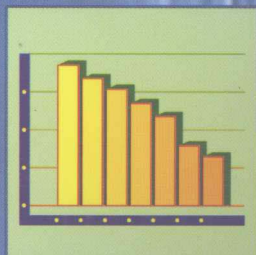
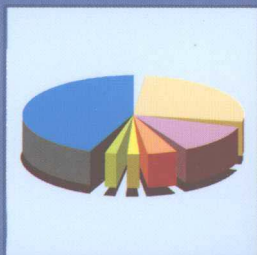
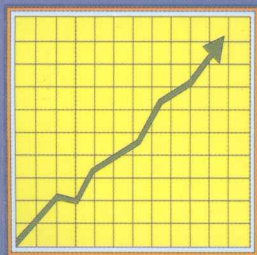


统计学原理

(修订版)

朱洪文 王淑杰 郑桂荣 编著

哈尔滨工业大学出版社



统计学原理

朱洪文 王淑杰 郑桂荣 编著

哈尔滨工业大学出版社

内 容 简 介

本书比较系统地介绍了统计学的基本原理。内容包括常见的一些统计方法,如统计资料的搜集整理方法、综合指标法、统计推断法、时间数列分析、相关回归分析、指数方法等。全书尽量避免繁琐的数学推导,而把重点放在这些统计方法的应用上,具有较强的实用性。

本书理论联系实际,简明易懂,可作为经济管理专业大中专教材,也可供从事社会经济管理工作的人员阅读参考。

图书在版编目(CIP)数据

统计学原理/朱洪文等编著.—3版.—哈尔滨:哈尔滨工业大学出版社,2004.10(2008.6重印)

ISBN 978-7-5603-1284-2

I.统… II.朱… III.统计学 IV.C8

中国版本图书馆 CIP 数据核字(2004)第 089388 号

出版发行 哈尔滨工业大学出版社

社 址 哈尔滨市南岗区复华四道街 10 号 邮编 150006

传 真 0451-86414749

印 刷 黑龙江省地质测绘印制中心印刷厂

开 本 880mm×1168mm 1/32 印张 8.625 字数 219 千字

版 次 2004 年 10 月第 3 版 2008 年 6 月第 5 次印刷

书 号 ISBN 978-7-5603-1284-2

定 价 18.00 元

再版前言

本教材于1998年出版后,曾于2000年修订再版。经过几年的应用,我们觉得有必要对书稿内容做进一步的修改、完善。因此,我们在2000年版基础上,结合近年来各领域对统计学基础知识要求的变化,做了适当的调整和修改。

本教材是根据“统计学”教学大纲的要求,结合高等院校非统计专业学习统计的需要编写而成的。全书力图体现出“大统计”的思想,不囿于社会经济统计学与数理统计学之分,只侧重于统计方法的实际应用,这是本书的一大特色。

全书共分九章。第一章介绍了统计学的基本理论;第二、三章介绍了统计资料的搜集和整理方法;第四章至第九章介绍了统计资料的各种分析方法。这些内容完全是统计学的基本原理和基本方法,因此本书定名为《统计学原理》。全书内容基本满足了经济管理专业本科生与专科生对于统计学的需求。对于其他类读者,可根据实际需要阅读相关内容。

参加本书编写的有哈尔滨工业大学朱洪文(第一、二、三、五、六章)、王淑杰(第四、七、九章)、长春金融商业专科学校郑桂荣(第八章)。最后由朱洪文负责总纂。

由于编写水平所限,虽经修改,书中还会有不妥之处,敬请读者批评指正。

编 者

2004年8月

目 录

第一章 总论

- 第一节 统计学的产生和发展…………… (1)
- 第二节 统计学的研究对象和研究方法…………… (3)
- 第三节 统计学学科的分类…………… (7)

第二章 统计资料的搜集

- 第一节 统计资料的来源…………… (9)
- 第二节 统计调查的组织形式…………… (12)
- 第三节 统计调查方案的设计…………… (17)

第三章 统计资料的整理与显示

- 第一节 统计整理的方法…………… (24)
- 第二节 频数分布…………… (29)
- 第三节 统计表与统计图…………… (34)

