

经济管理专业课程教材

应用统计学

◆徐建中 朱建新 主编

哈尔滨工程大学出版社
HARBIN ENGINEERING UNIVERSITY PRESS

工商管理核心课程教材

应用统计学

徐建中 朱建新 主编

哈尔滨工程大学出版社

图书在版编目(CIP)数据

应用统计学/徐建中主编. —哈尔滨:哈尔滨工程大学出版社, 2004

ISBN 7-81073-238-2

I. 应… II. 徐… III. 应用统计学 IV. C8

中国版本图书馆 CIP 数据核字(2004)第 011942 号

内 容 简 介

本书结合了社会经济统计学和数理统计学中有关理论的研究成果,详细介绍了在宏观、微观经济管理活动中对有关数据的搜集、整理、分析及在此基础上进行预测等工作的常用方法,并对统计工作的组织和管理以及统计结果的表达方式等问题进行了详细的阐述。

本书可以作为高等院校经济或管理类专业的研究生教材,也可供统计专业的本科学生、机关及企业管理干部、财会统计人员及经济研究人员参考。

哈尔滨工程大学出版社出版发行
哈尔滨市南通大街145号 哈工程大学11号楼
发行部电话:(0451)82519328 邮编:150001
新华书店经销
黑龙江教育厅印刷厂印刷

*

开本 787mm×960mm 1/16 印张 16.75 字数 302 千字

2004 年 2 月修订第 1 版 2004 年 2 月第 1 次印刷

印数:1 001—3 000 册

定价:17.00 元

前 言

在人类迈进 21 世纪的今天,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愈来愈广泛的应用于日常生活、研究和工作,这无疑对统计学的教学提出了更高的要求。统计学是高等院校经济管理专业的一门专业基础课程,因此编好这一教材就成了一项很有意义的工作。为了满足教学的要求,我们吸取了同类教材的特色之处,编写了这本书。

本书在第 1 次印刷的基础上,根据教学实际情况,对书中的某些章节及内容做了适当修改。

本书以现代社会经济统计学和数理统计学的有关理论为基础,对在日常生活、研究和工作常用的统计方法进行了系统的介绍和分析;注重理论和实践相结合、宏观和微观相结合、定性和定量相结合,在重视结构合理、理论具有一定深度的同时,在内容体系方面进行了一些新的探索和尝试;突出了实用性强,内容新颖、翔实的特点,并力求做到理论联系实际,更好地反映中国高等教育的特色。

本书各章执笔人是:徐建中(第 1、2、3、4、5、6 章)、朱建新(第 7、8、9、10、11 章)。由徐建中负责全书大纲的设计、书稿的组织和编纂。本书在编写过程中,参考了有关教材和研究成果,在此表示感谢!由于编者水平有限,加之时间仓促,书中难免存在缺点和疏漏,希望广大读者批评指正。

编 者

2004 年 3 月

目 录

第1章 总论	1
1.1 统计学的研究对象	1
1.2 统计学的类别	6
1.3 统计研究工作的程序和方法	11
1.4 统计学中常用的几个概念	13
第2章 统计数据的搜集和显示	17
2.1 统计数据的搜集	17
2.2 统计表和统计图	23
2.3 经济统计科学研究的特殊图示方法	34
第3章 次数分布	38
3.1 统计分组	38
3.2 次数分配的图示方法	47
3.3 茎叶法	55
第4章 集中趋势和离中趋势	60
4.1 集中趋势的测度	60
4.2 离中趋势的测度	79
4.3 偏斜度和峰度的计量	88
4.4 描述次数分配位置离差的中位数体系	91
第5章 时间序列分析	100
5.1 时间序列构成	100
5.2 长期趋势测定方法	107
5.3 季节变动分析	119
5.4 循环波动分析	126
5.5 时间序列综合分析的程序和应用	129
第6章 指数	133
6.1 指数的概念及作用	133
6.2 指数的编制方法	134
6.3 指数的应用	142
6.4 指数体系与因素分析	149

