现代远程教育系列教材

www.cmr.com.cn

统计学教程

■ 主编　金勇进

中国人民大学出版社

现代远程教育系列教材

统计学教程

主编　金勇进

中国人民大学出版社

图书在版编目（CIP）数据

统计学教程/金勇进主编.
北京：中国人民大学出版社，2004
（现代远程教育系列教材）
ISBN 7-300-06074-9

Ⅰ. 统…
Ⅱ. 金…
Ⅲ. 统计学-远距离教育-教材
Ⅳ. C8

中国版本图书馆 CIP 数据核字（2004）第 120104 号

现代远程教育系列教材
统计学教程
主编　金勇进

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		
印　　刷	北京东君印刷有限公司		
开　　本	787×965 毫米 1/16	**版　　次**	2004 年 11 月第 1 版
印　　张	19.5	**印　　次**	2005 年 3 月第 2 次印刷
字　　数	303 000	**定　　价**	21.00 元

版权所有　侵权必究　　印装差错　负责调换

总 序

人类迈进21世纪，全球性的科技革命正在越来越深刻地影响着人类的生活、工作和学习方式，教育领域当然也不例外。网络教育以其鲜明的时代特色、充满希望的生命力尤为引人瞩目，已经成为教育领域里突起的一支生力军。

对于网络教育，尽管目前大家看法不尽一致，但一般认为就是利用计算机、计算机网络和多媒体等现代信息技术传授和学习知识的一种全新教育方式。正如现代教育制度产生于工业社会一样，正在兴起的网络教育反映了知识经济社会对教育的新要求。

在知识经济社会，知识对于个人的发展和社会整体进步起着决定性的作用。由于知识更新速度加快，人们对于知识的渴求也越来越强烈，“继续教育”、“终身教育”等教育理念也越来越为社会所接受。建立在计算机网络技术、多媒体技术以及现代教育学、心理学等基础之上的网络教育应运而生，它突破时间、空间的限制，为一切具有学习热情、学习能力的人敞开了接受教育的大门。学校变得没有围墙，因此极大地拓展了教育的空间，充分体现了终身教育的先进教育理念，适应了学习化社会里人们个性化学习、多样化学习的需要。

此外，网络教育的出现对于目前仍处于精英教育阶段的高等教育走向普及化有着更为重要的现实意义。进入新世纪后，经济、科技和教育

全球化趋势日益加剧，各国经济实力和综合国力的竞争更加激烈，而经济的发展、综合国力的提高，最终都取决于科技和教育的发展。也就是说，国际竞争归根结底是人力资源开发和人才培养的竞争，这已经成为各个国家的共识，我国因此也将“科教兴国”定为长期发展的战略之一。目前，我国接受高等教育的毛入学率与发达国家相比尚有很大差距。要实现高等教育大众化的目标，单靠传统高等教育手段是很困难的，而网络教育恰恰可以大有作为。网络教育利用全社会优质的教育资源，通过计算机网络等现代信息技术，可以有效解决中国高等教育资源相对短缺和教育经费投入不足等问题。教育部为此批准了几十所具有教育资源优势、办学条件成熟的高校开办网络教育，目前，通过网络接受高等教育的人数在迅速增长，网上大学生已达上百万人。因此，可以毫不夸张地说，网络教育已经成为中国在新世纪增强综合国力、实现高等教育跨越式发展的重要手段。