第四章 综合指标

- 第一节 总量指标…………… (41)
- 第二节 相对指标…………… (44)
- 第三节 平均指标…………… (52)
- 第四节 标志变异指标…………… (74)

第五章 统计推断基础

- 第一节 概率分布…………… (85)
- 第二节 抽样分布…………… (99)

第六章 统计推断

- 第一节 参数估计…………… (104)
- 第二节 假设检验…………… (117)

第七章 时间数列分析

第一节 时间数列的基本问题	(127)
第二节 时间数列的水平指标	(131)
第三节 时间数列速度分析指标	(144)
第四节 长期趋势分析	(152)
第五节 季节变动分析	(164)
第六节 循环变动和不规则变动	(170)
第七节 时间数列因素分析应用	(172)

第八章 指数分析

第一节 指数的一般概述	(179)
第二节 综合指数	(183)
第三节 平均数指数	(194)
第四节 平均指标指数	(198)
第五节 指数体系与因素分析	(201)
第六节 指数数列	(211)

第九章 回归与相关分析

第一节 回归分析的基本问题	(217)
第二节 简单直线回归分析	(219)
第三节 简单线性回归的相关性检验	(231)
第四节 多元线性回归分析	(244)

参考文献	(249)
------	-------

附表 1. 随机数表	(250)
------------	-------

附表 2. 二项分布表	(251)
-------------	-------

附表 3. 标准正态分布表	(257)
---------------	-------

附表 4. t 分布表	(258)
---------------	-------

附表 5. χ^2 分布表	(260)
--------------------	-------

附表 6. F 分布表	(262)
---------------	-------

第一章 总 论

第一节 统计学的产生和发展

学习统计这门学科,必须首先弄懂“统计”一词在今天的意义以及它产生的历史。

统计,英文为 Statistics,包括三个方面的含义:统计活动、统计资料和统计学。统计活动(或称统计工作)是指各国政府或其他机构为满足政治、经济、社会等方面的需要以及科学研究的需要而进行的收集、整理、分析、编制有关数据的一系列活动。这些活动的程序一般包括统计调查、统计整理、统计描述、统计分析和统计显示等。统计资料是指统计活动产生的原始的或加工、编制过的统计数据材料,是统计活动的成果。统计学是指研究大量现象数量关系变化规律的方法论科学,是一门关于统计资料收集、显示、描述和分析方法的学科。统计学来源于统计活动,是统计活动经验的理论概括,又用理论和方法指导统计活动。以上三者之间具有密切的联系,习惯上把这三者统称为统计。

一切科学都来源于实践。统计也是随着科学的产生和发展以及适应国家管理需要而产生和发展的。

在原始社会里,人类最初的打绳结、堆石子等简单的计数活动,就蕴藏着统计的萌芽。随着奴隶社会国家的产生,计数活动进一步发展。在中国古代、古希腊、古罗马都有人口、财产和土地等方面的记载。到了封建社会末期,特别是进入资本主义社会以后,社会生产力迅速发展,社会分工越来越细,交通、航运、贸易日益发达,国际市场逐步形成,这时的统计已逐步扩展到工业、农业、贸

易、银行、保险、交通、邮电、海关等部门,并且出现了专业的统计机构和研究组织。统计逐步成为社会分工中一个独立部门和专业。

从 17 世纪中叶开始,欧洲出现的一些统计理论著述,标志着统计学的初步创立。从统计学的产生和发展过程来看,可以把统计学大致分为古典统计学、近代统计学和现代统计学三个时期。

1. 古典统计学时期

这是指 17 世纪中叶到 18 世纪中叶这一百年间的统计学萌芽时期。这时期主要代表人物有约翰·格朗特(1620~1674 年)、威廉·配第(1623~1687 年),以及阿亨瓦尔(1719~1772 年)等。他们的代表作《对死亡表的自然观察和政治观察》、《政治算术》和《近代欧洲各国国势学概要》的发表,奠定了统计学作为一门独立科学的基础,使统计学有了自己的名称、理论和方法。他们使用了如大量观察法、分类法、对比法和图表法等统计方法,来分析研究社会经济现象,满足国家管理的需要。与此同时,概率论也有了初步的发展,出现了杰克斯·伯努利(1654~1705 年)的《推测法》等著作。但这时的概率论与统计学还没有多少联系,呈独立发展的态势。

