

● 马军海 主编

管理统计基础



天津大学出版社

TIANJIN UNIVERSITY PRESS

管理统计基础

马军海 主编



天津大学出版社

Tianjin University Press

内 容 提 要

本书紧密结合我国管理和经济工作的实际情况,讲述了现代管理统计基础知识、基本理论和基本方法。全书共分 11 章,主要内容包括统计总论,统计调查、统计整理与统计描述,平均数指标,标志变异指标,统计的概率基础,参数估计,抽样调查,假设检验,方差分析,相关分析与回归分析,多元线性回归分析等。它可以作为普通高等学校管理学、经济学及相关专业的教材,也可以供管理人员使用。

图书在版编目(CIP)数据

管理统计基础/马军海主编. —天津:天津大学出版社,2005.2

ISBN 7-5618-2104-2

I. 管... II. 马... III. 管理学:统计学
IV. C931.1

中国版本图书馆 CIP 数据核字(2005)第 011919 号

出版发行 天津大学出版社
出 版 人 杨风和
地 址 天津市卫津路 92 号天津大学内(邮编:300072)
电 话 发行部:022-27403647 邮购部:022-27402742
网 址 www.tjup.com
印 刷 永清县晔盛亚胶印有限公司
经 销 全国各地新华书店
开 本 170mm × 240mm
印 张 19.25
字 数 422 千
版 次 2005 年 2 月第 1 版
印 次 2005 年 2 月第 1 次
印 数 1 - 4 000
定 价 25.00 元

前 言

本书是在参阅了大量的国内外优秀统计学著作和教科书的基础上,紧扣我国管理和经济问题的特点,结合近年来我们所承担的管理统计学课程的教学体会编著而成的。其显著特点是,较为详细地介绍了现代管理统计学的基础知识、基本理论和基本方法,不过多涉及特别高深的数学推导及其证明,而将重点放在现代统计的理论和方法在管理和经济领域中的应用上面。对于比较复杂的统计学方法我们使用了深入浅出的通俗语言来描述,这样可以提高初学者对管理统计学的学习兴趣,使初学者全面掌握管理统计学研究问题的方法和思维方式,提升学员应用统计学的基础知识和基本技能熟练解决实际问题的技巧和能力。

书共分 11 章,主要内容包括统计总论,统计调查、统计整理与统计描述,平均数指标,标志变异指标,统计的概率基础,参数估计,抽样调查,假设检验,方差分析,相关分析与回归分析,多元线性回归分析。各章后附有习题。

本书内容翔实生动、通俗易懂,既可以作为教材,又便于读者自学;既适合于管理学、经济学及其相关专业的学生学习,也适合在管理和经济及其相关领域工作的人员使用。

王建俯同志参加了本书第 1、第 2、第 10 和第 11 章的部分编写工作;张晓峰同志参加了本书第 5、第 6 和第 7 章的部分编写工作。

在本书的编著过程中,参阅了大量的统计学著作、教材和相关资料,在此我们特向这些作者表示深深的谢意。尽管我们花费了大量的时间和精力来完成此书的编著工作,但书中难免存在一些错误和缺点,欢迎来自各方面的批评和指教,以便使此书的结构、内容越来越完善。

目 录

第 1 章 统计总论	(1)
第 1 节 统计学概述	(1)
第 2 节 统计学的研究对象、基本概念及其内容	(5)
习题	(11)
第 2 章 统计调查、统计整理与统计描述	(12)
第 1 节 统计调查与统计整理的概念及其作用	(12)
第 2 节 统计调查的基本方法	(14)
第 3 节 统计整理与统计描述	(24)
第 4 节 相对指标	(37)
习题	(40)
第 3 章 平均数指标	(42)
第 1 节 平均数的概念与分类	(42)
第 2 节 算术平均数	(43)
第 3 节 调和平均数	(47)
第 4 节 几何平均数	(48)
第 5 节 中位数	(50)
第 6 节 四分位数	(52)
第 7 节 众数	(54)
第 8 节 平均数之间的关系及在应用中应当注意的问题	(55)
习题	(57)
第 4 章 标志变异指标	(61)
第 1 节 标志变异指标的概念和作用	(61)
第 2 节 标志变异指标	(61)
第 3 节 偏度与峰度	(72)
习题	(77)
第 5 章 统计的概率基础	(79)
第 1 节 基本概念	(79)
第 2 节 事件的概率	(79)
第 3 节 条件概率	(81)
第 4 节 事件的独立性	(83)
第 5 节 随机变量	(84)
第 6 节 离散型随机变量及其概率分布	(85)