在这场教育变革中，中国人民大学一直走在前列。中国人民大学是一所以经济、管理等人文、社会科学为主，兼有信息科学、环境科学等的综合性、研究型大学。建校60多年来，它秉承“永远奋进在时代前列、实事求是、兼容并蓄、服务现实、艰苦奋斗”等五大传统，发展成为中国人文、社会科学人才培养和科学研究的主要基地，在中国现代化进程中发挥着不可忽视的作用，同时也成为在国际上有着广泛影响的著名学府。长期以来，中国人民大学利用自身教育资源的优势，在办好全日制高等教育的同时，一直积极开展远程教育和继续教育，不遗余力地为中国高等教育事业做出自己的贡献。随着网络时代的到来，计算机网络等信息技术的发展为远程教育提供了新的技术手段和教育方式。中国人民大学敏锐地把握住了这一教育发展的契机，充分意识到由此可能带来远程教育的蓬勃生机，从1997年开始网络教育的筹备和试验工作，并于1998年成立了国内第一所网络教育学院。经过近几年的发展，中国人民大学网络教育学院已经成为国内技术手段先进、实力雄厚的网上大学。此次出版的“现代远程教育系列教材”，中国人民大学给予了高度重视，成立了专门的编审委员会，力争在两三年内陆续推出“网上人大”财经、管理类系列教材。

按照新世纪人们学习方式的变化和网络教育的特点，出版现代远程教育系列教材，不仅是一项创新工程，也是发展我校网络教育的基础工

此为试读，需要完整PDF请访问：www.ertongbook.com

程。相信“现代远程教育系列”教材的作者们，一定能根据中国人民大学网络教育的实践经验，创造出适合网络教学和网络学习的、有较高学术水平、反映时代要求的网络教材，把正在兴起的我国网络教育事业推向新的高度。

林　岗

2003年1月

前 言

当今的时代是信息的时代。谁要在激烈的市场竞争中占有主动，赢得胜利，没有信息是不可想像的。信息是什么？是数据，大量的数据。统计数据是最实实在在的信息。所以有人说，数据就是金钱，数据就是财富。现在，统计数据已经越来越多地从政府机构的保密室中走出来，来到大家中间。不是吗？统计信息经常出现在报纸、杂志、广播和电视节目中，还有大量的统计年鉴、各种专门性的统计季报、月报。近些年来还涌现出许多调查公司、咨询公司，专门从事数据的采集和数据分析工作。

统计是关于数据的科学，是一门非常有用的知识。话虽然这么说，但仍然有许多人对统计不甚了解。例如，他们会问：为什么样本的数据能够说明总体的情况？怎样选择一个能够说明总体情况的样本？统计检验是怎么一回事？怎样判断A事物与B事物是否有联系？等等。统计学可以告诉我们如何回答这些问题。

许多学过统计学的人抱怨说，统计学是一门难学的课程，那么多的公式，那么复杂的计算，计算出来的结果还不知道用来说明什么。这些抱怨不是没有道理。分析一下可以看出，学统计的人有两类，一类是把统计当做饭碗的，这些专门从事统计的专业人员需要具备扎实的统计学专业功底，但在学统计的人群中这类人毕竟是少数。另一类是大多数，

他们有自己的专业，统计不过是他们用于分析、研究本专业问题的一个方法，一个工具。他们不需要搞清、搞懂每个数学公式的来龙去脉，只需要了解公式的含义及用它来说明的问题就行了。我们这本书就是针对后一类非统计专业的读者写的。

这本书有这样几个特点：

（1）没有复杂的数学公式和数学推导。编写中力求简明易懂，强调应用实例阐明统计方法的基本原理和思想，力图让读者感到统计学既有用，又不难学。

（2）与计算机操作结合。统计方法的普及离不开计算机技术的应用。本教材选择了最常见的 Excel 软件，将其操作方法以附录方式附在相关章的后面，以利于读者有针对性地掌握和应用。

（3）尽量生动活泼。本书力图避免生硬式的传教，在每章开始有学习导航，列出该章学习重点，结尾有小结，说明本章内容线索。章中还采用一些其他方式，如“小词典”、“想一想”、“知道吗”等专栏，引发读者的兴趣和思考。

