

148

586

统计学教程

主 编 宋光辉

副主编 许涤龙 曾昭法
王亚雄 周四军



A1040945

湖南大学出版社

2003年·长沙

图书在版编目(CIP)数据

统计学教程/宋光辉主编. —长沙:湖南大学出版社,2002

ISBN7-81053-538-2

I. 统… II. 宋… III. 统计学-教材 IV. C8

中国版本图书馆CIP数据核字(2002)第064543号

统计学教程

Tongjixue Jiaocheng

宋光辉 主编

☐责任编辑 王和君 袁作兴

☐装帧设计 张毅

☐出版发行 湖南大学出版社

社址 长沙市岳麓山 邮码 410082

电话 0731—8821691 0731—8821315

☐经 销 湖南省新华书店

☐印 装 长沙县新华彩印厂

☐开本 850×1168 32开 ☐印张 17.75 ☐字数 445千

☐版次 2002年9月第1版 ☐2003年1月第3次印刷

☐印数 35 001—40 000册

☐书号 ISBN 7-81053-538-2/C·44

☐定价 30.00元

(湖南大学版图书凡有印装差错,请向承印厂调换)

前 言

《统计学教程》是在原《统计学》的基础上重新组织编写的。本教程基本上保持了原教材的框架,但为了适应统计改革和高等统计教育课程体系调整的要求,本教程在写作思想、内容体系和写作方法上都作了较大的改变。

本教程继承了我国统计学原理教材编写的传统经验,并推陈出新。在写作思想上更加注重统计基本理论、基本知识的介绍,着力于社会经济统计理论与数理统计方法的融合,强化统计分布、动态分析、抽样推断、相关分析等理论和方法,突出统计信息、咨询和监督的整体功能。在内容体系上,对有关章节作了较大的调整,删去了原教材的社会经济统计指标体系和国民经济核算体系两章,增加了统计预测和统计综合评价分析两章;调整分立了统计调查、统计整理、假设检验三章,并适当补充了一些适用性的内容,如问卷调查、统计图等。在写作方法上,力求简明扼要,深入浅出,理论联系实际;章后思考题与练习题,紧密结合教材内容,以加强统计思维与基本技能的训练。

本教程由宋光辉主编,负责全书总体结构的设计、全书总纂和修改定稿;许涤龙、曾昭法、王亚雄、周四军任副主编。全书共12章,各章执笔人:宋光辉(第一、二、三、九、十一章)、辛玲、陈曜(第四章)、陈黎明(第五章)、刘建辉(第六章)、王亚雄(第七章)、周四军(第八章)、张立军(第十章)、曾昭法(第十二章)。

本教程比较全面系统地阐述了我国社会经济统计的基本理论和方法,可以作为高等院校统计学专业和经济管理类各专业统计

学课程的教材,对实际统计工作者和从事经济管理工作的人员也具有广泛的参考价值。

由于本书兼顾了不同对象学习的需要,因此在采用本书作为教学用书时,可根据实际情况取舍内容。高等学校本科统计专业一般单独开设概率论与数理统计学、计量经济学课程,则本书第五章、第六章第二节的内容可以合并到那些课程中去讲授。高等学校财经类各专业(不含统计专业)在采用本书时,第三章第四节、第五章、第六章第二节、第七章、第八章第三节和第四节、第十章第二节和第四节、第十一章、第十二章的内容可以略讲或不讲。

本书编写过程中,参考了许多兄弟院校的教材及统计界同仁的学术研究成果,从中获得诸多教益,限于篇幅,不能一一列举,在此谨向这些作者表示诚挚的敬意和谢意!

