建议和要求是推动本书再版的主要动力。同时,我们非常感谢中国统计出版社对于本书再版的大力支持。

希望读者继续对本书予以宝贵的支持和批评指正。

吴喜之 赵博娟

2013年8月

## 第三版前言

根据作者和许多非参数统计课教师的实践,我们觉得有必要出本书的第三版。第三版首先纠正了在第二版中发现的错误和不妥之处,并且对内容作了部分的修订。这一版还去掉了第二版中与非参数统计关系不大的第二章,以减少教学的负担。我们保留了第二版中广受欢迎的使用 R、SPSS 或 SAS 等统计软件来分析有关数据的程序语句和各种选项的说明。

这里仍然强调,对于初学的或实际应用部门的读者,可以略去打星号(\*)的章节,这些章节至少包括:§ 1.6、§ 1.7、§ 2.3、§ 3.3、§ 4.2、第五章、第九章及第十章。第一章主要是用于介绍、回顾或参考的,可以由教师有选择地根据情况选择地讲,也可以完全不讲。实际上,对于任务课程,应该完全由任课教师来决定讲哪些内容以及如何讲,教学大纲都应该服务于实际教学的需要,而不应成对教师的束缚。教科书应该留给教师以较大的余地和自由。

这里必须对所有使用本教材的广大师生表示感谢,他们提出了不少宝贵的建议和意见,是推动并鼓励本书再版的主要动力,同时也要对中国统计出版社对于本书第三版的支持表示感谢。

希望各方面的读者继续对本书予以宝贵的支持,并提出批评和建议。

作者

2008 年 10 月

## 第二版说明

本书的第一版发行以来,在许多学校被选为“非参数统计”课程的教科书或参考书。各个学校的师生对本书提出许多宝贵的意见,并且指出了很多错误和不妥之处。没有他们的支持和鼓励,本书的第二版不可能面世。

和第一版相比较,第二版对许多内容完全重新写过,还进行了一些调整,同时加强了对概念和方法的解释,使得该书更加容易理解。第二版还对例题和习题都做了一些修订和增减,并且都在光盘给出了数据。此外,还增加了一些内容,特别是关于如何通过编程来理解方法,以及用软件来实现数值计算的内容。本书在课文中关于计算方法的叙述中主要使用了免费的,功能强大,需要自己动手写程序的 R 软件;力图清楚地用 R 语句来描述计算的细节。这也是一些“黑匣子”式的傻瓜软件所无法比拟的。R 软件是使用 S 语言来编程的(和 S-plus 的编程语言一样);在其问世的不到十年的时间,已经成为国外统计研究生的首选软件。它有强大的网上支持系统。多数最新的统计计算方法,在进入商业软件之前,就已经以 R 语言的形式在 R 网站上免费提供。使用本书的师生最好也使用 R 语言。掌握 R 软件对其他统计方向的学习和研究都会有很大的帮助,甚至会有一种到了自由天地的感觉。

为了适应分析实际数据的各种需要,本书还在每一节(除了少数介绍性章节之外)的最后加入了使用 R、SPSS 或 SAS 等统计软件分析有关数据的程序语句和各种选项的说明。这里要指出的是,编者尽量使本书提供的 R 程序是直接根据公式或定义写成的;这里的 R 程序没有按照专业化编写软件所通常遵循的高效率和漂亮输出的原则;这是因为那将使得显示基本公式和概念的语句淹没在为了形式和效率而加入的大量其他语句之中,而使得有关程序难以读懂。希望本书在编程上起着一个抛砖引玉的作用,鼓励读者编出更加高效、更加漂亮的程序。

本书选择的与内容有关的 SPSS 软件选项和 SAS 软件语句(或选项)的原则是容易理解和掌握;当然,由于编者知识有限,对于有些方法,没有找到(因此也无法提供)合适及方便的 SPSS 或 SAS 方法;希望读者提出建议,使得再版时予以弥补。

由于使用软件比查表更加方便和可靠,有人说,你自己都不查表,为什么要教学生去查表呢?的确,编者除了在最初等的统计课教学过程中曾经涉及到少数统计表之外,从来都是使用软件。“己所不欲,勿施于人”,本书原本不想再提供任何统计分布表,但为了部分没有计算条件的读者易于理解,还是提供了少数最常用的表格,以备不时之需。希望有条件的读者尽量使用计算机,而不去查表。实际上,如果没有计算机的支持,很难对有一定规模的数据在任何统计方向进行较深入的分析。

