

统计学基础

刘玉玫 编著

$$F(x) = P\{X \leq x\}$$

probability distribution

 中国统计出版社
China Statistics Press

统计学基础

刘玉玫 编著

 中国统计出版社
China Statistics Press

(京)新登字 041 号

图书在版编目(CIP)数据

统计学基础/刘玉玫编著.

-北京:中国统计出版社,2002.8

ISBN 7-5037-3879-0

I. 统…

II. 刘…

III. 统计学

IV. C8

中国版本图书馆 CIP 数据核字(2002)第 056975 号

作 者/刘玉玫

责任编辑/杨映霜

封面设计/刘国宁 张建民

出版发行/中国统计出版社

通信地址/北京市西城区月坛南街 75 号 邮政编码/100826

办公地址/北京市丰台区西三环南路甲 6 号

电 话/(010)63459084、63266600-22500(发行部)

印 刷/科伦克三莱印务(北京)有限公司

经 销/新华书店

开 本/850×1168mm 1/32

字 数/230 千字

印 张/9.5

印 数/1-3000 册

版 别/2002 年 8 月第 1 版

版 次/2002 年 8 月第 1 次印刷

书 号/ISBN 7-5037-3879-0/F·1465

定 价/17.80 元

中国统计版图书,版权所有,侵权必究。

中国统计版图书,如有印装错误,本社发行部负责调换。

序

统计学是一门“对数据进行收集、分类、分析和解释的科学。”1994年版《兰登书屋大学词典》(《The Random House College Dictionary》)简言之,统计学是关于数据的科学。数据含有信息,这是不争的事实。但是任何数据都不免受到随机性的干扰,要从数据中提取有用信息,必须借助统计技巧消除随机干扰,得出建立在统计证据基础之上的科学结论。因此统计学成为各学科必不可少的研究工具,统计知识成为各类专业人员的必备知识。君不见,诺贝尔经济学奖得主有多少本身就是统计学家,又有多少在其研究中大量运用或依靠了统计!

统计使人豁达。随着我国市场经济的不断发展,统计分析方法受到各界人士的日益重视,各行各业的相关人士都迫切要求学习并掌握数据分析知识及应用。近年,我国出现了越来越多的统计学著作,其中一部分来自于国外专家,另一部分出自于国内学者。但是,这些著作要么侧重于理论知识,强调公式推导;要么侧重于实际应用,忽略基本理论。能够将这两方面紧密结合、浑然一体的著作的确不多。刘玉玖副教授编著的《统计学基础》,既阐述了统计的基本方法,又应用了经济中现实问题,同时还结合了计算机的具体操作,可以说不失为一本较好的将理论知识与实际应用相结合的著作。

最初的统计学是国家重大事项的记述,后来,威廉·配第采用了数量分析的方法对比英国、法国和荷兰的“财富和力量”,约翰·

格朗特利用大量数据研究人口的变动规律,阿道夫·凯特勒运用概率论原理分析社会经济现象。统计学产生和发展的历史告诉我们,统计学是关于数据的应用科学。遗憾的是以往我国的统计学教材,一方面,大多注重理论知识的灌输,忽视运用理论知识解决实际问题;另一方面,大多教材的对象为财经院校学生,而非广大的普通院校学生。

当时间推移到今天的 21 世纪,我国已正式加入 WTO,并且已经步入经济全球化的行列,与时俱进,教育是基础,教材是关键,因此,为了解决上述问题,我国统计学教材必须进行革新,编著出适应时代适应市场的教科书。刘玉玫老师编著的《统计学基础》,主要有基础性、实用性和工具性的特点。基础性表现在理论知识的阐述注重基本原理和基本思想,适用于普通高校经济管理专业的学生;实用性体现在引进适量的典型案例及其分析,并运用计算机技术进行数据处理,以提高学生和读者理解问题和解决问题的能力;工具性表现在叙述方法强调实际例解,避免过多的理论推导。