本书是“现代远程教育系列教材”中的一本，但其使用对象不仅仅是网络学院的学生，还可以用于各类院校统计学课程的教材，也是各类人员了解统计学知识的参考书。在本书的写作过程中，我的学生们帮助我进行了大量的资料搜集和整理，对于他们所付出的辛勤劳动，我表示深深的谢意。他们是：王华博士、栾文英博士、刘欣硕士、刘薇硕士、龚娴硕士、周翠翎硕士。本书的出版得到了中国人民大学网络学院和中国人民大学出版社的大力支持，在此一并表示感谢！

金勇进

2004 年 9 月

目录

绪　论 …… 1

第一章　数据的搜集与整理 …… 8

第一节　数据的计量与类型 …… 9

第二节　统计数据的来源 …… 14

第三节　统计数据的整理与显示 …… 20

第二章　统计数据的描述 …… 40

第一节　绝对数与相对数 …… 41

第二节　集中趋势的描述 …… 42

第三节　离散趋势的描述 …… 52

第四节　数据分布形状的描述 …… 61

第三章　概率与概率分布 …… 67

第一节　概率基础 …… 68

第二节　概率分布 …… 79

第四章　参数估计 …… 105

第一节　统计推断的基本概念 …… 106

第二节　参数估计基本方法 …… 113

第三节　总体均值和总体比例的估计 … 122

第四节　两个总体均值及两个总体比例之差的估计 …… 128

第五章　假设检验 …… 139

第一节　假设检验的一般问题 …… 140

第二节　假设检验的方法 …… 147
第三节　假设检验方法的总结 …… 157

第六章　方差分析 …… 165
第一节　方差分析的基本问题 …… 166
第二节　单因素方差分析 …… 169
第三节　双因素方差分析 …… 173

第七章　列联分析 …… 182
第一节　列联表 …… 183
第二节　χ^2 分布与 χ^2 检验 …… 188
第三节　列联表中的相关测量 …… 194

第八章　相关分析与回归分析 …… 203
第一节　相关分析 …… 204
第二节　一元线性回归 …… 209
第三节　多元线性回归与一元非线性回归 …… 222

第九章　时间序列分析 …… 231
第一节　时间序列分析基础 …… 232
第二节　长期趋势分析 …… 242
第三节　季节变动分析 …… 249
第四节　循环波动分析 …… 255

第十章　指数 …… 261
第一节　指数的性质与分类 …… 262
第二节　加权指数 …… 266
第三节　指数体系 …… 278
第四节　几种常用的价格指数 …… 287

绪 论

一、统计学及其方法体系

统计作为一种社会实践活动已有悠久的历史。最初，统计只是为统治者管理国家搜集资料、提供依据。在我国，夏禹时代（公元前两千多年）就有了人口数量的记载；为了赋税、徭役和兵役的需要，历代都有田亩和户口等的记录。在国外，古巴比伦、埃及和罗马帝国也有人口和资源的详细记录；到中世纪，西欧各国都有人口、军队、领地、职业、财产的统计。外语中“统计”一词和“国家”一词来自同一词源。到现在，统计活动已经在生产、科学研究以及各类管理领域有了相当广泛和深入的应用。但不论是哪个领域的统计活动，概括起来，它们都是要收集、整理和分析统计数据，并探索数据的内在数量规律性。

目前比较公认的看法是，统计有三种含义，即统计活动、统计数据和统计学。专门收集数据资料的活动称为统计活动，其成果则表现为统计数据，它是统计研究对象的数量状况及其内在规律性的客观、真实反映，其规律性是由统计研究对象的内在必然联系决定的；统计学是一门关于数据资料的收集、整理、描述、显示和分析推断的科学，其目的是探索数据内在的数量规律性。统计数据正是统计学所要研究和处理的对象。由此三者有机地结合起来。

统计学提供了一系列众多的统计方法，专门用来收集数据、整理数

此为试读，需要完整PDF请访问：www.ertongbook.com

据、显示数据的特征进而分析和探索（或推断）出事物总体的数量规律性。抽样调查和试验设计是统计学的两个重要分支，这两方面的理论专门用于对资料数据收集方式方法的研究；统计分析推断方面则包括了统计估计、假设检验、非参数统计、方差分析、相关分析、回归分析、统计决策、时间序列分析等丰富的内容，形成了庞大而严密的方法论体系。