一些读者提出,本书内容对于每周两学时的课程似乎太多。我觉得,对于初学者或者实际应用部门的人来说,可以略去不讲的章节(打\*号的)至少包括:§ 1.6、§ 1.7(正态记分部分)、§ 3.3、§ 4.3、§ 5.2、第六章、第十章及第十一章<sup>①</sup>。总的来说,第一章主要是用于介绍、回顾或参考的,可以有选择地在需要时讲、也可以完全不讲,这应该根据学生的需要由老师自己安排。实际上,对于任何课程,应该由任课教师来决定讲哪些内容以及如何讲,因为他们最了解他们所面对的学生。教科书编者的思维方式不见得和老师的一致,而老师最好按照自己的理解来讲述。一个好的教科书,应该给教师以较大的余地和自由。

希望读者继续对本书予以宝贵的支持和批评指正。

吴喜之

2006年4月

---

① 注意:第二版和第三版章节不尽相同。

# 第一版前言

本书的目的是用简明的语言,不多的数学工具并通过大量例子来尽可能直观地介绍非参数统计的基本方法。它可以作为统计学专业本科一学期(2学时)的应用非参数统计课程的教材,也可以作为实际工作者自学或查阅的参考书。所需要的预备知识为统计学教程中的最基本的内容。读者只要知道总体和样本,随机变量及分布,统计量,检验和估计的基本概念等即可以看懂本书。虽然计算机并不是学会本书内容所必需的,但是不能想象,一个不会用计算机的统计工作者如何在实践中生存。

本书在引进每一个方法时,都通过数据例子来说明该方法的意义和使用过程。所有例题的计算和绘图都是由笔者完成的。笔者还核算了每一章后面的所有习题。由于这些习题都只涉及基本概念和方法。相信读者完全可以独立完成。由于本书的基本原理和方法广泛适用于许多不同的领域,这里的例子和习题尽量取自不同的领域和学科,以扩展读者的思路。

本书的第一章引言部分包含以下几类内容:(1)对统计和非参数统计以及计算机软件应用的一般论述;(2)对一些初等统计内容,特别是对本书常遇见的问题作了回顾;这些问题有一般的检验与置信区间问题,特别的 $\chi^2$ 检验问题,探索性数据分析问题等;(3)初等统计不一定有的问题,比如渐进相对效率(ARE)和局部最优势(LMP)检验,顺序统计量,秩,线性秩统计量和线性记分问题。这里的(1)和(2)可以根据使用者的情况酌情处理,最好先浏览一下,而在需要时再读。第(3)部分内容在书中多次涉及;但由于仅与理论推导和对方法的评价有关,可作为有兴趣的人的参考。

从第二章到第七章依次序为关于位置的单样本,两样本和多样本

模型,尺度问题,相关与回归问题以及分布及一些  $\chi^2$  检验问题。这些一般都可以讲;但是如果时间安排不开,可以对正态记分部分,仅作举一反三的式的介绍。这并不是它不常用,而是因为其思想仅仅是别的统计量的推广。最后两章为非参数密度估计和回归与稳健统计简介。这两部分中每一部分都可以构成数倍于本书厚度的专著。它们在统计中占有重的地位,这里的内容仅打算让读者作一初步了解。

本书在编写过程中始终得到国家统计局教育中心的关心和帮助。苏州大学的汪仁官教授极其认真地审阅了全书,并提出了宝贵的意见;自然,所有的意见都是非常合理的而且均被采纳了;这使我回忆起 36 年前敬爱的汪老师为我们仔细批改数学分析作业的感人情景。本书的大部分内容和例子曾在人民大学讲过,在此也对积极参与课堂教学的统计 96 级同学一并表示感谢。