第 7 章 随机变量及概率分布	155
7.1 随机变量	155
7.2 离散型随机变量的概率分布	156
7.3 正态分布	161
7.4 抽样分布	166
第 8 章 参数估计	172
8.1 参数估计概述	172
8.2 总体均值的区间估计	177
8.3 两个总体均值之差的估计	179
8.4 总体比例和两个总体比例之差的区间估计	183
8.5 正态总体方差及两个正态总体方差比的区间估计	185
第 9 章 假设检验	188
9.1 假设检验的基本问题	188
9.2 总体均值的假设检验	192
9.3 两个总体均值的假设检验	197
9.4 总体比例的假设检验	200
9.5 总体方差的假设检验	203
第 10 章 抽样调查	205
10.1 抽样调查的一般问题	205
10.2 抽样误差和抽样估计	207
10.3 抽样调查的组织方式	211
10.4 样本容量的确定和对总量指标的推算	219
第 11 章 相关与回归	223
11.1 相关分析	223
11.2 一元线性回归	226
11.3 多元线性回归	237
附表 1 二项概率分布的部分数值表	245
附表 2 正态分布的密度函数表	246
附表 3 正态分布的分布函数表	247
附表 4 随机数表	248
附表 5 t 分布表	249
附表 6 χ^2 分布表	250
附表 7 F 分布表	251
参考文献	260

第 1 章 总 论

1.1 统计学的研究对象

1.1.1 统计学的产生和发展

人类的统计实践是随着计数活动而产生的。因此,对统计发展的历史可追溯到远古的原始社会,也就是说,距今已有 3000 多年的漫长岁月。但是能使人类的统计实践上升到理论上予以概括总结的程度,即成为一门系统的学科——统计学,却是近代的事情,距今只有 300 多年的短暂历史。回顾一下统计科学的渊源及其发展过程,对于我们了解统计学的研究对象和性质,学习统计学的理论与方法,以及提高我们的统计实践和理论水平,都是十分必要的。

最早的统计是作为国家重要事项的记录,比如,从公元前 21 世纪夏禹立国开始,中国的统计不仅详细记录重大历史活动成果,而且明显地被新兴的奴隶制国家用作治国的手段;又如,在古代埃及、希腊和罗马的历史中,也有关于国情国力的记载。统计实践经过漫长的历史长河,尤其是经过封建社会末期经济的发展,客观上需要从理论上加以概括和总结。从统计学的产生和发展过程来看,可以把统计学划分为三个时期:统计学的萌芽期、统计学的近代期和统计学的现代期。

(1) 统计学的萌芽期

统计学的萌芽期始于 17 世纪中叶至 18 世纪中、末叶,当时主要有国势学派和政治算术学派。

① 国势学派

统计学的萌芽最初在当时欧洲经济发展较快的意大利孕育良久,但最终却在 17 世纪的德国首先破土成芽。这个时期的代表人物是康令(H. Conring, 1606—1681 年)、阿亨瓦尔(G. Achenwall, 1719—1772 年)和斯廖采尔(1735—1800 年)等。他们在大学中开设了一门新课程,最初叫做“国势学”。我们把从事这方面研究的德国学者称为国势学派。他们所作的工作主要是对国家重要事项的记录,因此又被称为记述学派。这些记录记载着关于国家组织、人口、军队、领土、居民职业以及资源财产等事项,几乎完全偏重于品质的解释,而忽视了量的分析。严格地说,这一学派的研究对象和研究方法都不符合统计学的要求,只是登记了一些记述性材料,用以说明

管理国家的方法。

当然,国势学派对统计学的创立和发展还是作了不少贡献的:首先,国势学派为统计学这门新兴的学科起了一个至今仍为世界公认的名词“统计学(Statistics)”,并提出了至今仍为统计学者所采用的一些术语如“统计数字资料”、“数字对比”等。国势学派建立的最重要的概念就是“显著事项”,它事实上是建立统计指标和使统计对象数量化的重要前提;其次,国势学派在研究各国的显著事项时,主要是系统地运用对比的方法来研究各国实力强弱,统计图表实际上也是“对比”思想形象化的产物。