具体而言，统计学根据统计方法的构成可以分为描述统计学和推断统计学。描述统计学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数量特征；推断统计学则研究如何根据样本数据去推断总体数量特征，是在样本数据进行描述的基础上，对统计总体的未知数量特征做出以概率形式表述的推断。推断统计学已经成为现代统计学的核心内容。不论何种方法，由客观事物本身的特点和统计方法的特性共同决定了都要通过对数据的大量观察和处理来研究和探索数据的规律性。

二、统计学的产生和发展

虽然关于统计活动早有记载，但统计学的真正萌芽是发生在17世纪的欧洲，距今只有三百多年的历史。当时在德国和英国分别形成了国势学派和政治算术学派，这两个学派的主旨是要搜集并反映国家或社会总体的各方面信息，并尝试以量化的形式加以体现，从中发现变化规律，成为社会经济统计学的初始形态。但自从概率论被引入统计学科体系之后，数理统计学方法论体系不断得以扩充和完善，日益成为西方统计学的主流学派，作为统计学原始形态的社会经济统计学反而退居其次。即便如此，社会经济统计学仍然以相当的规模存在并有所发展，尤其是受意识形态的影响，在很多国家社会经济统计学还曾经一度占据了统计学的主导地位。在我国当前的统计学科体系建设中，社会经济统计学与数理统计学并存、进而融合为“大统计”的趋势已经日益彰显。

可见，统计学是遵循两条脉络发展的，一是由国势学和政治算术到当代的社会经济统计学，一是由概率论发展而来的数理统计学，两者时有交叉，又分歧不断。不过总体上统计学已经日益从实质性学科向方法论学科转变，成为可以为经济、社会、医学、农业、管理等众多领域提

供定量分析方法的学科体系。下面我们将统计学的发展大致划分为三个阶段来加以阐述。

（一）古典统计学时期

古典统计学时期约自17世纪中叶至19世纪初期，这段时期国势学派和政治算术学派出现并提出了“统计学”的概念，进行了众多有关统计学的实践活动。另外由于概率论对数理统计学的产生起到了至关重要的作用，因而对这一时期概率论学派的发展也一并提及。

1. 国势学派

国势学派用记述的方法研究一国的地理、人口、财政、军事、政治和法律制度等国家大事，其创始人是德国人 H. Corning（1606—1682）。至1723年，德国人 M. Schmertzel（1679—1747）在 Jena 大学创设统计学讲座。随后，法国人阿亨瓦尔（1719—1772）在 Gottingen 大学开始正式讲授统计学；“统计学”（statistik）这一名词由阿亨瓦尔首次提出，并定义其为国家显著事项的学问，言下之意是通过这门科学，可了解国家理乱兴亡之迹。最初的“统计”只是文字记载。丹麦人 J. D. Ancherson（1700—1765）首创以表式分栏排列一国的土地、人口、宗教、军事、货币及度量衡等数字，这被称为“表式统计学派”，亦属国势学派。

2. 政治算术学派

政治算术学派以英国人为主。英国是资本主义发展较早的国家，也较早地利用数字对人口和经济等方面进行记载和推断。其创始人是格朗特（John Graunt，1620—1674）和威廉·配第（Willian Petty，1623—1687）。前者于1662年出版了《关于死亡表的自然观察与政治观察》一书，他通过大量观察的方法，研究并发现了人口与社会现象中重要的数量规律性（如一切疾病和事故在全部死亡原因中占有稳定的百分比等），并对伦敦市总人口做出了比较科学的估计。后者的《政治算术》（1676）用大量的数字资料对英国、法国、荷兰三国的经济实力进行比较，用数字、重量、尺度等定量的方法进行分析比较；其对国民收入估算的方法，不仅对经济学而且对统计学的发展都具有重大意义。他们虽未创立“统计学”之名，但所用于探索社会和经济现象数量规律性的方法却具有“统计学”之实，因而公认配第是政府统计的创始人，格朗特是人口统计的创始人。稍后，德国人 Halley（1656—1742）编制了生命表。