2. 近代统计学时期

这是指 18 世纪中叶到 19 世纪末叶的一百多年的统计学过渡时期。进入近代统计时期,欧洲的社会经济又有了很大发展,出现了保险、卫生、农业、工商业等不断细分化和定量化的统计。这些统计的出现,在客观上要求统计提供更有效的调查整理和分析资料的方法。在这同时,以拉普拉斯(1749~1827 年)为代表的数学家们又进一步发展了概率论,提出了大数定律、正态分布和误差理论等概念和计算方法。使运用概率论等数理方法研究法律、政治、道德、经济等社会问题的人日渐增多。正是在这种条件下,综合了社会经济理论和概率论方法的近代统计学便应运而生了。其主要代表人物是比利时的统计学家阿·凯特勒(1796~1874 年),他的著作有《论人类》、《概率论书简》和《社会物理学》等。他对统计学的贡献是巨大的。他把自然科学的研究精神和研究方法(实验法、归

纳法)带到社会现象的研究中来,并用数理方法研究社会现象,使前期的统计学逐步从实质性科学过渡到方法论科学。从而推动了统计学这一方法论科学的应用。他还发展了大量观察法并为数理统计学奠定了基础,成为数理统计学的奠基人。

3. 现代统计学时期

这是指 20 世纪初迄至今日的统计学发展时期。这一时期数理统计学由于同自然科学、工程技术科学紧密结合,被广泛应用而得到迅速发展,进入了鼎盛时期,并先后建立了推断的理论和方法、统计假设理论,创立了决策理论,发展了样本分布理论,提出了实验设计理论和方法等,进一步拓宽了统计学的范围。在 20 世纪 40 年代以后,数理统计学已经逐步成为现代统计学中的主流,占有重要地位。

但是,我们也应看到,以社会现象为研究对象的社会科学统计学在许多国家依然存在。只是由于两种社会制度和两种意识形态的存在,而产生了分化:一种是属于资本主义思想体系的社会统计学;另一种是属于社会主义思想体系的社会经济统计学。前者,由于片面强调数学化,已逐渐失掉了独立性而成为数理统计学的附庸;后者,在马列主义统计理论指导下,不断总结社会主义国家的统计经验,取得了一定的进展。如在分组理论方面、指数理论方面都有不少建树,被认为是统计学史上又一次质的飞跃。但总的看来,社会科学统计学发展还是比较缓慢的,这不仅是因为社会现象本身更为复杂,而且作为社会科学,它还受到社会政治变故的影响。

第二节 统计学的研究对象和研究方法

一、统计学的研究对象

从上述统计学的发展历史来看,统计学是从研究社会经济现象开始的。经过 300 余年的演变与发展,统计学趋于成熟,成为一

门研究大量社会经济现象和自然现象的数量方面的方法论科学。作为一门方法论科学,统计学的研究对象是变量及其数据。具体地说,它是以客观现象的数量方面为研究对象,研究如何收集、整理、显示、描述、分析和解释客观现象的数量表现、数量特征、数量关系,以期认识数据的规律性。从当前世界各国统计研究的情况看,统计学的特点有以下三个方面。

1. 统计学研究的是客观现象的数量方面

统计学是通过对客观现象的规模、水平、构成速度等数量关系的调查研究,说明客观事物发展规律在一定时间、地点条件下的数量表现,包括数量的多少、现象之间数量关系,以及质量互变的数量界限。数量性是统计研究对象的主要特点,这一特点必须建立在大量观察这一基本原则基础上,只有对大量实际数字资料进行综合研究,才能反映事物的本质和规律性。