笔者相信,本书的出版,对于我国统计学教材的革新是一个启示。只要大家都关心并支持统计学教材的建设,我国统计学教材市场必将出现越来越多日臻完善的统计学著作。

邱 东

2002 年夏于东北财经大学

前 言

随着我国市场经济的不断发展,统计分析方法的应用越来越广泛,无论是国民经济管理还是企业经营决策,无论是在读大学生还是科学研究人员都越来越依赖于数据分析方法,他们迫切希望寻找到一本理论知识与实际应用技巧相结合的工具书。近年来,虽然我国统计学教材逐渐改变了理论过多、实际不足的现象,但仍缺少适合普通高校经济管理专业的教材。本书编者在总结十余年教学经验,汲取国内外同类优秀教材精华的基础之上,完成了本书的编写工作。本书既可用作高校经济管理专业学生、MBA 学员的教材,也可作为其他专业学生和各领域实际工作者的参考书。

本书在编写过程中力求由潜入深、循序渐进,言简意赅、通俗易懂,其特色表现在:一方面,注重理论知识的基础性,对于统计学的基本理论和基本思想进行了较为全面的阐述,同时为了检验基本原理的掌握程度,每章都附有练习题供学生和读者使用。另一方面,强调统计方法的工具性和实用性,避免过多的公式推导,借助实例阐述基本原理,并运用大量生动案例分析,以便提高学生和读者应用统计方法分析和解决实际问题的能力,更为重要的是,本书独辟蹊径,将应用最为广泛的 Excel 软件分析系统与典型案例有机结合起来,形成完整的一章进行专门介绍,使学生和读者既学会了统计分析方法,也了解了如何运用计算机操作进行数据分析,这在我国目前统计学教材中尚属首次出现。

全书共分 10 章,第 1 章叙述统计的发展及其应用,第 2 章说

明数据的收集方法,第3章、第4章讲述描述统计方法,第5章、第6章阐述推断统计方法应用,第7章、第8章、第9章为统计分析及预测方法的基础知识,第10章介绍计算机在统计中的具体应用方法。其中,第1章至第9章为刘玉玫撰写,第10章及书中计算机应用部分为侯彤璞撰写。

东北财经大学邱东教授对本书提出了宝贵的意见,在此表示衷心的感谢。在本书的编写过程中,编者参阅了众多国内外专家的有关著作,颇受启示,在此深表谢意,同时非常感谢东北大学的金秀副教授、辽宁石油化工大学的于兆河老师、大连大学的单艺斌教授、鞍山科技大学的姜玉华副教授以及其他给予本书编写及出版大力帮助和支持的所有人士。

由于编者水平所限,书中难免有不当之处,恳请专家和读者批评指正。

编著者

2002年7月8日

目 录

第 1 章 绪论	(1)
第 1 节 统计学的发展	(1)
第 2 节 统计学的分科	(4)
第 3 节 统计学在经济管理中的应用	(6)
第 4 节 统计学与其它学科的关系	(6)
第 5 节 统计学中的若干基本概念	(8)
习题一	(10)
第 2 章 数据的搜集	(11)
第 1 节 数据的特征	(11)
第 2 节 数据的类型	(12)
第 3 节 数据的搜集	(14)
案例 1 第一次全国农业普查方案(节选)	(23)
习题二	(25)
第 3 章 统计表与统计图	(26)
第 1 节 数据的概括	(26)
第 2 节 定性数据的图形描述	(29)
第 3 节 定量数据的图形描述	(31)
案例 2 个人投资者选择证券组合的理由	(33)
习题三	(37)
第 4 章 数据描述的数字方法	(38)
第 1 节 集中趋势的测度	(38)
第 2 节 离散程度的测度	(45)
第 3 节 偏度与峰度的测度	(50)

案例 3 大学毕业生的表现	(52)
习题四	(57)
第 5 章 概率与概率分布基础	(60)
第 1 节 随机事件	(60)
第 2 节 随机事件的概率	(63)
第 3 节 离散型概率分布	(68)
第 4 节 连续型概率分布	(72)
案例 4 机票预售数量的统计研究	(79)
习题五	(81)
第 6 章 抽样与抽样估计	(82)
第 1 节 抽样的基本概念	(82)
第 2 节 简单随机样本的抽样分布	(87)
第 3 节 总体参数的估计	(92)
第 4 节 其它抽样方法及抽样分布特征	(104)
案例 5 公司总经理的平均年收入是否有效	(114)
习题六	(115)
第 7 章 回归分析	(119)
第 1 节 回归分析的基本概念	(119)
第 2 节 简单线性回归分析	(123)
第 3 节 多元线性回归分析	(139)
第 4 节 一元非线性回归分析	(147)
案例 6 质量控制中的简单线性回归分析	(149)
案例 7 子女身高与父母身高的回归分析	(153)
习题七	(156)
第 8 章 时间序列分析	(159)
第 1 节 时间序列的成分	(159)
第 2 节 时间序列的水平分析	(162)
第 3 节 时间序列的速度分析	(167)
第 4 节 长期趋势分析及预测	(171)
第 5 节 季节变动分析	(182)
案例 8 囤积粮食可以创高价吗	(184)