第7节 连续随机变量及其分布	(89)
第8节 大数定理及中心极限定理	(98)
习题	(102)
第6章 参数估计	(104)
第1节 点估计	(104)
第2节 区间估计	(108)
第3节 大样本的情况下两个总体均值之差的区间估计	(114)
第4节 总体成数的区间估计	(118)
第5节 单侧置信区间	(120)
习题	(121)
第7章 抽样调查	(123)
第1节 抽样调查的基本概念和作用	(123)
第2节 随机抽样的样本数目和代表性误差	(126)
第3节 简单随机抽样	(130)
第4节 分层抽样	(136)
第5节 等距抽样	(142)
第6节 整群抽样	(145)
第7节 多阶段抽样	(147)
习题	(151)
第8章 假设检验	(155)
第1节 假设检验的概念和原理	(155)
第2节 总体均值的假设检验	(161)
第3节 两个总体均值比较的假设检验	(168)
第4节 总体比例的假设检验	(175)
第5节 总体方差的假设检验	(180)
习题	(184)
第9章 方差分析	(189)
第1节 单因素方差分析	(192)
第2节 双因素方差分析	(197)
习题	(209)
第10章 相关分析与回归分析	(213)
第1节 相关分析	(214)
第2节 一元线性回归分析	(224)
第3节 相关与回归的区别与联系	(239)
习题	(241)
第11章 多元线性回归简介	(246)
第1节 多元线性回归模型	(246)

第2节 多元回归模型的相关检验	(251)
习题	(260)
附录1 ××市第五次人口普查实施办法	(264)
附录2 统计报表制度	(268)
附录3 利用 Excel 进行回归分析简介	(272)
附录4 随机数表	(276)
附录5 二项分布数值表	(279)
附录6 D.W.统计量临界值表	(282)
附录7 泊松分布表	(284)
附录8 标准正态分布表	(286)
附录9 t 分布临界值表	(288)
附录10 χ^2 分布临界值表	(290)
附录11 F 分布临界值表	(293)
参考文献	(299)

第1章 统计总论

在人们的日常生活中,统计问题以及与统计有关的问题随处可见。例如:人们在选购某种商品时一般要货比三家,要了解所要购买的同类商品的质量、产地、式样、商家的信誉以及市场的占有率等不同指标,然后再优中选优;股票持有者所关心的是上市公司的经营业绩情况;企业管理人员每天要掌握企业的生产销售情况和利润;而国家统计局所关心的是国家经济的增长速度,人们的收入水平、消费水平等等。这些数据的获得都离不开统计知识。总而言之统计已经渗透到人们的日常生活中的各个方面。统计是人们认识社会发展水平、发展速度和发展规模的有力工具。随着社会政治、经济的快速发展,统计在人们的日常生活和社会经济活动中的作用显得越来越重要。

统计学是一种方法,一种工具。狭义的统计学是指以数字表示的事实或数据;广义的统计学是指搜集、整理、表现、分析及解释数据资料,并通过科学的方法,在不确定的情况下从样本数据中获得结果,将杂乱无章的数据变成有价值的信息,依次来推断总体的性质与事物的发展变化规律,从而帮助人们做出正确决策的一门学科。

人类的行为与自然界的现象及科学的研究大都会产生大量的数据资料,要想将这些大量数据资料转化成有用的信息,只有依靠统计学的方法。在日常生活中,人们几乎每时每刻都要面对各种社会的经济和政治问题,而且,随着社会的发展与进步,这些问题也显得越来越复杂,越来越不确定。而统计学方法是帮助人们避免风险做出合理的决策的重要工具。