2. 统计学研究的是整体现象的数量特征

客观世界是十分复杂的,任何一整体都是由许多个体所组成。这些个体由于受种种偶然因素的影响,其数量特征并不能代表整体的特征。因此,要反映整体的本质和规律性,必须从总的方面观察分析事物的数量特征,而不能着眼于个别事物。凯特勒曾说过,人者分而观之,人心之不同各如其面,几无规律可循,然合人人而观之,则相去不远,其间自有不变的规律在。这说明只有从整体上观察,才能做出正确的判断。当然,要了解整体的特征和规律性,离不开收集个体的数据,但这仅是研究整体的一种手段而已。

3. 统计学所研究的是客观事物的具体的数量方面

任何客观事物都具有质和量的两个方面,并构成客观事物的统一体。因此,对客观现象数量方面进行研究,必须与现象的质密切结合,研究在一定质的规定下具体的数量表现,而不是抽象的数字。这也是统计学与数学的不同之处。统计学的数字都是在一定的时间、地点、条件下的具体数量,而数学则是研究抽象的数量关系和空间几何形式,是舍弃了具体对象的抽象的量。两者有质的

差别。客观事物的质与量是相互转化的,因此,统计学还要注意研究客观现象由量变到质变的数量界限,定性和定量两个方面结合起来,以达到对现象的本质及其规律性的正确认识。

统计研究对象的数量性、整体性和具体性,决定了统计学是既区别于其他实质性科学,又区别于数学的一门独立的方法论科学。

二、统计学的研究方法

统计学根据所研究对象的性质和特点,在长期的统计实践经验的基础上,总结了一系列特有的研究方法,如大量观察法、统计分组法、综合指标法、时间数列理论、相关分析法、归纳推断法、统计预测法等等。这些统计方法构成统计学有机结合的方法体系,成为统计学研究的中心内容。以下重点阐述大量观察法、综合指标法和归纳推断法,其他方法分别在各章节论述。

1. 大量观察法

统计学的研究对象具有数量性、整体性、具体性的特点,决定了统计研究方法必须采用大量观察法。所谓大量观察法,是指对所研究事物的全部或足够数量进行观察的方法。各种客观现象是受多种因素相互影响的结果,组成客观现象整体的个别单位往往受偶然因素的影响,如果只选择一部分单位进行观察,是不能代表整体的一般特征的,必须观察事物的全部或足够数量单位并加以综合分析。这样可使事物中非本质的偶然因素的影响互相抵消或削弱,整个现象的一般特征就能显示出来。例如,在人口统计方面,生男或生女在个别家庭中存在着差异,但经过大量观察后,男婴女婴的出生数则趋向均衡。这就说明,同质的大量现象是有规律的,尽管个别现象受偶然因素的影响,出现差异,但观察数值达到一定数量后就会呈现出规律性来。

大量观察法的实质也是符合概率论中大数定律要求的。概括地说,大数定律就是:研究由大量的相互独立(或关系很微弱)的随机因素所构成的总体时,如果每个因素对总体的影响很小,则这些

个别因素的影响将相互抵消,而使整体特征(如频率、平均值等)呈现出稳定性。大数定律从统计理论上进一步证明了事物的偶然性和必然性的辩证规律,并因此成为统计学重要的基础理论之一。

2. 综合指标法

统计指标是用来说明客观事物的主要形式。综合说明大量现象的数量关系,概括地表明其一般特征和规律性,必须使用综合指标,即通过对大量原始资料的分组、整理、汇总,得出综合指标,从而排除个别的、次要的和偶然因素的影响,显示出普遍的、主要的结果。这种方法就是综合指标法。

常用的综合指标有总量指标、相对指标、平均指标等,例如我国 1992 年工业总产值为 36 802 亿元,1992 年我国国民收入使用额为 19 715 亿元,积累额占国民收入使用额的 34.3%,这些都是综合指标,它们概括地描述了我国 1992 年工业和国民收入的综合数量特征。