② 政治算术学派

统计理论在英国与德国几乎同时产生,由于两国的社会背景、经济水平和思想渊源的不同,其统计理论亦各具特色。在英国当时从事统计研究的人被称为政治算术学派。虽然政治算术学派与国势学派研究都与各国的国情、国力这一内容有关,但国势学派主要采用文字记述的方法,而政治算术学派则采用数量分析的方法,因此,从严格意义上来说政治算术学派作为统计学的开端更为合适。17世纪的英国学者威廉·配第(W. Petty, 1623—1687年)在他所著的《政治算术》一书中,对当时的英国、荷兰、法国之间的“国富和力量”进行数量上的计算和比较,做了前人没有做过的从数量方面来研究社会经济现象的工作。正是在这个意义上,马克思称配第是“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”(《资本论》第1卷,第302页)。

配第的朋友约翰·格朗特(J. Graunt, 1620—1674年),当时也从事了50多年来伦敦市人口出生和死亡的计算工作,写出了第一本关于人口统计的著作。从此,统计的含义从记述转变为专指“量”的方面来说明国家的重要事项。这就为统计学作为一种从数量方面认识事物的科学方法开辟了广阔的发展前景。

政治算术学派在统计发展史上有着重要的地位。首先,它并不满足于社会经济现象的数量登记、列表、汇总、记述等传统的统计过程,而是进而要求把这些统计经验加以全面系统地总结,并从中提炼出某些理论原则。这个学派在搜集资料方面,较明确地提出了大量观察法、典型调查、定期调查等思想,在处理资料方面,较为广泛地运用了分类、制表及各种指标来浓缩与显现数量资料的内涵信息。其次,政治算术学派第一次有意识地运用可度量的方法,力求把自己的论证建筑在具体的、有说服力的数字上面,依靠数量的观察来解释与说明社会经济生活。然而,政治算术学派在很大程度上毕竟还处于统计核算的初创阶段,它只能从简单的、粗略的算术方法来对社会经济现象进行计量和比较。

(2) 统计学的近代期

统计学的近代期始于18世纪末至19世纪末,在这一时期的统计学主要有数理

统计学派和社会统计学派。

① 数理统计学派

数理统计学派最初的统计方法是随着社会政治和经济的需要而逐步得到发展的,直到概率论被引进之后,才逐渐形成一门成熟的科学。在统计发展史上,最初卓有成效地把古典概率论引进统计学领域的,是法国的天文学家、数学家、统计学家拉普拉斯(P.S.Laplace, 1749—1827年)。他发展了对概率论的研究,推广概率论在统计中的运用,明确了统计学的大数法则,并进行了大样本推断的尝试。

随着资本主义经济的发展,逐步形成近代文化高度发达的一个突出标志,就是统计事业走向昌盛,正是由于比利时统计学家、数学家、天文学家凯特勒(A. Quetelet, 1796—1874年)的努力,才完成了统计学和概率论的结合。从此,统计学才开始进入更为丰富发展的新阶段。国际统计学界有人称凯特勒为“统计学之父”,就在于他发现了大量现象的统计规律性和创造了大量的统计方法。凯特勒把统计学发展中的三个主要源泉,即德国的国势学派、英国的政治算术学派和意大利、法国的古典概率学派加以统一、改造并融合成具有近代意义的统计学,促使统计学向新的境界发展。可以说,凯特勒是古典统计学的完成者,又是近代统计学的先驱者,在统计发展史上具有承上启下、继往开来的地位。

同时,凯特勒也是数理统计学派的奠基人,因为数理统计就是在概率论的基础上发展起来的。随着统计学的发展,对概率论的运用逐步增加,同时自然科学的迅速发展和技术不断进步更对数理统计的方法有进一步的要求。这样,数理统计学就从统计学中分离出来自成一派。它从19世纪末以来逐步形成,由于它主要由英美等国发展起来,故又称英美数理统计学派。

② 社会统计学派

自凯特勒后,统计学的发展开始变得丰富而复杂起来,由于在社会领域和自然领域统计学被运用的对象不同,统计学的发展,呈现出不同的方向和特色,19世纪后半叶,正当致力于自然领域研究的英美数理统计学派刚开始发展的时候,在德国竟异军突起,兴起了与之截然不同的社会统计学派。这个学派是近代各种统计学派中比较独特的一派。由于它在理论上比政治算术学派更加完善,在时间上比数理统计学派提前成熟,因此它很快占领了“市场”,对整个世界统计学界影响较大,流传较广,直至今日。

