

21
世纪统计学系列教材

教育部推荐教材

平群进
俊晓勇
贾何金
编著

统计学

(第二版)



中国人民大学出版社

教育部推荐教材
21 世纪统计学系列教材

统 计 学

(第二版)

贾俊平 何晓群 金勇进 编著

中国人民大学出版社

图书在版编目 (CIP) 数据

统计学/贾俊平等编著. 2 版.
北京: 中国人民大学出版社, 2004
(21 世纪统计学系列教材)

ISBN 7-300-03589-2/F·1082

- I. 统…
- II. 贾…
- 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 100920 号

教育部推荐教材

21 世纪统计学系列教材

统计学 (第二版)

贾俊平 何晓群 金勇进 编著

出版发行	中国人民大学出版社		
社 址	北京中关村大街 31 号	邮政编码	100080
电 话	010-62511242 (总编室)	010-62511239	(出版部)
	010-82501766 (邮购部)	010-62514148	(门市部)
	010-62515195 (发行公司)	010-62515275	(盗版举报)
网 址	http://www.crup.com.cn http://www.ttrnet.com (人大教研网)		
经 销	新华书店	版 次	2000 年 9 月第 1 版
印 刷	北京雅艺彩印有限公司		2004 年 10 月第 2 版
开 本	787×965 毫米 1/16	印 次	2006 年 1 月第 5 次印刷
印 张	30.75	定 价	38.00 元(含光盘)
字 数	562 000		

版权所有 侵权必究

印装差错 负责调换

《21 世纪统计学系列教材》编委会

编委会主任 易丹辉

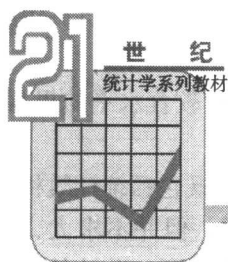
编委会委员 (按姓氏笔画排序)

尹德光 冯士雍 张尧庭

陈希孺 吴喜之 赵彦云

柯惠新 袁 卫 倪加勋

顾 岚 袁寿庄 耿 直



总 序

改革开放以来，高等统计教育有了很大的发展。随着课程设置的不断调整，有不少教材出版，同时也翻译引进了一些国外优秀教材。作为培养我国统计专门人才的摇篮，中国人民大学统计学系自 1952 年创建以来，走过了风风雨雨，一直坚持着理论与应用相结合的办学方向，培养能够理论联系实际、解决实际问题的高层次人才。随着新知识经济和网络时代的到来，我们在教学科研的实践中，深切地感受到，无论是自然科学领域、社会科学领域的研究，还是国家宏观管理和企业生产经营管理，甚至人们的日常生活，信息需求量日益增多，信息处理技术更加复杂，作为信息技术支柱的统计方法，越来越广泛地应用于各个领域。

面对新的形势，我们一直在思索，课程设置、教材选择、教学方式等怎样才能使学生适应社会经济发展的客观需要。在反复酝酿、不断尝试的基础上，我们决定与统计学界的同仁，共同编写、出版一套面向 21 世纪的统计学系列教材。

这套系列教材聘请了中科院院士、中国科技大学陈希孺教授，上海财经大学数量经济研究院张尧庭教授，中国科学院数学与系统科学研究所冯士雍研究员等作为编委。他们长期任中国人民大学的兼职教授，一直关心、支持着统计学系的学科建设和应用统计的发展。中国人民大学应用统计科学研究中心 2000 年已成为国家级研究基地，这些专家是首批专职或兼职研究人员。这一开放性研究基地

的运作，将有利于提升我国应用统计科学研究的水平，也必将进一步促进高等统计教育的发展。

这套教材是我们奉献给新世纪的，希望它能促进应用统计教育水平的提高。这套教材力求体现以下特点：

第一，在教材选择上，主要面向经济类统计学专业。选材既包括统计教材也包括风险管理与精算方面的教材。尽管名为统计学系列教材，但并不求大、求全，而是力求精选。对于目前已有的内容较为成熟、适合教学需要、公认的较好的教材，并未列入本次出版计划。

第二，每部教材的内容和写作，注意广泛吸收国内外优秀教材的成果。教材力求简明易懂、内容系统和实用，注重对统计方法思想的阐述，并结合大量实际数据和实例说明统计方法的特点及应用条件。

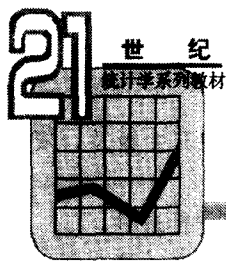
第三，强调与计算机的结合。为着力提高学生运用统计方法分析解决问题的能力，教材所涉及的统计计算，要求运用目前已有的统计软件。根据教材内容，选择使用 SAS、SPSS、TSP、STATISTICA、EViews、MINITAB、Excel 等。