综合指标法是统计分析的基本方法,其他统计分析方法都以它作为基础。时间数列法、指数法、抽样法等均离不开综合指标的对比分析。

3. 归纳推断法

根据调查中各单位的特征,得出全部整体的某种信息,这种由个别到一般的推理方法,就是统计研究常用的方法——归纳推断法。综合指标虽然能够概括反映整体的一般数量特征,但它必须是从各个单位数量特征归纳计算而得。当我们所观察的只是部分或有限的单位,而所需要判断的客观现象范围却是大量的甚至是无限的,那么就必须根据部分单位的数量特征,归纳推断出客观现象全部范围的数量特征。例如,一个医学研究工作者要确定一种新药是否比旧药疗效更好,在医院里把 100 个病人随机划分为两组,一组用新药,另一组用旧药。从开始用药起,就从每个病人那里取得各种医学资料,经过一段时间的观察记录,对两组资料进行分析和比较,并以一定的置信标准来推断这两种药的疗效优劣。

这种根据部分数据来判断整体数量特征的方法叫做统计推断法。

统计推断是现代统计学的基本方法,它广泛应用于农业产量调查、工业产品质量检查与控制,以及根据时间数列进行预测所做的参数估计和假设检验等。

第三节 统计学学科的分类

一、统计学学科体系

统计学是一门多科性的学科,在长期的发展过程中形成了许多分支学科和交叉学科。根据研究问题的方法不同,对统计学的主要内容进行划分,统计学学科体系可以分为以下两个部分:(1)社会经济统计学。包括社会统计学、经济统计学、科技统计学、环境与生态统计学等。其中社会统计学还包括人口统计学、教育统计学、司法统计学、社会保障统计学等;经济统计学包括国民经济统计学、工业统计学、农业统计学等部门统计学。(2)数理统计学。它可以分为理论数理统计学和应用数理统计学。前者包括概率论、抽样理论、实验设计、估计理论、假设检验、决策理论、非参数统计、序贯分析、多元分析、时间数列,以及博弈理论等;后者包括生物计量学、统计力学、遗传统计学、工程统计学、天文统计学、政府统计学、品质管理学、经济计量学等。统计学学科体系除以上主要内容,还包括统计史、统计法学、国际统计学等。

二、应用统计学

在现代自然科学和社会科学相互影响和渗透,边缘科学(交叉科学)不断涌现的发展趋势中,统计学作为一门方法论科学,被广泛地应用于解决自然科学和社会科学的各个领域中的问题,由此形成了应用于解决各学科问题的统计学,即应用统计学。它包括一般性的统计方法,适用于各个领域,如参数估计、假设检验、方差

分析、相关与回归分析等方法,还包括特有的分析方法,只适合于某一专业领域,如指数分析法、时间数列分析法等。应用统计学区别于理论统计学的不同之处在于:它不对统计数学原理进行推导,而是侧重于阐明统计的思想,并将理论统计学的结论作为工具应用于各个领域。诸如在气象观察、药理试验、市场调查、质量控制、财务管理和投资决策等研究中,都应用统计方法作为分析问题和解决问题的工具。

应用统计学与一般统计学的不同之处有三点:

第一,从研究对象来看,它是以社会经济现象或自然现象中具体的指标数值作为主要研究对象,这些指标都具有质的含义。

第二,从研究方法来看,它是研究如何运用统计学的理论和方法来研究现实问题。

第三,从研究的目的来看,它以了解掌握社会经济现象和自然现象和数量特征、数量关系和统计规律为目的。

第二章 统计资料的搜集

第一节 统计资料的来源

一、统计总体的概念

当我们确定某一调查任务后,首先必须明确调查的范围,也即资料的来源范围。搞清资料的来源范围,既可以避免搜集范围以外的无用资料,还能使调查者对资料范围的特征认识清楚,防止范围以内的资料遗漏。这个资料来源的范围,在统计学上通常称为统计总体。

所谓统计总体就是根据一定的目的和要求所确定的研究事物的全体,它是由客观存在的具有某种共同性质的许多个别事物构成的整体,简称为总体。例如,当我们需要研究某个地区工业企业生产情况时,该地区的所有工业企业就成为我们研究的一个统计总体。它们都具有本地区及工业企业的特点,这是它们组成一个总体所具有的共同性质。而其他地区的工业企业或本地区的非工业企业都不构成这次研究的总体。确定了这个总体也就确定了资料的来源范围,所以我们搜集的资料只能是本地区的工业企业的资料,而不是其他。