编者水平有限;欢迎各方面能对本书的错误和不当之处予以批评指正。

吴喜之

1999 年 11 月 20 日

## 目 录

|                                             |    |
|---------------------------------------------|----|
| 第一章 引言 .....                                | 1  |
| 1.1 统计的实践 .....                             | 1  |
| 1.2 关于非参数统计 .....                           | 2  |
| 1.3 假设检验及置信区间的回顾 .....                      | 3  |
| 1.4 $\chi^2$ 检验简单回顾 .....                   | 7  |
| 1.4.1 基于随机化模型的 $\chi^2$ 检验 .....            | 8  |
| 1.4.2 关于离散分布的列联表 $\chi^2$ 检验 .....          | 9  |
| 1.5 熟悉手中的数据和数据变换 .....                      | 9  |
| 1.6 渐近相对效率(ARE)、局部最优势(LMP)检验* .....         | 11 |
| 1.7 顺序统计量,秩,线性秩统计量及正态记分* .....              | 13 |
| 1.8 计算机统计软件的应用 .....                        | 16 |
| 1.9 习题 .....                                | 18 |
| 第二章 单样本问题 .....                             | 21 |
| 2.1 广义符号检验和有关的置信区间 .....                    | 21 |
| 2.1.1 广义符号检验:对分位点进行的检验 .....                | 23 |
| 2.1.2 基于符号检验的中位数及分位点的置信区间 .....             | 25 |
| 2.2 Wilcoxon 符号秩检验,点估计和区间估计 .....           | 28 |
| 2.2.1 Wilcoxon 符号秩检验 .....                  | 28 |
| 2.2.2 基于 Wilcoxon 符号秩检验的点估计和置信区间 .....      | 33 |
| 2.3 正态记分检验* .....                           | 35 |
| 2.4 Cox-Stuart 趋势检验 .....                   | 38 |
| 2.5 关于随机性的游程检验 .....                        | 40 |
| 2.6 习题 .....                                | 43 |
| 第三章 两样本数据 .....                             | 47 |
| 3.1 两样本和多样本的 Brown-Mood 中位数检验 .....         | 48 |
| 3.2 Wilcoxon(Mann-Whitney)秩和检验及有关置信区间 ..... | 52 |
| 3.2.1 Wilcoxon(Mann-Whitney)秩和检验 .....      | 52 |

|                                           |            |
|-------------------------------------------|------------|
| 3.2.2 $M_X - M_Y$ 的点估计和区间估计 .....         | 56         |
| 3.3 正态记分检验* .....                         | 57         |
| 3.4 成对数据的检验 .....                         | 59         |
| 3.5 McNemar 检验 .....                      | 61         |
| 3.6 Cohen's Kappa 系数 .....                | 63         |
| 3.7 习题 .....                              | 65         |
| <b>第四章 多样本数据 .....</b>                    | <b>69</b>  |
| 4.1 Kruskal-Wallis 秩和检验 .....             | 69         |
| 4.2 正态记分检验* .....                         | 73         |
| 4.3 Jonckheere-Terpstra 检验 .....          | 75         |
| 4.4 区组设计数据分析回顾 .....                      | 77         |
| 4.5 完全区组设计;Friedman 秩和检验 .....            | 79         |
| 4.6 Kendall 协同系数检验 .....                  | 83         |
| 4.7 完全区组设计:关于二元响应的 Cochran 检验 .....       | 85         |
| 4.8 完全区组设计:Page 检验 .....                  | 88         |
| 4.9 不完全区组设计;Durbin 检验 .....               | 91         |
| 4.10 习题 .....                             | 93         |
| <b>第五章 尺度检验* .....</b>                    | <b>97</b>  |
| 5.1 两独立样本的 Siegel-Tukey 方差检验 .....        | 97         |
| 5.2 两样本尺度参数的 Mood 检验 .....                | 99         |
| 5.3 两样本及多样本尺度参数的 Ansari-Bradley 检验 .....  | 102        |
| 5.4 两样本及多样本尺度参数的 Fligner-Killeen 检验 ..... | 105        |
| 5.5 两样本尺度的平方秩检验 .....                     | 107        |
| 5.6 多样本尺度的平方秩检验 .....                     | 109        |
| 5.7 习题 .....                              | 110        |
| <b>第六章 相关和回归 .....</b>                    | <b>112</b> |
| 6.1 Spearman 秩相关检验 .....                  | 113        |
| 6.2 Kendall $\tau$ 相关检验 .....             | 116        |
| 6.3 Goodman-Kruskal's $\gamma$ 相关检验 ..... | 121        |
| 6.4 Somers'd 相关检验 .....                   | 123        |
| 6.5 Theil 非参数回归和几种稳健回归 .....              | 125        |