社会统计学派由德国大学教授尼斯(K.G.A. Knies, 1821—1898年)首创,主要代表人物为恩格尔(C.L.E. Engel, 1821—1896年)和梅尔(G.V. Mayr, 1841—1925年)。他们认为,统计学的研究对象是社会现象,目前在于明确社会现象内部的联系和相互关系;统计方法应当包括社会统计调查中资料的搜集、整理,以及对其分析研究。他们认为,在社会统计中,全面调查,包括人口普查和工农业调查,居于重要

地位。以概率论为根据的抽样调查,在一定的范围内具有实际意义和作用。

(3) 统计学的现代期

统计学的现代期自 20 世纪初到现在的数理统计时期。自 20 世纪 20 年代以来,数理统计学发展的主流从描述统计学转向推断统计学。如 19 世纪和 20 世纪初的统计学教科书中主要讲描述统计学中的一些基本概念、资料的搜集、资料的整理、资料的图示和资料的分析等,后来逐步增加概率论和推断统计的内容,直到 20 世纪 30 年代,推断统计学才使数理统计进入现代范畴。

现在,数理统计学的丰富程度完全可以独立成为一门宏大的学科,但它还不能完全代替一般统计方法论。传统的统计方法虽然比较简单,但在实际统计工作中运用的频率仍然很大,正如四则运算与高等数学的关系一样。不仅如此,数理统计学主要涉及资料的分析和推断方面,而统计学还包括各种统计调查、统计工作制度和核算体系的方法理论、统计学与各专业相结合的一般方法理论等。由于统计学比数理统计在内容上更为广泛,因此,数理统计学相对于统计学来说不是一门并列的学科,而是统计学的重要组成部分。

从世界范围看,自 20 世纪 60 年代以后,统计学的发展有三个明显的趋势:第一,随着数学的发展,统计学依赖和吸收数学营养的程度越来越迅速;第二,向其他学科领域渗透,或者说,以统计学为基础的边缘学科不断形成;第三,随着应用的日益广泛和深入,特别是借助电子计算机后,统计学所发挥的功效日益增强。

统计发展史说明,统计学是从设置指标研究社会经济现象的数量开始的。随着社会的发展与实践的需要,统计学家对统计方法的不断丰富和完善,统计学也就随着不断发展和演变。从当前世界各国统计研究状况来看,统计学已不仅为研究社会经济现象的数量方面,也为研究自然技术现象的数量方面提供各种统计方法;它即研究事物随机现象的数量方面,又为研究事物随机现象的数量方面提供各种统计方法。从统计学的发展趋势分析,它的作用与功能已从描述事物现状、反映事物规律,向进行样本推断,预测未来变化方向发展。

1.1.2 统计学的研究对象

(1) 统计的涵义

统计的起源很早,已经有几千年的发展历史了,它是适应国家行政管理和社会经济发展的需要而产生和发展起来的。最早作为学名使用的“统计”,是由 18 世纪德国政治学教授阿亨瓦尔(G. Achenwall)提出的,他认为统计学是关于国家应注意事项的学问。从此以后,各国相继沿用“统计”这个词。而随着统计活动的发展,数字资料的大量增多,“统计”一词,也就成了记述国家和社会状况的数量关系的总称。

目前,统计一词应用广泛,一般具有三种涵义,即统计工作、统计资料和统计科

学。

统计工作,是指由统计机构对社会经济现象的数量方面进行有计划、有步骤地调查整理和分析研究的活动。

统计资料,是指反映客观实际情况的数字资料、信息或其它有关资料。

统计科学,是指研究如何搜集、整理、分析与预测社会经济现象数量方面资料的科学。统计科学有两大类,一类是社会经济统计学,包括统计学原理、部门或专业统计学和统计史。统计学原理论述统计的基本理论和方法,部门或专业统计学是统计学原理在国民经济各个领域中的具体应用,它又可分为工业统计学、农业统计学和商业、基本建设、物资、劳动、交通运输、财政金融、文教卫生以及国民经济综合平衡等统计学。另一类是数理统计,它是应用数学的一个分支。数理统计学虽然是属于自然科学,但它的许多方法已为社会经济统计学所采用。