3. 概率论学派

概率论学派最早起源于对赌博中掷骰子输赢问题的研究，其创始人包括法国的帕斯卡尔（1623—1662）和费马特（1601—1665），他们以通信的方式讨论赌博时的概率问题，在数学家们对机会游戏研究的基础上，将赌博中出现的各种具体问题归纳为一般的概率原理，为后来概率论和统计学的发展奠定了重要的基础。C. Huygens（1629—1695）著有《骰子赌博理论》，A. De Moivre（1667—1754）发现正态方程式。同一时期，瑞士数学家贝努里（1654—1705）提出二项分布理论。此后，1814年法国人P. S. Laplace（1748—1827）发表了《概率分析论》一书，构筑了古典概率理论的完整体系，并用于自然和社会现象的研究；S. D. Poisson（1781—1840）提出泊松分布；德国人K. F. Gauss（1775—1855）提出最小二乘法，他还从观察天象中发现了误差正态曲线。

（二）近代统计学时期

近代统计学时期，约自19世纪初期至20世纪初期。这一时期以比利时人凯特莱（1796—1874）为起点，他发表了《社会物理》一书，提出了偶然误差的概念，并指出某一学科的统计方法也可以应用于其他学科。在统计研究方面，他先研究天文、气象方面的统计资料，后又用统计数字研究植物界和人类社会；他以概率论作为理论基础，用大量观察和综合平均的方法进行研究，从而把概率论、国势学派和政治算术学派观察群体现象并进行数量分析的方法融合为一门统计学，奠定了近代统计学的基础。凯特莱于1851年在比利时首都布鲁塞尔召开了第一届国际统计学会议，国际统计学会议又于1855年在伦敦召开，改称为“国际统计学会”，这一名称一直沿用至今。

此外，K. G. A. Knies（1812—1898）和恩格尔（1821—1896）以大量观察法寻求社会现象规律，称为“社会统计学派”。英国人高尔顿（1822—1911）发现了百分位数，他还从研究遗传学和优生学中创立了“回归分析”的概念。皮尔逊（1857—1936）不仅完成了描述统计学的体系，还提出经验分布函数、相关分析、动差法、卡方检验和大样本理论，有人由此认为他是近代统计学的创始人。

（三）现代统计学时期

现代统计学时期自20世纪初至今。大工业的发展对产品质量检验问题提出了新的要求。1907年，英国人戈赛特（1876—1937）提出了小样

本 T 统计量理论，使统计学进入了现代统计学（主要是推断统计学）的阶段。E. Borel（1871—1956）奠定了现代概率理论的基础。英国人 R. A. Fisher（1890—1962）提出样本相关系数的分布、t 检验、实验设计和方差分析等方法和概念。A. Wald（1902—1950）提出决策理论和序贯抽样法。J. V. Neumann 和 O. Morgenstern 提出博弈论，使决策理论更加系统化。Neyman 和 Deming 提出抽样调查法，对质量控制及生产管理贡献很大。N. Wiener 的控制论和 C. E. Shannon 的信息论使推断统计学的理论更加健全。美国的大学自 1950 年把统计学设为独立的学系，1955 年开始颁授统计学的高级学位。

从 20 世纪 50 年代起，统计理论、方法和应用进入了一个全面发展的新阶段。统计学受计算机科学、信息论、混沌理论等现代科学技术的影响，新的研究领域层出不穷，如多元统计分析、现代时间序列分析、随机过程、非参数统计、贝叶斯统计、线性统计模型等等。据美国学者估计，现代统计学是以指数式加速度发展的，新的研究分支不断增加，统计应用领域不断扩展，几乎所有科学研究都离不开统计方法。自然科学、工程技术、农学、医学、军事科学、社会科学都离不开数据，对数据的研究和分析必然要用到统计方法。统计方法在各学科领域的应用又进一步促进了统计方法研究的深入和发展。统计方法与数学、哲学一样成为所有学科的基础。