感谢中国人民大学出版社的同志们，他们怀着发展我国应用统计科学的热情和提高统计教育水平的愿望，经过反复论证，使这套教材得以出版。感谢参与教材编写的同行专家、统计学系的教师。愿大家的辛勤劳动能够结出丰硕的果实。我们期待着与统计学界的同仁，共同创造应用统计辉煌的明天。

易丹辉

2000 年 8 月

于中国人民大学



第二版前言

《统计学》第一版出版后，得到社会各界的厚爱，有幸被一些兄弟院校作为教材使用，先后获得“教育部推荐教材”、“第四届国家统计局优秀统计教材”等殊荣。在3年多的时间里，先后9次印刷。这些都是对我们的极大鼓舞。经过几年的教学实践，我们发现书中仍有许多不足之处。特别是随着教学手段的进步，教学内容和教学重点也在发生着变化，这些都对教学内容提出了一些新的要求。正是基于这样的原因，才使得对教材的修订显得十分必要。

第二版是在第一版的基础上重新编写的。同第一版相比，第二版具有以下变化：

第一，在内容上增加了一些新的章节，并对一些章节的内容进行了扩充，如增加了“抽样分布”一章，将回归分析分为“一元线性回归”和“多元线性回归”两章等。同时，考虑到中学所学过的统计知识，将部分内容进行了大量缩减。

第二，将统计计算完全与计算机结合。根据教材内容的特点，我们仍然选用Excel作为计算工具，并且将Excel的应用与教材内容完全结合。大部分数据表格和计算结果也都以Excel形式给出。

第三，在每一章的后面都给出了本章的中英文对照主要术语，以便使读者复

习和掌握一些关键性术语。

第四，每章后面的习题量有所增加，类型更加多样化，并较多地采用了实际数据。

第五，对教学辅助光盘进行了全面更新，实现了与 Excel 的完全链接，使教学和学习都更加方便，并增加了项目教学的一些实际案例。

考虑到不同院校的教学特点，对部分内容可根据实际情况进行选讲。

第二版仍由贾俊平任主编，各章的作者也进行了调整。第1章、第3章、第4章、第10章、第11章、第12章、第13章、第14章由贾俊平编写；第5章由何晓群编写；第6章、第7章由贾俊平和何晓群编写；第2章、第8章、第9章由金勇进编写。

第二版的顺利完成，要特别感谢下列同志：

感谢编委会的专家吴喜之教授对第二版初稿的认真审阅，特别是指出了书中的多处原则性错误，作者对这些错误进行了纠正，受益匪浅。

感谢编委会主任易丹辉教授的大力支持以及对时间序列分析和预测一章提出的修改建议。

感谢我的同事杜子芳先生，在每次编写中，他总是给予特别多的建设性建议。

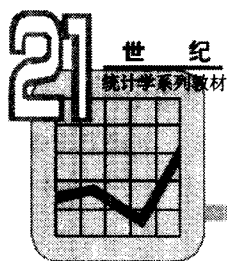
感谢曾讲授过该门课程的王燕老师、王星老师、张云老师、姚嘉秋老师、赵以立老师、陈龙老师、刘文卿老师、李静萍老师、黄向阳老师等，他们多年的教学经验为本书写作积累了宝贵财富。特别感谢王燕同志对第二版写作提出的宝贵意见以及张云同志对第一版中错误的纠正。

感谢中国人民大学出版社对本书出版的大力支持。感谢编辑同志对书稿的认真校对。

尽管我们做了各种努力，仍难以保证让读者和同行满意。您的建议始终是我们进一步修改的基础。

贾俊平

2004年6月于中国人民大学



第一版前言

在人类迈进 21 世纪、面对知识经济时代的今天，我们每天对信息的需求量和处理量日益增多，作为数据处理和分析技术的统计方法也越来越广泛地应用于自然科学和社会科学研究、生产和经营管理及日常生活中，这无疑对统计教材的编写提出了更高的要求。在过去的一段时期里，我国高等财经院校的统计教材一直沿用苏联的社会经济统计学体系。随着我国市场经济体系的建立，原有的统计学教材在内容上已不能满足社会经济发展的需要。尽管近年来高等财经院校的统计学教材在内容和体系上都有所改进，但仍存在一些不足。由于统计学是财经院校经济管理类各专业的一门基础必修课程，因此编好这一教材就成了一项很有意义的工作。