统计的这三种涵义,具有密切的联系,统计的活动过程是对社会经济现象进行调查,占有实际资料,取得感性认识,然后再加以综合分析,做出结论,从而上升到理性的认识。所以,统计工作是对社会进行调查研究的过程,而统计资料则是统计工作的成果。统计科学和统计工作的关系是理论和实践的关系,统计科学是统计实践经验的理论概括,反过来,统计科学所论述的理论和方法又指导着统计工作的开展。

(2) 统计学的研究对象

统计学是一门社会科学,是在质与量的辩证统一中研究大量社会经济现象的规模、水平、结构、速度、比例关系和利用程度等数量方面,研究社会经济现象的发展规律在具体地点、时间和条件下的数量表现。

统计学主要研究以下几个方面的内容:

第一,统计学是研究社会经济现象的

①研究地理环境、物质资源和一定历史条件下的人口情况。如人口数目、民族、职业、性别、年龄等构成情况以及计划生育、人口密度、人口变动和人口的文化教育水平等等。

②研究工业、农业、交通运输、基本建设和国内外贸易以及国民收入等情况。即研究生产过程各个要素的数量、构成及其利用情况;还研究社会扩大再生产过程,即研究社会产品的生产、分配、交换与消费、生产力水平、生产中新技术的推广等等。

③研究社会文化服务事业的工作情况。例如,医疗机构的数量及构成;各类学校的数目,其中教师、学生数目及构成;图书馆、影剧院及书报、刊物、影片的出版发行数目等等。

④研究有关国家政治生活。例如,某市人民代表大会人数及其构成;工会会员

及其他社会团体成员的数目及其构成等;另外还有司法行政统计等等。

总之,统计学的研究范围极其广泛,包括政治、经济、文化生活、科学技术、艺术、卫生等方面,既包括生产力也包括生产关系,既包括经济基础也包括上层建筑。

从上述的情况来看,统计学所研究的对象是社会经济现象而不是自然现象。

第二,统计学是研究社会经济现象的数量方面

统计学所研究的是社会经济现象的数量方面,但它决不是离开社会经济现象的质量方面而孤立地去研究社会经济现象的数量方面的。因为质与量是对立的统一,任何质都是具有一定量的质,任何量都是具有一定质的量,二者是不可分割的。因此,统计在从数量方面来研究社会经济现象时,必须坚持质与量的辩证统一。

统计反映和研究社会经济现象的数量方面,就是要反映和研究社会经济现象在一定时间、地点和条件下的规模、水平、结构、普遍程度、速度、比例关系及其发展变化规律。例如,1980年我国工农业总产值为7100亿元,2000年达到28400亿元,也就是为1980年的四倍。这些都是社会经济现象的数量方面。但它们不是纯粹的数字,而是社会经济现象发展实际过程在数量上的表现。

第三,统计学是研究大量社会经济现象的数量方面

统计区别一般调查研究的另一个特点在于,它所研究的主要是大量社会经济现象的数量方面。社会经济现象的发展变化,是受客观经济规律支配的,但个别单位、个别事物在一定时间内的发展变化,也会受偶然因素的影响。只有经过大量的观察,综合的分析,才能排出偶然因素的影响,正确反映社会经济现象发展规律的数量表现。例如,社会主义工业劳动生产率的不断提高是社会主义经济发展的客观规律,但个别企业,个别单位由于某些因素的影响,在一定时间内,可能出现劳动生产率停滞或下降的现象,但经过反复大量的观察,就能揭示出劳动生产率不断提高的规律性。

当然,社会主义统计学还必须注意对个别现象,特别是典型现象的研究。就人们认识事物的规律来说,往往是由个别到一般,就事物的表现来说,甚至有时个别典型现象所表现出来的数量特征,比大量现象所表现出来的共同特征更为丰富、生动和深刻。因此,统计研究要把研究大量现象和典型事物结合起来进行。

1.2 统计学的类别

1.2.1 描述统计学与推断统计学

随着统计学的不断发展,统计学的内容越来越丰富。但就其基本内容来说,分为描述统计和推断统计两大类。

(1) 描述统计

描述统计就是指如何从已知的观察资料,搜集、整理、分析、研究并提供统计资料的理论和方法,用以说明研究现象的情况和特征。描述统计包括各种数据处理,这些数据的处理是用来总括或描述数据的重要特征的,而不必深入一层地去试图推论数据本身以外的任何事情。因此,描述统计的主要作用是通过对现象进行调查或观察,然后将所得到的大量数据加以整理、简缩、制成统计图表。并就这些数据的分布特征(如集中趋势、离中趋势等等)计算出一些概括性的数字(如平均数、标准差、相关系数等等)。