习题八	(188)
第 9 章 指数	(190)
第 1 节 指数的概念与分类	(190)
第 2 节 综合指数	(192)
第 3 节 平均指数	(197)
第 4 节 几种重要的经济指数	(200)
第 5 节 指数体系与因素分析	(206)
案例 9 国债指数的编制与分析	(211)
习题九	(219)
第 10 章 计算机在统计中的应用	(222)
第 1 节 中文 Excel 及其统计应用功能	(222)
第 2 节 在描述统计中应用 Excel	(253)
第 3 节 在推断统计中应用 Excel	(258)
第 4 节 SAS 软件简介	(265)
附录 常用统计表	(268)
主要参考书目	(292)

第 1 章 绪 论

第 1 节 统计学的发展

一、统计活动的产生和发展

统计作为一种社会实践活动已有悠久的历史。它是适应社会经济的发展,适应国家管理的需要而产生和发展的。自从有了国家,便有了统计实践活动。

我国在原始社会末期,在奴隶制形成的过程中,就已经出现了统计的萌芽。早在公元前 22 世纪的夏禹王朝,将中国分为九州,人口 1355 万人,这是我国最早的土地与人口调查的资料。到了西周已建立了统计报告制度——日成、月要和岁会,这项制度经春秋战国,至秦统一中国,形成了比较完善的“上计”报告制度。同时,也逐渐形成了“举事必成,不知计数不可”和“强国知十三数”的统计思想。到了封建社会,统计不仅已略具规模,而且已居于当时世界先进水平。秦始皇最早建立了全国规模的人口调查登记制度,东汉进行全国性土地测量,唐代有了计口授田制度,宋、明有田亩鱼鳞册等等。

在西方,统计也有很大发展。埃及在公元前 27 世纪,进行过全国人口和财产调查。16 世纪初,伦敦出现“死亡公报”公布死亡人数,其记录的人数,从开始的瘟疫死亡,扩充为受洗礼死亡,到世

纪末的疾病死亡。在 17 世纪至 18 世纪资本主义上升时期,包括人口、工业和农业的“国情普查”逐渐形成制度,商业、工业、农业、海关、外贸、物价等方面的统计,也得到广泛发展。19 世纪中叶之后,数学,特别是古典概率的引入,使统计活动有了很大飞跃。20 世纪 30 年代各国普遍采用了随机抽样的方法,20 世纪 40 年代起电子计算机的应用,不仅提高了统计数据处理的效率和准确性,更为数据的积累、存储、更新、检索、加工及其一体化的自动化创造物质技术基础,20 世纪 50 年代,以国民经济为整体的统计研究方法(国民经济帐户体系及投入产出法)的推广应用,大大提高了统计的认识能力。

二、统计学的产生和发展

封建社会生产的落后,限制了统计实践活动和统计理论的发展。资本主义经济的发展,特别是现代化大生产,促进了统计活动和统计科学的发展。

国势学派或记述学派的创始人是 17 世纪中叶德国的康令(H. Conring, 1606—1681),他在西尔姆斯特大学开设“国势学”课程,其内容是分门别类,记述包括土地、人口、立宪、社会秩序、行政、财政等等的“国家显著事项”。阿亨瓦尔(G. Achenwall, 1719—1772)是国势学派的主要继承人,他在《近代欧洲各国国势学论》中首先使用“统计学”这个名称。该学派的特点是以文字记述为主,但并未把数量分析作为基本特征。

政治算术学派的代表人是英国的威廉·配第(W. Petty, 1623—1687),他的代表著作《政治算术》的问世,标志着古典政治经济学的产生,也标志着统计学的诞生。这本书运用数量分析的方法对比英、法、荷三国的“财富和力量”,以批驳英国当时的悲观论断。他的“用数字、重量和尺度来表达自己的问题”的观点,奠定了统计学的方法论基础。马克思评价他是“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”。政治算术学派的另