在新的世纪里,我国的统计事业正在蓬勃发展,社会经济统计领域正在发生深刻的变革,特别是我国加入世界贸易组织(WTO)和国际货币基金组织(IMF)数据公布通用系统(GDDS),许多新情况、新问题需要我们深入探讨,企盼本书能为此作一些探索性研究,但限于编者的水平,恐难达此目的。书中的疏漏及错误在所难免,尚祈同仁及读者批评指正。

· 编 者

2002年8月

目 次

第一章 绪论	(1)
第一节 统计与统计学	(1)
第二节 社会经济统计学的研究对象	(14)
第三节 统计工作的组织	(22)
第四节 统计指标及其测量尺度	(28)
思考与练习	(40)
第二章 统计调查	(42)
第一节 统计调查设计	(42)
第二节 统计调查方式	(48)
第三节 问卷调查	(56)
第四节 统计调查质量控制	(69)
思考与练习	(73)
第三章 统计整理	(74)
第一节 统计整理概述	(74)
第二节 统计分组	(82)
第三节 绝对数和相对数	(94)
第四节 统计表和统计图	(105)
思考与练习	(122)

第四章 统计特征值	(127)
第一节 集中趋势的测度.....	(127)
第二节 离散趋势的测度.....	(143)
第三节 偏态与峰度.....	(149)
思考与练习.....	(154)
第五章 统计分布	(157)
第一节 频率与概率.....	(157)
第二节 频率分布.....	(172)
第三节 概率分布.....	(183)
第四节 概率分布的特征值.....	(198)
思考与练习.....	(209)
第六章 抽样推断	(212)
第一节 抽样推断概述.....	(212)
第二节 随机抽样的概率分布.....	(220)
第三节 参数估计.....	(226)
第四节 抽样设计.....	(241)
第五节 多阶段抽样.....	(251)
思考与练习.....	(257)
第七章 假设检验	(260)
第一节 假设检验概述.....	(260)
第二节 u 分布检验法和 t 分布检验法.....	(262)
第三节 χ^2 分布检验法和 F 分布检验法.....	(270)
第四节 统计检验的两类错误.....	(272)
思考与练习.....	(275)

第八章 时间数列	(278)
第一节 时间数列概述.....	(278)
第二节 时间数列发展水平与速度指标.....	(282)
第三节 长期趋势的测定方法.....	(296)
第四节 季节变动的测定方法.....	(312)
思考与练习.....	(318)
第九章 统计指数	(323)
第一节 统计指数的概念与分类.....	(323)
第二节 统计指数的编制方法.....	(332)
第三节 指数体系与因素分析.....	(342)
第四节 指数数列.....	(353)
第五节 统计指数的应用.....	(355)
思考与练习.....	(365)
第十章 相关与回归	(370)
第一节 相关分析概述.....	(370)
第二节 等级相关与品质相关.....	(382)
第三节 简单线性回归模型.....	(386)
第四节 多元线性回归模型.....	(401)
思考与练习.....	(413)
第十一章 统计预测	(416)
第一节 定性预测.....	(416)
第二节 平滑预测.....	(422)
第三节 趋势线预测.....	(431)
第四节 季节预测.....	(444)
第五节 回归分析预测.....	(457)

思考与练习	(469)
第十二章 统计综合评价分析	(473)
第一节 统计综合评价分析概述	(473)
第二节 指标赋权方法	(479)
第三节 综合评价分析的常规方法	(484)
第四节 无标准排序评价法	(492)
第五节 灰色关联度综合评价法	(495)
第六节 模糊综合评价分析	(501)
第七节 主成分分析	(518)
思考与练习	(528)
附录 常用统计数值表	(533)
附表 1 泊松分布表	(533)
附表 2 标准正态分布表	(535)
附表 3 正态分布分位数表	(538)
附表 4 t 分布表	(541)
附表 5 χ^2 分布表	(543)
附表 6 F 分布表	(546)
附表 7 Spearman 等级相关统计量	(554)
附表 8 Kendall T 统计量	(557)
附表 9 相关系数 $\rho=0$ 的临界值表	(559)

第一章 绪 论

第一节 统计与统计学

一、统计的基本涵义

日常生活中,人们常说的统计,一般包括三个方面的涵义,即统计工作、统计资料和统计学。

统计工作,即统计实践活动,是人们利用各种科学的统计方法,搜集、整理、分析和提供有关社会、政治、经济、文化和各种自然现象等数字资料工作的总称,是人们对客观事物数量方面进行调查研究的一种认识活动。

统计资料,亦称统计信息,包括原始信息和再生信息,是统计工作活动过程所取得的成果,是反映被调查研究的客观事物及其过程的数字资料以及与之相联系的文字资料、图表资料等。