|                                                   |            |
|---------------------------------------------------|------------|
| 6.6 习题 .....                                      | 131        |
| <b>第七章 分布检验和拟合优度 <math>\chi^2</math> 检验 .....</b> | <b>134</b> |
| 7.1 Kolmogorov-Smirnov 单样本检验及一些正态性检验 .....        | 135        |
| 7.1.1 Kolmogorov-Smirnov 单样本分布检验 .....            | 135        |
| 7.1.2 关于正态分布的一些其他检验和相应的 R 程序 .....                | 138        |
| 7.2 Kolmogorov-Smirnov 两样本分布检验 .....              | 141        |
| 7.3 Pearson $\chi^2$ 拟合优度检验 .....                 | 142        |
| 7.4 习题 .....                                      | 144        |
| <b>第八章 列联表 .....</b>                              | <b>146</b> |
| 8.1 二维列联表的齐性和独立性的 $\chi^2$ 检验 .....               | 146        |
| 8.2 低维列联表的 Fisher 精确检验 .....                      | 149        |
| 8.3 两个比例的比较 .....                                 | 152        |
| 8.4 Cochran-Mantel-Haenszel 估计 .....              | 155        |
| 8.5 对数线性模型与高维列联表的独立性检验简介 .....                    | 157        |
| 8.5.1 处理三维表的对数线性模型 .....                          | 158        |
| 8.5.2 假设检验和模型的选择 .....                            | 159        |
| 8.6 习题 .....                                      | 161        |
| <b>第九章 非参数密度估计和非参数回归简介* .....</b>                 | <b>164</b> |
| 9.1 非参数密度估计 .....                                 | 164        |
| 9.1.1 一元密度估计 .....                                | 165        |
| 9.1.2 多元密度估计 .....                                | 168        |
| 9.2 非参数回归 .....                                   | 169        |
| <b>第十章 稳健统计方法简介* .....</b>                        | <b>176</b> |
| <b>附表 .....</b>                                   | <b>180</b> |
| <b>参考文献 .....</b>                                 | <b>197</b> |

# 第一章 引言

## 1.1 统计的实践

虽然“统计学(statistics)”的定义在当今世界的百科全书和统计教科书中于文字、侧重点或描述方式有所出入,但就其所包含的总体内容和应用领域来说则差不多. 比如“不列颠百科全书(Encyclopædia Britannica)”(2008)把统计定义为:“收集、分析、展示和解释数据的科学”<sup>1</sup>. 以Kotz和Johnson(1983)主编的13卷“统计科学百科全书(Encyclopedia of Statistical Sciences)”是迄今最完整的关于统计的具有权威性的百科全书. 它说“统计学”这个术语表示“涉及收集、展示和分析数据的普遍方法和原理的领域”. 它还列举了四十多个运用统计的领域,它们包括:精算,农业,动物学,人类学,考古学,审计学,晶体学,人口统计学,牙医学,生态学,经济计量学,教育学,选举预测和策划,工程,流行病学,金融,水产渔业研究,遗传学,地理学,地质学,历史研究,人类遗传学,水文学,工业,法律,语言学,文学,劳动力计划,管理科学,市场营销学,医学诊断,气象学,军事科学,核材料安全管理,眼科学,制药学,物理学,政治学,心理学,心理物理学,质量控制,宗教研究,社会学,调查抽样,分类学,博彩和气象改善.

统计在每一个应用领域都有它自己的目标和特点,有的还有自己的名字,比如生物统计和商务统计. 各个应用统计领域既有个性又有共性. 多数普遍应用的统计方法最初是为某一个应用领域而发展的,然后为其它领域所利用. 这些统计方法和原理逐渐形成统计学的基础.

统计作为一门科学,随着其应用的发展和深入,涉及大量的数据及复杂的模型,因而也需要先进的计算机和越来越多的数学. 事实表明,数学和计算机的大量运用加速了统计学的发展,也更新了统计学的面貌. 当前,统计是计算机数值计算的最重要的用户. 今天的统计学如果没有计算机是不可想象的.

统计应用的广泛性既造就了一批为各个具体应用领域服务的,并懂得该领域内容的统计学家,同时也造就了一些相对独立于某一两项具体应用,从事于研究具有普遍性的统计方法或原理的统计学家. 后者所研究的内容有时也被称为“数理统计”. 他们对目前广泛应用的大量的统计模型有着重要的贡献. 然而这些似乎“脱离”某一两个具体应用领域的表面现象以及他们所使用的复杂的数学工具,使得有些人认为统计(或数理统计)就是数学或数学的一个分支. 实际上,也的确有许多人把统计学当成数学来研究. 这些自然要引起一些争论. 这没有关系,在数

<sup>1</sup>原文为: The science of collecting, analyzing, presenting, and interpreting data.