一创始人约翰·格朗特(J. Graunt, 1620—1674)的代表作是《对死亡公报的自然和政治观察》,它根据“死亡率公报”,分析伦敦人口出生与死亡的数量关系,他提出一个要在多年内形成的规律,是需要进行多次观察的,体现了统计学大数法则的观点。该学派以数量分析为特征,尽管一直未正式采用“统计学”的命名,但从其研究的内容和方法来看应是统计学的正统。

数理统计学派产生于 19 世纪中叶,创始人是比利时的统计学家凯特勒(A. Quetele, 1796—1874),他的代表著作是《社会物理学》,书中论述了社会生活中随机现象也具有必然的规律性,可以根据概率计算表明规律的存在形式。凯特勒对统计学发展的最大贡献,就是把古典概率引入统计学,所以有人称他为“现代统计学之父”。随着概率论研究方法在各个领域应用的迅速发展,数理统计学又逐渐形成两个分支,早期着重于现象总体数量特征的描述和比较,称之为描述统计学,以皮尔逊(K. Pearson, 1857—1936)为代表,到了 20 世纪 20 年代,统计方法发展到以随机样本为基础推断有关总体数量特征的方法,称之为推断统计学,创始人是费歇尔(R. A. Fisher 1880—1962)。

社会经济统计学,是在马克思主义奠基人广泛使用统计方法进行资本主义的经济分析过程中,逐步建立并发展起来的,它适应社会主义建设的需要。列宁强调指出,社会经济统计是认识社会的有力武器。

毋庸置疑,数理统计方法有其科学性,但在其后续发展中,曾一度造成理论研究与实际应用脱节的现象。社会经济统计方法有其历史性,但贬低数理统计方法在社会经济现象中的应用,大大限制了学科的发展。近年来,经过国际统计界的反思,数理统计方法与社会经济统计方法的融合日趋密切,数理统计方法已成为解决社会经济问题必不可少的手段。

第 2 节 统计学的分科

一、统计的涵义

“统计(Statistics)”一词最早是在德国使用的,那是阿亨瓦尔为自己所授“国势学”课程的命名—Statistik。18 世纪末,一位英国人将德语“Statistik”译为“Statistics”传入英国,并赋以用数字作为语言来表述事物的新内容。后来,这个词很快流传到美国,时至今日,已广泛传播到世界各国,约 20 世纪初才传入我国。

今天,“统计”一词已被人们赋予多种含义,在不同场合,“统计”一词可以具有不同的理解。它可以是统计工作,也可以是统计资料,还可以是统计学。

统计工作,是根据科学的方法搜集、整理和分析统计数据的实践活动。

统计资料,是统计工作过程所获得的各种有关资料的总称,是统计工作的成果。

统计学,是一门搜集、整理和分析统计数据,探索数据的内在规律性,以达到认识客观事物的目的的科学,即是关于数据的科学。它是统计工作实践的理论概括和科学总结。

二、统计学的分科

随着统计方法在各个领域的广泛应用,统计学已发展成为具有多个分支学科的学科体系。

1. 描述统计学和推断统计学

按统计方法的不同构成,可以将统计学划分为描述统计学(Descriptive Statistics)和推断统计学(Inferential Statistics)。

描述统计学,是研究如何取得反映客观现象的数据,并通过图表形式对所搜集的数据进行加工处理和显示,进而概括并分析得

出反映客观事物规律性的数量特征。它包括统计数据的搜集方法、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。

推断统计学,是研究如何根据样本(Sample)数据去推断总体(Population)数量特征的方法,它是在对样本数据进行描述的基础上,对统计总体的未知数量特征作出以概率形式表述的推断。

描述统计学和推断统计学是统计方法的两个组成部分。描述统计学是整个统计学的基础,推断统计学则是现代统计学的主要内容,并且,它在现代统计学中的地位和作用越来越重要,已成为统计学的核心内容。这是因为,在对现实问题的研究中,主要是采用抽样调查取得样本的数据资料。但是,这并不等于说描述统计学不重要,如果没有描述统计学搜集可靠的统计数据并提供有效的样本信息,即使再科学的统计推断方法也难以得出切合实际的结论。从描述统计学发展到推断统计学,既反映了统计学发展的巨大成就,也是统计学走向成熟的重要标志。