统计学,即统计理论,是统计工作实践经验的总结和理论概括,是研究如何对客观事物数量方面进行调查、整理和分析的原理、原则与方式、方法的科学。

二、统计学的产生与发展

统计起源很早,它是随着社会生产的发展和国家管理的需要而产生和发展起来的。早在原始社会,人类就有了一般的计数活动。在奴隶制社会的形成过程中,就有了社会经济统计的萌芽。我国在公元前 2000 多年的夏禹时代,就有了人口、土地等的历史

记载。禹时,分中国为九州,人口 1355 万人,土地约 2438 万顷。封建社会我国的统计已略具规模。战国时期秦国的商鞅(约公元前 390~338 年)就提出国家之强盛必须知“境内令、口(府)之数,壮男、壮女之数,老弱之数,宦士之数,以言谈取食之数,利民之数,马、牛、刍、藁之数”等所谓十三数;至秦汉,已有地方田亩和户口的记录;唐宋则有计口授田、田亩鳞册等土地调查和计算;明清时期,建立了经常的人口登记和保甲制度等等。在世界,古埃及在公元前 3000 年已有人口、财产数字方面的记载。古希腊在公元前 600 年就进行了人口普查;古罗马公元前就建立了出生、死亡登记制度。不过在奴隶社会和封建社会里,由于生产力水平很低,社会统计工作也只是为了适应奴隶主和封建主王朝实现赋税、徭役、征兵等需要而进行的人口、土地、财产等方面的登记和简单的计算工作。17 世纪,欧洲资本主义发展,工业、商业、农业、外贸、海关、物价、交通、航运等都进入了一个空前的发展阶段,统计也适应社会经济的发展而得到广泛的发展,统计工作开始从国家管理领域扩展到社会经济活动的许多领域。出现了许多专业的统计机构和研究组织,统计逐步成为社会分工中的一个独立部门和专业。统计工作的发展,统计实践经验的日益丰富,加上数学及其他科学发展的影响,统计作为一门科学也应运而生。从 17 世纪下半叶开始,欧洲出现了一些统计理论著述,并在学术争鸣中逐步奠定了统计学的科学基础,20 世纪初至今,统计学在广采博纳,兼收并蓄中逐步发展成为现代统计科学体系。

从统计学的产生和发展过来看,可以大致分为古典统计学、近代科学和现代统计学三个发展时期,每一个发展时期又有不同的学术派别和学术特点。

(一)古典统计学时期

古典统计学时期是指 17 世纪中叶到 18 世纪中叶统计学的初创时期。这个时期主要有两个大学派,即政治算术学派和记述学

派。

1. 政治算术学派

统计学史认为,统计学是 17 世纪中叶在英国产生的。英国人威廉·配第(W·Petty,1623~1687)所著《政治算术》一书的问世,标志着古典统计学的诞生。《政治算术》成书于 1671-1676 年,1690 年正式出版。书中的“政治”是指政治经济学,“算术”是指统计方法。配第在书中运用大量的统计资料,对英、法、荷三国的国情国力,作了系统的对比分析,反映英国的潜在力量,提出英国社会经济发展的方向道路。配第主张一切论述都通过数量的表述,以数字、重量和尺度来说话,并开始用朴素的统计图表形式来概括数字资料,这些方法为统计学的创立奠定了方法论的基础。马克思称威廉·配第是“政治经济学之父”,在某种程度上也可以说是统计学的“创始人”。

与威廉·配第同时,英国人约翰·格朗特(J·Granunt,1620~1674)的《关于死亡表的自然与政治的观察》(1662 年)也被认为是统计学初创时期的代表作。当时,伦敦时疫流行,死亡情况严重,引起社会不安。格朗特根据英国政府定期发布的关于出生和死亡公报,对伦敦人口的出生率、死亡率、性别比例和人口发展趋势,作了分类计算和分析,证明没有必要悲观。他广泛应用各种方法,批判地整理和运用统计资料,最早编制死亡率表。尽管从现代统计学的观点来看,书中应用的方法是比较粗糙的,但其所采用的数量对比分析在当时具有独创性和新颖性,对统计学的创立,与《政治算术》一样,起了同等重要的作用。