借助于这些概括性的数字,就可以使我们从杂乱无章的资料中取得有意义的信息,便于对不同的总体进行比较,从而得出结论。总之,这些工作的主要目的之一就是使反映客观现象的统计数据可以一目了然,条理清晰,使用方便。

例如,一个随机样本包括 200 名顾客,他们同意对某一新产品提出意见。表 1-1 给出了他们回答的汇总,图 1-1 给出了这一信息的简单图示,即圆形图。

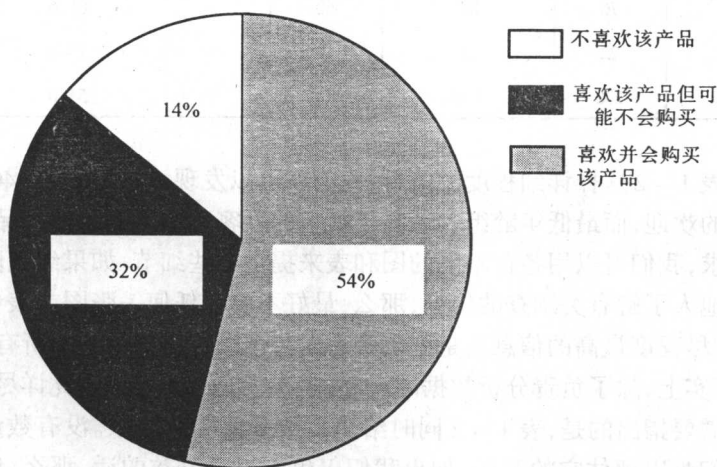


图 1-1 200 名顾客对该产品意见的圆形图

表 1-1 200 名顾客对新产品的意见

意见	百分比(%)
十分喜欢并购买该产品	54
喜欢该产品但可能不会购买	32
不喜欢该产品	14
合计	100

表 1 - 1 和图 1 - 1 是相当粗略的汇总,即描述了顾客对该产品的总的欢迎程度。表 1 - 1 和图 1 - 1 构成了进一步分析的良好起点,它们清楚地表明该产品是相当流行的。然而,有些问题要求更为具体详细的信息。对表 1 - 1 中第一个问题来说,我们也许希望更具体地了解这次被调查者人数的情况。这一更具体的信息如表 1 - 2 所示。

表 1 - 2 被访的 200 名顾客按年龄分组表

年龄	样本年龄组中的人数		十分喜欢并购买该产品的人数	
	人数	比重(%)	人数	占各年龄组人数的比重(%)
15 - 19	10	5	1	10.0
20 - 29	20	10	4	20.0
30 - 39	72	36	38	52.7
40 - 49	76	38	62	81.6
50 及以上	22	11	3	13.6
合计	200	100	108	54.0

有了表 1 - 2 这样详细程度的信息,我们就可以发现,该产品最受 40 ~ 49 岁年龄组顾客的欢迎,而最低年龄组和最高年龄组的顾客不特别喜欢该产品。根据详细程度的要求,我们可以用各自不同的图和表来提供某些细节。如果编制图或表的重点是让其他人了解有关调查的信息,那么,最好不要在任何一张图或表中包括太多的信息,详尽程度最高的信息就是列出带有对每个顾客情况说明的所有 200 名顾客的意见。事实上,除了负责分析数据的人外,多半没有必要知道如此详尽的结果。

尤其需要指出的是,表 1 - 2 同时给出了数量值和百分数。没有数量值的百分数会使人们难以评估它的意义。如果我们仅仅给出百分数的话,那么,应该同时给出被调查的总人数。由小的数量值计算出来的百分数很容易误解。在一个仅有 2 个人的部门中,当其中 1 人被分配去执行一项新的研究项目时,当然可以说该部门有 50% 的人致力于该项目。但它所产生的效果,则与 10 人中投入 5 人或 400 人中投入 200 人(50%)就不同了。