范围确定以后,就需要确定具体资料由谁来承担的问题。在统计研究中,把统计资料的原始承担者称为总体单位。总体单位是指那些组成总体的个别单位,统计研究所需要的资料都是来自于各个总体单位。在上例中,本地区所有工业企业是一个总体,而

每一家工业企业就是总体单位,所需要搜集的资料全部来源于每一家工业企业,它们是原始资料的承担者。

在研究客观事物时,我们总是希望获得有关总体的全部资料,即把所有的总体单位的资料全部搜集到。但是这在实践中往往是不可能的。统计总体可能是包含无限个总体单位的无限总体,如浩瀚宇宙中的无数个星体,它们是不可能全部被调查的;统计总体也可能是包含有限个总体单位的有限总体。但是即使是有限总体,也会由于各种情况的限制,使我们很难对全部总体单位进行调查,这时我们只能从总体中抽取一部分总体单位进行调查,据此来推断总体。这些被抽取的部分总体单位所组成的整体,称为样本。在社会经济的各种调查中,很多时候都是根据样本的特征来推断总体特征的。因此,根据实际情况,在确定调查资料的范围时,还需要确定能够代表总体特征的样本的范围。

总体和总体单位是相对的,它们是随着研究的目的不同,而随之改变的。例如,当需要了解某市的工业企业生产情况时,每一家企业都是这次研究的总体单位;而当我们研究其中某一家企业的职工情况时,则该企业的每一位职工就是总体单位,而该企业就是我们研究问题的总体。

二、搜集统计资料的途径

搜集统计资料应本着“准确、及时、完整和方便”的原则进行,脱离了这些特点的统计资料则毫无意义。本文着重介绍以下两种途径。

1. 来源于公开的出版物

世界各国或地区都有定期的统计出版物,如《中国统计年鉴》、《中国统计摘要》、《北京社会经济年鉴》,以及《美国统计摘要》、《日本统计月报》、《联合国统计年鉴》等,这些出版物能提供较系统的大量数据,是统计资料的一个主要来源;另外,一些部门或专门组织根据管理的需要不定期地出版有关行业内部或特定研究的统计

报告,这些报告能提供比较详细深入的统计资料,这也成为统计资料的一个来源。

2. 统计调查与实验

为了研究某一问题,如果没有现存的统计资料可以引用时,则必须自己搜集资料。搜集资料的方法主要有两种,即统计调查和实验方法。

实验方法就是控制一种或多种因素保持不变,只考查某种所研究因素的变化情况。这种取得资料的形式,称为实验。例如,企业管理者想知道在职技术培训能否提高劳动效率。那就可以将工人随机分为两组。一组人员进行培训,另一组人员不进行培训,然后比较两组工人的劳动效率的高低。但是由于在实验方法中对非研究因素进行控制的要求常常难以实现,尤其在社会经济方面进行实验更难以控制,因此,实验方法主要用于自然科学的资料取得上。

对于社会经济现象,因为各种因素变动复杂,一般采取统计调查的方式来搜集资料。所谓统计调查就是根据研究的目的,对客观存在的事实进行登记,以取得实际统计资料的工作过程。统计调查与实验方法的区别在于,它是以实际发生的结果为依据,然后根据各种因素的变化情况来确定对结果所发生的影响。如果单从准确角度看,统计调查方法不如实验的方法。但是由于社会各种因素不可控性的存在,在对社会经济现象的研究时,一般仍然应用统计调查的方法取得资料。

三、统计资料的审核

在搜集统计资料时,必须同时进行资料的审核工作。所谓审核,就是对原始资料进行审查与核实,目的在于保证资料的客观性、准确性和完整性。计算机人员有句话:进垃圾出垃圾。因此,如果原始资料有错误,必然导致结论的错误。轻者会使统计分析的结果毫无意义,重者可能导致错误的决策,引发严重的后果。

审核统计资料应根据不同的资料来源,选择不同的审核重点