2. 记述学派

记述学派亦称国势学派,产生于 17 世纪中叶的德国。它是由德国人海·康令(H·Conring,1606~1682)所创立的。康令最先在德国赫尔莫斯达德大学讲授《欧洲最近国势学》,奠定了国势学的基础。所谓国势学是以文字记述国家显著事项之学,是用对比分

析的方法研究关于国家组织、人口、军队、领土、财产等国情国力,比较各国势力强弱为德国君主服务的。后来德国许多大学都开设国势学这门课程,故人们称国势学派为德意志大学教授派。国势学派的主要继承人及中期代表人物高·阿亨瓦尔(G·Achenwall, 1719~1772)在《近代欧洲各国国势论》一书中,首次使用“统计学”(Statistik)一词作为国势学的名称。不久这个词传到英国,被译为英语 Statistics,取代了“政治算术”这一名称。到 19 世纪中叶,欧美许多国家建立了政府统计机构,发行统计刊物,成立了统计协会,一提到统计,人们一般都把它同用数量方法研究社会经济问题的活动联系起来。在这种历史背景下,德国经济学家兼统计学家克尼斯(K.G.A.knies, 1821~1897)于 1850 年发表《作为独立科学的统计学》著名论文,指出阿亨瓦尔的所谓统计学的研究对象和研究方法是历史学的一个组成部分,而政治算术学派是以社会经济客观现象的数量关系为依据来探索规律性的科学,主张把统计学(Statistics)作为政治算术的科学命名。克尼斯的建议得到经济学界和统计学界的普遍赞同,从而正式确立了统计学这门科学的名称。

(二)近代统计学时期

近代统计学时期是从 18 世末到 19 世纪末统计学的重大发展时期。这一时期又形成了许多学派,其中主要是数理统计学派和社会统计学派。

1. 数理统计学派

数理统计学派产生于 19 世纪中叶,它是在概率论已有相当发展的基础上,把概率论引进统计学而形成的。1665 年法国数学家帕斯卡(B·Pascal)和费尔马(P·Fermat)研究概率问题,其理论源于赌博及机会游戏,提出了“分赌难题”的一般解法。1713 年瑞士数学家丁·贝努里(J·Bernoulli, 1654~1705)的《推测术》出版,提出了贝努里大数定律。1763 年英国统计学家贝叶斯(Bayes theo-

rem)首先将归纳推理用于概率理论,提出了“逆概率公式”。(“假设概率公式”),即贝叶斯定理,创立贝氏统计理论。1809年德国数学家高斯(C.F.Gauss,1777~1855)和法国数学家、统计学家拉普拉斯(P.S.Laplace,1749~1827)分别研究和建立误差理论而发展正态分布理论,1812年,拉普拉斯的《概率分析理论》出版,系统地将数学分析应用于概率理论,推广其研究与应用。1840年比利时统计学家、数学家凯特勒(L.S.T.Quetelet,1796~1874),将概率理论应用于人口学、社会学及保险学,创立“平均人”的概念。凯特勒是一位多才多艺的科学家(统计学家、数学家、天文学家、人类学家和物理学家),又是社会活动家。他继承德国的国势学派和英国的政治算术学派的传统,把统计学作为管理国家行政的方法,扩展到作为研究社会内在矛盾及其规律性数量表现的科学认识方法。他积极主张把概率论引进统计,以研究社会现象的偶然变化中规律性的表现。他的统计代表作《社会物理学》等,认为社会现象如同自然现象一样,也具有内在必然性,一切事物都受“大数法则”的支配,正态分布可适应许多学科领域,统计学是应用于任何科学研究方法的最一般方法,是一门通用科学。凯特勒把概率论引进统计学,开辟了统计学的新领域,提高了统计计量上的准确性,使统计学产生了质的飞跃。欧美统计学界称凯特勒是近代“统计学之父”。

数理统计学在发展初期,着重于现象总体特征的描述和比较,人们称之为描述统计学,以高尔登、K·皮尔逊等人为代表;进入20世纪20年代,以费希尔等人为代表,发展成为用随机样本推断有关总体,人们称之为推断统计学。20世纪70年代后,统计与计算机密切结合,人们称之为当代统计学。