上述例子说明了描述统计的基本思想,即对某项研究中所搜集的数量信息进行整理和表述。

(2) 推断统计

推断统计则是指只凭样本资料以推断总体特征的技术和方法。推断统计可以

利用样本资料来代替总体资料,在观察资料的基础上深入一步地分析、研究和推断,以推知资料本身以外的情况和数量关系,从而对不肯定的事物做出决断,为进行决策提供数据依据。由于推断统计节省时间、人力和物力,因而备受人们的重视和欢迎。例如,在管理现象日益复杂、市场情况瞬息万变的环境中,有许多事情要求对不肯定事物做出科学的决断,因而就要求必须在不完全观察资料的基础上,对所关心的数量关系做出可靠的估计,以便进行有效的决策。

推断统计主要有两种类型,即估计和检验。在这两种类型中,有关总体中某个样本的信息已经取得,所要做的推断是对整个总体的某个数值做出估计,这样的问题属于估计这一类型。例如,推断总体平均数或总体比率(如在一大批产品中合格品所占的百分比)等。如果所作的推断是在几个可供选择的行动方案中进行选择,这样的问题属于检验这一类型。例如,在工业生产管理中可以用于检验两种不同的工艺方法是否优于原有的工艺方法。又如,在药品生产试用阶段,就可以通过大量观察检验新药与旧药之间是否存在显著性的差异,从而判定新药较旧药的疗效是否更好、更安全,并确定其临床推广的应用价值。描述统计是推断统计的前提,而推断统计是描述统计的发展。

1.2.2 理论统计学与应用统计学

理论统计学是指统计学的数学原理,它扎根于纯数学的一个领域——概率论。从广义来说统计理论可以认为是包括概率论的,此外还包括一些并不属于传统概率论的内容,如随机化原则的理论,各种估计的原理,假设检验的原理以及一般决策的原理,这些原理可以看成是概率论公理的扩增。在统计实践中常常会遇到一些新问题,使原有的统计方法不适应,需要统计学家针对新问题去建立一个与实际情况相适合的统计模型,创造新的统计方法去分析,这就要靠统计理论来指导。

与理论统计学相对应的是应用统计学。前面已经提到统计学是应用性很强的一门学科,因此将统计学的基本原理应用于各个领域就形成各种各样的应用统计学。它包括一整套统计分析方法,有的是一般性的统计方法,它适用于各个领域,如参数估计、假设检验、方差分析、相关与回归等。有的则是某一专业领域中特有的分析方法,例如经济统计学中的指数分析法等。近几十年来,由于统计研究的范围越来越广,一些科学实验也日趋复杂,统计方法也相应地复杂化和专门化,在应用统计方法中必须对因模型和实际情况的不一致而引起的各种误差的性质和大小作出判断,或提出改进的措施。由于统计的方法更加专门化了,它们就缺少通用性,一个统计学家要熟悉所有的专门方法已不可能。适应这种发展的需要,要求既要熟悉统计知识又要熟悉某一领域业务的应用统计学家就应运而生,同时也产生了相应的应用统计学。这类统计学的特点不着重于统计数学原理的推导,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具体领域。本书主要是属于应用统计学,特别是侧重于经济管理方面的应用。

1.2.3 统计学与其他学科的关系

(1) 统计学和数学的关系

统计学是研究客观现象数量方面的,它是应用数学的一个分支,因此与数学的关系十分密切,且与其他的应用数学有一定的共性。如和数学中的有关定理一样,统计中的一些分布也是客观现象数量方面的一种抽象。统计学中也要使用很多数学方法,学好统计学,尤其是理论统计学需要有坚实的数学基础。但另一方面,统计学与其他的数学分支相比又有其特殊性。

① 统计方法处理的数据必须是受到偶然性的影响而产生差异的数据。偶然现象在统计学中常称为随机现象,因此统计学及其理论基础概率论不同于其他数学分支的一个特点在于它是研究随机现象的一门学科。

② 在方法上数学常常是用演绎的方法,即在作结论时,是从一些假设命题、已知的事实出发,按一定的逻辑推理去得到有关的结论。而统计学在本质上是用归纳的方法,它是根据观察到的大量个别情况,“归纳”起来去推断总体的情况,这一点与概率论的方法也有区别。因此目前国际上也有—种趋势,把统计学看成为与数学独立的学科。