构建统计学教材的内容体系，关键是要对统计学科有一个全面科学的理解和认识。我们认为，统计学是一门收集、整理和分析统计数据的方法科学，其目的是探索数据的内在数量规律性。取得统计数据是进行统计分析的基础和前提，离开了统计数据，统计方法也就失去了用武之地。如何取得较为准确的统计数据是统计学研究的内容之一。统计数据的整理是数据收集与数据分析之间的一个必要环节，是通过对统计数据的加工处理使其系统化、条理化，符合统计分析的需要。统计数据的分析是统计学的核心内容，是通过统计描述和统计推断探索数据

内在规律的过程。基于这一认识,鉴于财经院校经济管理各专业统计教学的实际需要,本教材的内容体系包括描述统计(统计数据的搜集、数据的整理与显示、数据分布特征的描述)、推断统计(概率与概率分布、抽样与参数估计、假设检验、方差分析、列联分析等)、经济管理常用的统计方法(相关与回归分析、时间序列分析、指数)等几部分。本书既可作为高等院校财经管理类各专业本科生教材,也可作为广大实际工作者的参考书。

本教材力求体现以下特点:

第一,在内容上,立足于“大统计”的角度,从统计数据出发,以统计数据的处理和分析为核心,并根据统计教学的实际需要构建本教材的内容体系。在写法上,力求简明易懂,注重于对统计方法思想的阐述,结合大量的实际数据和实例说明统计方法的特点、应用条件和适用场合。

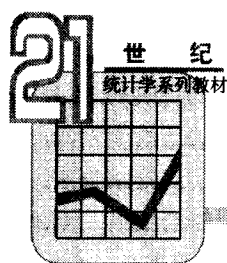
第二,强调计算机的应用。在大部分章的最后都给出了用 Excel 进行计算和分析的步骤,可以使學生利用相关章节学习的方法,通过计算机实现其计算和分析任务,这不仅可减轻学生的计算负担,也可以提高其运用统计方法分析和解决问题的能力。我们之所以选用 Excel 主要出于以下几点考虑:一是 Excel 较为普及,为大多数人所熟悉。尽管目前的统计软件如 SAS、SPSS、STATISTICA、MINITAB、TSP 等有更强的统计功能,但不如 Excel 普及。二是就本书所介绍的统计方法来看,Excel 中的统计功能已能够满足计算和分析的需要。三是作为 Office 的组成部分之一,Excel 能够很好地与 Office 中的其他软件结合,比如与 PowerPoint 结合制作教学幻灯片等。

作为 21 世纪统计学系列教材之一,本教材的编写得到了中国人民大学统计学系全体教师的支持和帮助。编写大纲拟定后,教材编委会及全体教师进行了广泛讨论,并进行了多次修改。袁寿庄教授、倪加勋教授、尹德光教授审阅了本书的初稿,并提出了许多宝贵意见,在此一并表示感谢。

本教材各章执笔人是:贾俊平(第一、二、三、四、十一、十二章及各章后的附录)、何晓群(第五、六、十章)、金勇进(第七、八、九章)。全书由贾俊平任主编,负责全书大纲的设计、书稿的组织和编纂。尽管几年来我们对统计学教材的编写一直进行着各种努力和尝试,很想奉献给读者一本满意的教材,但仍有可能达不到各方面的要求。书中的不当或疏漏之处在所难免,恳请同行和读者提出宝贵意见。

贾俊平

2000 年 7 月于中国人民大学



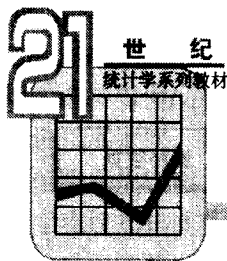
目 录

第1章 导论	1
1.1 什么是统计学	1
1.2 统计数据类型	5
1.3 统计中的几个基本概念	7
本章主要术语	11
思考与练习	12
第2章 数据的搜集	13
2.1 数据的来源	13
2.2 调查数据	16
2.3 实验数据	27
2.4 数据的误差	33
2.5 数据文件	39
本章主要术语	41
思考与练习	41
第3章 数据的整理与显示	43
3.1 数据的预处理	43