近代统计学数理统计学派的代表人物,英国统计学家、生物学家高尔登(F.Galton,1822~1911)于1880年出版《正常人遗传性身高的回归》、《自然遗传》等著作,应用统计方法研究遗传因果关

系,首创相关与回归的概念及其系数的计算方法,并致力于探索简缩数据的新途径和新方法,把大量已知数归纳为几个简单的公式用于描述比较,从而设计出中位数、四分位数、百分位数、四分位偏差,相关与回归等新方法。1885年,英国数理论统计学家、生物统计学派的代表人物皮尔逊(K. Pearson, 1857~1936),将统计学与概率论融为一体,求出能通用的数学公式描述总体,提出了一系列曲线;皮尔逊Ⅰ型分布(β)、皮尔逊Ⅲ分布(γ 分布);推导出卡方(χ^2)分布,用以检验观察与期望值之间的差异显著性;首次将“动差”一词移植到数理统计学;继承和发展了高尔登回归和相关的理论,并提出了复相关与净相关的计算方法,将相关与回归的理论扩展到许多领域,为大样本理论奠定了基础。1900年又提出了假设检验和 χ^2 分布。

此外,1885年英国的艾奇沃斯(F. Y. Edgeworth, 1845~1926)的《统计方法》、1901年鲍莱(A. L. Bowley, 1869~1957)的《统计学原理》和《统计学基础教程》(1911)、1911年尤尔(G. U. Yule, 1871~1951)的《统计学原理导论》等著作,系统地介绍了数理统计学的初期发展阶段描述统计学的基本方法。

2. 社会统计学派

在凯特勒的影响下,近代统计学无论在理论和方法上都有系统的发展。但是,由于研究的领域、内容和方法各有侧重,因而形成了近代统计学的数理统计学派和社会统计学派。

社会统计学派是19世纪下半叶与数理统计学派大体同时产生与发展的。社会统计学派以德国为中心,由德国大学教授克尼斯(K. G. A. Knies, 1821~1898)首创,主要代表人物有恩格尔(C. L. E. Engel, 1821~1896)以及梅尔(G. V. Mayr, 1841~1925)等人。

恩格尔是德国统计学家、经济学家,曾任萨克逊王国统计局长和普鲁士王国统计局长,既有丰富的统计实践经验,又有深湛的统

计理论。他提出了统计调查、整理和分析三阶段的统计方法。他的主要著作有《统计学独立的科学和方法》、《比利时工人家庭的生活》等,他所提出的著名恩格尔规律,对社会统计的研究有很大的影响。

梅尔是社会统计学派的另一代表人物,他认为统计学是根据数量的大量观察,对人类社会生产的状态及其生产的规律,作有系统的说明和研究。他明确提出社会统计学的研究对象是社会经济现象的规律,并强调社会统计学同时具有“实质性科学”与“方法论科学”的性质,认为脱离规律的认识,就不能获得科学的认识。所以,统计学不仅要确定事实数量的记述,而且还必须研究其规律性。他主张统计技术与统计科学结合起来。

社会统计学派在 19 世纪中叶到 20 世纪有很大发展,只是由于数理统计学派的兴起,逐渐失去其优势。但并没有消失,在社会主义国家中还得到了发展,即便在某些资本主义国家,如西欧各国以及日本等国也都有很大的影响。

(三)现代统计学时期

现代统计学时期是指从 20 世纪初迄至今日的统计学发展时期。20 世纪是人类历史上科学技术发展最为迅猛的时期,社会生产力得到了空前的提高,全球经济有了巨大的发展。统计学在 20 世纪进入了一个新的发展时期。这一时期的主要统计学派有西方的现代数理统计学派和东方的社会经济统计学派。

1. 现代数理统计学派

20 世纪,数理统计学同自然科学、工程技术科学紧密结合而被广泛应用,统计理论和方法有了巨大发展。首先,在随机抽样的基础上,建立了推断统计的理论和方法,所谓推断统计,即通过随机样本来推断总体数量特征的方法。推断统计学的先驱者是爱尔兰统计学家戈塞特(W.S.Gosset, 1876~1937)。他曾任都柏林市靳尼斯啤酒厂酿造化学技师,长期从事试验和统计分析工作。他