2. 理论统计学和应用统计学

按统计方法的不同作用,可以将统计学分为理论统计学(Theoretical Statistics)和应用统计学(Applied Statistics)。

理论统计学,是指研究统计学的一般理论和统计方法的数学原理的学科,它的立足点是统计方法的研究。由于现代统计学用到了几乎所有方面的数学知识,特别是概率论和数理统计,所以从事统计理论和方法研究的人员需要有坚实的数学基础。理论统计学是统计方法的理论基础。

应用统计学,是指研究如何应用统计方法去解决实际问题的科学。由于在自然科学及社会科学研究领域中,都需要通过数据分析来解决实际问题,因而,统计方法的应用几乎扩展到了所有的科学研究领域。例如,生物统计学、医疗卫生统计学、农业统计学、经济统计学和社会统计学等等。应当指出,所有这些应用统计学的不同分支,都应用的是描述统计和推断统计的主要方法。

第 3 节 统计学在经济管理中的应用

在经济全球化的今天,衡量成功的管理者和决策者的标准,就是看他是否能够充分理解并有效利用统计信息。因为统计信息的作用越来越重要,它的应用范围已经涉及到会计、金融、生产、营销及经济各个领域。

例如,在会计领域,会计师事务所为了经济、快捷的审核客户资产负债表中的应收帐款是否真实可靠,通常的做法就是抽选一些应收款项作为样本进行审核,然后审计师便可得出认同或拒绝的结论。

在金融行业,掌握了股票市场包括市盈率和股利的大量金融信息后,通过比较个股价格与平均市价,金融顾问就可以评判该个股的价值是被高估还是低估,以便作出买入、抛出或持有的决定。

在生产过程中,产品的质量控制是统计的最重要应用,特别是通过控制图可以实时监测产品的技术指标,及时了解何时有必要调整生产过程,使产品质量在要求的误差范围之内,以确保达到规定标准。

在市场营销领域,采用市场调查的各种手段(如超市自动收银台的电子扫描仪所保留的信息),了解商品销售的有关数据并加以分析,有助于各种商品未来营销策略的制定。

在经济领域及其各个方面,统计的重要应用是利用各种各样的信息进行预测,如经济学家利用价格指数、失业率及生产能力利用率等相关指标,采用适当的数学模型,可以预测通货膨胀率。

第 4 节 统计学与其他学科的关系

一、统计学与数学的关系

统计学与数学有密切的关系,但又有本质的区别。统计方法与数学方法一样,都不能独立地直接研究和探索客观现象的规律,

而是给各学科提供了一种研究和探索客观规律的数量方法。统计学日益依靠数学方法进行严密的定量分析。虽然表面上看统计学与数学都是研究数量规律的,但实际上却有着明显的差别。首先,数学研究的是抽象的数量规律,而统计学则是研究具体的、实际现象的数量规律;数学研究的是没有量纲或单位的抽象的数,而统计学研究的是有具体计量单位的数据。其次,统计学与数学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,而统计学则是演绎与归纳相结合,占主导地位的是归纳。

二、统计学与其他学科的关系

统计学是一门应用型方法论学科,同时由于几乎所有的学科都要研究和分析数据,因而,统计学与几乎所有的学科都有着或多或少的联系。这种联系表现为,统计方法是其他学科借以探索学科内现象的数量规律性的有效手段,而对这种数量规律性的解释并进而研究各学科内在的规律,只能由各学科的深入研究来完成。比如,利用统计方法对吸烟和不吸烟者患肺癌的数据进行分析,得出吸烟是导致肺癌的原因之一,但为什么吸烟能导致肺癌,这就需要医学进行解释了。

由此可见,统计方法是一种有用的定量分析的工具,但它并不是万能的,不能解决你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,首先取决于所选择的统计方法是否正确,其次还要在定性分析的基础上进行定量分析,然后再应用各学科的专业知识对统计分析的结果作出合理的解释,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,使用统计方法的复杂程度也不尽相同,但统计方法在各学科的研究中会发挥越来越重要的作用。