3.2 分类和顺序数据的整理与显示·····	48
3.3 数值型数据的整理与显示·····	55
3.4 统计表·····	70
本章主要术语·····	72
思考与练习·····	72
第4章 数据分布特征的测度 ·····	77
4.1 集中趋势的测度·····	77
4.2 离散程度的测度·····	89
4.3 偏态与峰态的测度·····	98
本章主要术语·····	102
思考与练习·····	103
第5章 概率与概率分布 ·····	107
5.1 随机事件及其概率·····	107
5.2 概率的性质与运算法则·····	111
5.3 离散型随机变量及其分布·····	118
5.4 连续型随机变量的概率分布·····	131
本章主要术语·····	140
思考与练习·····	140
第6章 抽样分布 ·····	143
6.1 三种不同性质的分布·····	143
6.2 一个总体参数推断时样本统计量的抽样分布·····	145
6.3 两个总体参数推断时样本统计量的抽样分布·····	153
本章主要术语·····	156
思考与练习·····	156
第7章 参数估计 ·····	159
7.1 参数估计的一般问题·····	159
7.2 一个总体参数的区间估计·····	165
7.3 两个总体参数的区间估计·····	172
7.4 样本容量的确定·····	182
本章主要术语·····	185
思考与练习·····	186
第8章 假设检验 ·····	190
8.1 假设检验的基本问题·····	190

8.2 一个正态总体参数的检验	199
8.3 两个正态总体参数的检验	209
8.4 假设检验中的其他问题	221
本章主要术语	225
思考与练习	226
第9章 列联分析	229
9.1 分类数据与列联表	229
9.2 拟合优度检验	233
9.3 独立性检验	239
9.4 列联表中的相关测量	241
9.5 列联分析中应注意的问题	245
本章主要术语	248
思考与练习	248
第10章 方差分析与试验设计	251
10.1 方差分析引论	251
10.2 单因素方差分析	257
10.3 方差分析中的多重比较	267
10.4 双因素方差分析	269
10.5 试验设计初步	280
本章主要术语	285
思考与练习	286
第11章 一元线性回归	289
11.1 变量间关系的度量	289
11.2 一元线性回归	297
11.3 利用回归方程进行估计和预测	312
11.4 残差分析	316
本章主要术语	322
思考与练习	323
第12章 多元线性回归	328
12.1 多元线性回归模型	328
12.2 回归方程的拟合优度	332
12.3 显著性检验	334
12.4 多重共线性	338

12.5 利用回归方程进行估计和预测·····	341
12.6 虚拟自变量的回归·····	343
本章主要术语·····	353
思考与练习·····	353
第 13 章 时间序列分析和预测 ·····	358
13.1 时间序列及其分解·····	358
13.2 时间序列的描述性分析·····	360
13.3 平稳序列的平滑和预测·····	365
13.4 有趋势序列的分析和预测·····	371
13.5 复合型序列的分解·····	384
本章主要术语·····	390
思考与练习·····	391
第 14 章 指数 ·····	396
14.1 引言·····	396
14.2 加权指数·····	397
14.3 指数体系·····	404
14.4 几种常用的价格指数·····	407
14.5 多指标综合评价指数·····	411
本章主要术语·····	416
思考与练习·····	417
附录一 各章习题答案 ·····	419
附录二 常用统计表 ·····	436
表 1 二项分布表 ·····	436
表 2 泊松分布表 ·····	448
表 3 标准正态分布表 ·····	450
表 4 正态分布分位数表 ·····	453
表 5 t 分布表 ·····	456
表 6 χ^2 分布表 ·····	458
表 7 F 分布表 ·····	461
表 8 随机数表 ·····	473
参考文献 ·····	476



第 1 章

导 论

在日常生活中，我们经常会接触到“统计”这一术语，在有关媒体中也经常会看见一些报导使用统计数据、图表等。难免很多人可能会将“统计”一词与统计工作联系起来：一提到统计首先想到的是统计工作，想到我们的政府统计机构、企业中从事统计工作的人，等等。这种理解是自然的，但又不全面的。在本章中，我们将讨论有关统计学的一些基本问题，包括统计的含义、统计数据及其分类、统计中一些常用的基本概念等。

1.1 什么是统计学

一、统计学与统计规律

统计学是随着人类社会的发展和社会管理的需要而发展起来的。目前，随着统计方法在各个领域的应用，统计学已发展成为具有多个分支学科的大家族。统计学家们给统计学下的定义繁多。比较有代表性的是不列颠百科全书的定义：“统计学是收集、分析、表述和解释数据的科学。”这一定义揭示了统计学是一套处理数据的方法和技术。

统计学的定义告诉我们，统计离不开数据。统计研究的过程首先要有数据，在拿到数据后，为满足分析的需要，还要对数据进行一定的整理，而后再对数据进行分析和解释。如图 1—1 所示，可将统计研究的过程描述为：