上述三个阶段的划分只是大致上的，其实统计学的发展是渐进的、错综的，并没有明确的时间界限。而且从发展趋势看，统计理论研究的分支仍会不断增加，统计学将越来越具有交叉学科的性质，统计学应用的范围将更加广泛。一些过去与数字毫无联系的学科，如政治学、历史学、法学等，也正在并将更多地应用统计方法进行研究和分析。目前阻碍统计方法推广应用的主要是定性资料，如国民经济行业、部门等分类资料，不同政治观点、思想准则等等。如何对这些定性问题进行量化处理分析，是统计学家们正在研究的重大课题之一。探索性数据分析的应用前景也很广阔，这种方法重视数据的直观显示、原始数据中信息的提取及对特殊数据的识别和具体分析。贝叶斯统计理论研究将在统计推断、预测和决策等领域继续显示其强健的生命力，成为统计实践中有力的工具。

三、统计学与其他学科的关系

（一）统计学与数学的关系

统计学与数学有着密切的联系。现代统计学用到了几乎所有的现代数学知识；统计方法与数学方法一样，并不能独立地直接研究和探索自然现象和社会现象的规律，而是给各学科提供了一个研究和探索客观规律的数量方法。研究理论统计学的人需要较深的数学功底，使用统计方法的人要具有良好的数学基础。但如果把统计学理解成是应用数学的一个分支是不妥当的。

统计学与数学存在本质的区别：首先，虽然表面上看统计学与数学都是研究数量规律，跟数字打交道的，但是，数学研究的是抽象的数量规律性，而统计学研究的是具体的、实实在在的实际现象的数量规律性；数学研究的是没有量纲或单位的抽象的数，统计研究的是有具体实物或计量单位的数据。其次，统计学与数学在研究中所使用的逻辑方法也是不同的，数学研究所使用的是纯粹的演绎，而统计学是演绎与归纳相结合，占主导地位的是归纳。数学家可以坐在屋里，凭借聪明的大脑从假设命题出发推导出结果，统计学家却必须深入实际收集数据，并与具体实际问题相结合，因为没有对大量数据的科学归纳，统计学家不可能得出任何有益的结论。实际上，数学只是为统计理论和统计方法的发展提供了数学基础，而统计学的主要特征是研究数据。

（二）统计学与其他学科的关系

统计学是一门应用性很强的学科。由于几乎所有学科都要研究和分析数据，所以统计学与几乎所有学科领域都有或多或少的联系。统计学的方法可以帮助其他学科探索其学科内在的数量规律性，但若要对这种数量规律性做出内在必然联系的解释，并从中把握该学科研究实体的实际规律，则只能由各学科的具体研究来完成。

可以这样说，统计方法仅仅是一种有用的定量分析工具，它不是万能的，不能解决想要解决的所有问题。能否用统计方法解决各学科的具体问题，首先要看使用统计工具的人能否正确选择统计方法，其次还要在定量分析的同时进行必要的定性分析，也就是要在用统计方法进行定量分析的基础上，应用该学科的专业知识对统计分析的结果做出合乎规律的解释和分析，这样才能得出令人满意的结论。尽管各学科所需要的统计知识不同，所使用的统计方法的复杂程度各异，统计学也不能解决

各学科的所有问题，但统计方法在各学科的研究中将会发挥越来越重要的作用。

了解这些特点，对于我们正确看待统计学的理论地位与实际功用，并有效地学习统计学的有关理论与方法是至关重要的。

此为试读，需要完整PDF请访问：www.ertongbook.com