(2) 统计学与其他专门学科的关系

统计方法有很广泛的实用性,其一般的数据分析方法适用于其他任何科学中的偶然现象,因此它与很多专门学科都有关系。但是统计方法只是从事物的外在数量表现去推断该事物可能的规律性,它本身不能说明何以会有这个规律性,这是各专门学科的任务。例如,用统计方法分析一些资料得出,吸烟与某些消化道疾病有关,这是从吸烟和不吸烟者的发病率的对比得出的结论,它不能解释吸烟何以会增加患这类疾病的危险性,这是医学这一专门学科的任务。所以统计方法只是一种工具,应用它进行定量分析时必须和定性分析结合起来。尤其是将统计方法应用于社会经济领域更应如此,因为社会经济现象比自然现象更为错综复杂,而又不可能像自然科学那样在实验室内排除其他因素进行试验。统计分析方法虽然是一个强有力的工具,但必须慎重使用,任意解释数据而不了解其背景是危险的做法。例如在原苏联沙俄时代曾有统计资料表明,医生接生的婴儿死亡率高于接生婆接生的死亡率。能否由此说明医生的医务水平以及卫生条件不如接生婆呢?实际情况是当时医生还不普及,人们只有遇到难产时才去找医生接生,因此二者有不可比的因素。同时也需要指出,定性和定量分析相结合也不等于先有了概念再去搜集数据,因为人们总是能精心选择数据来支持某项假设。虽然统计方法中也常用假设检验的方

法,但必须遵循科学的方法,不能是主观有意识地挑选数据。而且统计的结论往往带有不确定性,不能证明一项假设绝对正确,只是表明这项假设与取得的数据结论并不矛盾。统计分析中还应注意,若因两个现象恰好一起变动就作出它们之间有关系的结论,这样做也常常是有危险的。例如近年来我国电视机的拥有量和刑事犯罪案件都在增加,能否由此说明电视机的普及鼓励了犯罪呢?显然,这样的结论是危险的,刑事犯罪案件的增加是受多种因素影响的结果,因此有必要在应用统计方法时全面周密地考虑全部有关材料,要把统计学知识和其他有关专业知识紧密结合起来,才能更好地发挥统计这一工具的作用。

1.3 统计研究工作的程序和方法

应用统计学,研究和阐明了对反映社会各领域集体大量现象的数字资料,如何具体地应用统计描述和统计推断的种种具体统计方法的问题。应当指出:对于某一集体大量现象具体地应用统计描述方法和推断统计方法来进行统计分析、研究的工作,是一个步骤明确、相互衔接的调查(或观察)、表列、计算、分析研究的具体的工作过程。从事统计研究的科学工作者,只有理解并且掌握这一复杂的统计研究工作的程序和方法,才能获得统计分析、研究的科学结论。本节就是利用图表的方式具体地阐明统计分析、研究过程的必要的步骤和应该采用的统计描述或统计推断的统计调查、表列、计算、分析的具体统计方法。

下面首先列出统计调查、整理、计算、分析、研究工作图表,如图1-2,用以说明统计分析,研究全过程的八个具体步骤。

对于图1-2中八个步骤的具体内容,扼要说明如下。

(1) 统计调查或观察的原始材料

例如,某一国家全国人口普查按普查日(时点),对每一人口进行调查的项目,即性别、年龄、文化程度、婚姻状况、行业和职务等等全部填答的材料。再例如,某城市选样居民家庭收入支出的调查、登记的原始记录,就需要分月份、按日登记。

(2) 统计组织、整理

按统计表列分析的需要,对调查或观察所获得的原始材料,进行分类整理。例如,在全国人口普查的统计整理工作中,首先,就某一调查项目所获得的具体答案,按某项目预先规定的分类、分组(例如,性别按男子和女子)进行分类。然后,进一步就各类、组分到的个数进行综合计数,如按性别计算的男子和女子的人数,又例如按就业人口的具体行业(或称为部门)分类计算的农业(分为作物栽培、动物饲养、林业及渔业)、矿业、制造业、建筑业、交通运输、教育、行政等等部门计算的就业人数。这里需要指出的是,每个人的具体行业往往为一个具体的生产单位或服务单位