图 1—1 统计研究的过程

数据收集也就是取得统计数据。数据整理是将数据用图表等形式展示出来。数据分析则是通过统计方法研究数据，其所用的方法可分为描述统计方法和推断统计方法。描述统计（descriptive statistics）研究如何取得反映客观现象的数据，并通过图表形式对所收集的数据进行加工处理和显示，进而通过综合、概括与分析得出反映客观现象的规律性特征。内容包括统计数据的收集方法、数据的加工处理方法、数据的显示方法、数据分布特征的概括与分析方法等。推断统计（inferential statistics）主要研究如何根据样本信息来推断总体的特征。数据解释是对分析的结果进行说明，说明结果为什么是这个样子，结果所隐含的事物的特征是什么，从数据中得出了哪些规律性的结论，等等。

可见，统计学是一门有关统计数据的科学，统计学与统计数据有着密不可分的关系。在英文中，“statistics”一词有两个含义：当它以单数名词出现时，表示作为一门科学的“统计学”；当它以复数名词出现时，表示“统计数据”。从中可以看出，统计学与统计数据之间有着密不可分的关系。统计学是由一套处理统计数据的方法所组成，这些方法来源于对统计数据的研究，目的也在于对统计数据的研究。离开了统计数据，统计方法就失去了用武之地，统计学也失去了它存在的意义。而统计数据不用统计方法去分析也仅仅是一堆数据而已，不能得出任何有益的结论。

统计学提供了探索数据内在规律的一套方法。那么，什么是统计数据的内在数量规律性？为什么统计方法能通过对数据的分析找出其内在的数量规律性？我们用下面的几个例子来说明。

就单独的一个家庭来观察，每个家庭的新生婴儿的性别可能是男性，也可能是女性。如果不对生育人口进行任何限制，有的家庭的几个孩子可能都是男孩，而有的家庭的几个孩子也可能都是女孩。从表面上看，新生婴儿的性别比例似乎没有什么规律可循，但如果对大量的家庭新生儿进行观察，就会发现新生儿中男孩略多于女孩，大致为每出生 100 个女孩，相应地就有 107 个男孩出生。这个性别比例 107:100 就是新生儿性别比的数量规律，古今中外这一比例都大致相同，这是由人类自然发展的内在规律所决定的。人类社会要发展，就要保持男女人数上的大致相同。尽管从新生儿来看男性婴儿略多于女性，似乎并不平

衡，但由于男性婴儿的死亡率高于女性，到了中年时，男女人数就大体相同了。进入中老年后，男性的死亡率仍然高于女性，导致男性的平均预期寿命比女性短，老年男性反而少于女性。生育人口在性别上保持大体平衡，保证了人类社会的进化和发展。对人口性别比例的研究是统计学的起源之一，也是统计方法所探索的数量规律性之一。

我们都知道投掷硬币和掷骰子的游戏，随机地投掷一次硬币或骰子，出现正面、反面或某个点数是不确定的，完全是偶然的。但我们进行多次的重复投掷，就会发现投一枚均匀硬币出现正面和反面的次数大体相同，即比值接近于 $1/2$ 。投掷的次数越多，就越接近于 $1/2$ 这一稳定的数值。同样，在掷骰子时，出现1点~6点的比例也逐渐接近于 $1/6$ 。这里的 $1/2$ 和 $1/6$ 就是掷硬币和掷骰子出现某一特定结果的概率，也就是投掷硬币或骰子时所呈现的数量规律性。

在进行农作物试验时，如果其他试验条件相同，我们会发现某种粮食作物的产量会随着某种肥料施肥量的增加而增加。当最初增加施肥量时，产量增加较快，以后增加同样的施肥量，粮食产量的增加逐渐减少。当施肥量增加到一定数值时，产量不再增加。这时如果再增加施肥量，产量反而会减少。粮食产量与施肥量之间的这种数量关系，就是我们所要探索的数量规律性。如果我们能从大量的试验数据中用统计方法找出产量与施肥量之间的数量关系，就可以确定出最佳的施肥量，以求得最大的效益。

上述例子说明，就一次的观察或试验来说，其结果往往是随机的，但通过多次观察或试验得到大量的统计数据，利用统计方法是可以探索出其内在的数量规律性的。

二、统计的应用领域

目前，统计方法已被应用到自然科学和社会科学的众多领域，统计学也已发展成为由若干分支学科组成的学科体系。可以说，几乎所有的研究领域都要用到统计方法。下面的表1—1列出了统计的一些应用领域，目的是让我们通过简单浏览形成这样一个概念：统计学非常有用！

表 1—1 统计的应用领域

actuarial work (精算)	hydrology (水文学)
agriculture (农业)	industry (工业)
animal science (动物学)	linguistics (语言学)
anthropology (人类学)	literature (文学)
archaeology (考古学)	manpower planning (劳动力计划)