引进数理统计方法,提出小样本理论,创立了用小样本代替大样本的方法,为小样本理论奠定了基础。1908 年开始,他提出了小样本的 t 分布,发表了《平均数的概差》等论著,在抽样数目减少的情况下,进行统计推断的新途径;并研究样本标准差、样本平均差对样本标准差之比,以及相关系数的抽样分布;在田间试验中,导入了各种实验配置方法。戈塞特的论著虽然不多,但具有很高的学术价值。他所得到的 t 统计量的精确分布形式,是统计量精确分布理论中一系列重大理论成果的开端。英国统计学家费希尔(R.A.Fisher, 1890~1962),是现代数理统计学和数学遗传学的创始人之一,是推断统计学的代表人物。他曾任商业与通用投资公司的统计员、英国剑桥大学优生学和遗传学教授,曾长期在罗萨姆斯特德(Rothamsted)实验站从事研究工作,推进了数理统计的研究,充实了数理统计的内容。他创立了“实验设计”,研究自然现象所用的随机模型。他首创方差分析这一有效的统计方法,建立了点估计与区间估计的统计理论;对 t 分布进行数理论证;最先提出了 F 分布理论,使小样本分布理论在数理统计应用上占据了重要地位;将随机法导入田间试验,提出了随机区组或拉丁方格法。其《研究人员应用的统计方法》一书,是描述统计向推断统计发展的划时代的数理统计学巨著,对现代统计学界产生了深远的影响。他的《自然选择的遗传理论》(1903 年)、《实验设计》(1935 年)、《近亲繁殖理论》(1949 年)、《统计学方法和科学推断》(1956 年)以及与耶茨(F. Yates)合著的《为生物学、农学、医学研究应用的统计表》(1938 年)等著作,系统地阐述了现代数理统计学(推断科学)的基本方法。欧美统计学界认为,从 20 世纪 20 年代开始,数理统计的发展进入了费希尔时代(Fisher Era)。1936 年,波兰统计学家尼曼(T. Neyman, 1894~1981)以及 E.S. 皮尔逊等人建立了严格的统计假设理论和区间估计理论,把检验问题和区间估计明确提为一个数学上的最优问题。后来,美国统计学家瓦尔德

(A. Wald, 1902~1952)研究统计决策函数,将统计学中的估计和假设理论加以归纳,创立了“决策理论”。美国威尔克斯(S.S. Wilks, 1906~1964)、英国的威尔特(J. Wishart, 1898~1956)等统计学家对样本分布理论加以充实和发展;美国的科克伦(W. G. Cochran, 1909~1980)等 1957 年又提出了实验设计的理论和方法,进一步拓宽了统计学的范围。抽样调查自 1891 年挪威统计局长基也(A. N. Kiaer, 1838~1919)首创人口抽样调查后,逐步受到统计界的重视。1963 年 W. G 科克伦提出抽样技术,解释各种类型的抽样设计。中国著名数学家、统计学家许宝禄,不仅在概率论方面、研究样本方差的渐近展开对中心极限定理中误差大小的阶做出了精确估计,并在独立和极限分布论中取得重大成果;在数理统计,特别是在多元分析与统计推断方面也做出了重大贡献;其《抽样论》丰富了我国抽样调查的技术与理论,对我国抽样调查技术方法的推广也起了重大的作用。

20 世纪 60 年代以后,数理统计学的发展有三个明显的趋势:一是随着数学的发展,数理统计学越来越广泛地应用数学方法;二是数理统计学的发展出现了许多新的分支和以数理统计为基础的边缘学科,新分支如抽样调查理论、非参数统计、多变量分析和时间序列分析等;边缘学科如经济计量学、工程统计学、天文统计学等;三是电子计算机的发展,使数理统计的应用日益广泛深入。20 世纪数理统计学的飞速发展,使数理统计学派成为国际现代统计学的主流派。

2. 社会经济统计学派

社会经济统计学是在俄国十月革命胜利以后,首先在苏联建立起来的。二次世界大战以后,各社会主义国家都先后建立和发展了社会经济统计学。社会经济统计学是以辩证唯物主义和历史唯物主义以及马克思主义的政治经济学为指导,研究大量社会现象的数量方面,通过对社会现象中各种数量关系的研究来认识社