许多的社会现象与自然科学的研究,都可以将统计学的方法作为研究问题的工具。如今统计学已经被广泛地应用于物理、心理、生物、保险、教育、医疗、经济、农业、企管、会计、金融、营销、生产、社会、工程、政府管理等各个领域,并且收到了非常良好的效果。

统计学的产生和发展与人类社会的文明、社会的进步是紧密相关的。德国的斯勒兹(A. L. V. Schlozer 1735—1809年)曾经说过:“统计是动态的历史,历史是静态的统计。”历史伴随着统计而延伸,统计伴随着历史而发展。随着电子计算机的普及和互联网技术的飞速发展,现代社会已步入信息时代。统计作为信息的主体的作用也越来越明显。

第1节 统计学概述

一、统计(Statistics)的涵义

“统计”一词,英语为 statistics,用作复数名词时意思是统计资料,作单数名词时指的是统计学。“统计”一词最基本的含义是对客观事物的数量指标进行核算和分析,是人们对客观事物的数量表现、数量关系和数量变化进行描述和分析的一种计量活动。一

般来说,统计这个词包括三个含义:统计工作、统计资料和统计学。这三者之间存在着密切的联系。统计资料是统计工作的成果,统计学是统计实践经验的理论概括和深化,统计学来源于统计工作,统计学形成后,它对统计工作又起着理论指导作用。三者之间是理论与实践的辩证统一的关系。原始的统计工作即人们收集数据的原始形态,在文字产生以前就已经存在。例如,中国古代的结绳记事法,虽然距今已有几千年的历史,但它已蕴含着原始统计的基本思想。

而统计作为一门科学,则是从 17 世纪才开始的。众所周知,每一门科学都有其创立、发展的客观条件。统计科学当然也不例外。它是伴随着统计工作的发展并融合了社会经济理论、计量经济方法应运而生的一门边缘性学科。

二、统计的发展及统计学派

1. 统计的发展

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,适应国家宏观管理的需要而产生和发展的。自从有了国家,人们便开始了统计的实践活动。迄今为止统计已经有几千年的历史。我国在周朝(前 1046—前 256 年)时就有大量关于政府收入与支出的记载。这是我国较早的统计数据。在国外,统计的观念据称最早出现于埃及与巴比伦的人口普查的简单统计资料中,那时候统计只是用来作为计数的工具,后来则用于国家与其君主的财产登记。虽然这时统计已应用到社会生活的某些方面,但是早期人们并没有使用“统计”一词。最早作为学术名词使用“统计”的是 18 世纪德国政治学教授亨瓦尔(G. Achenwall)。他在 1749 年所著的《近代欧洲各国国家学纲要》一书的绪言中,他把国家学定名为“Statistika”(统计)这个词。原意是指“国家显著事项的比较和记述”或“国势学”,认为统计是关于国家应注意事项的学问。此后,各国相继沿用“统计”这个词,英国译为 Statistics。它源自拉丁文的 status,意思指各种现象的状态和状况。那时候统计是被用来表示国家的数据资料。1903 年钮永建、林卓南等翻译了四本横山雅南所著的《统计讲义录》,把“统计”这个词传到我国。1907 年彭祖植编写了《统计学》,这是我国最早的一本统计学书籍。“统计”一词就成了记述国家和社会状况数量关系的总称。

统计分析最初发源于英国。1662 年英国人 John Graunt 发表一篇名为《Natural and Political Observations Made upon the Bills of Mortality》的论文。他观察教堂记载的每周的死亡名单数据后,利用简单的统计方法加以分析,得出有关死亡的原因及疾病流行发展的趋势的粗略结论。这是人类首次根据大量原始资料应用统计学的方法来进行实证研究的案例。基于此 John Graunt 被后人称为现代统计学之父。

Graunt 论文发表之后,一些数学家对于统计学发生了浓厚的兴趣,开始涉足此研究领域,并且在此后的研究工作中作出了极其重要的贡献。比如,17、18 世纪时,著名学者 P. Fermat, B. Pascal 及 J. Bernoulli 等人对概率论的研究与杰出贡献,促进了人口统计、社会收入统计、保险与观测误差衡量的进步与发展。19 世纪时学者 P. S. Laplace, A. D. Moivre 及 C. F. Gauss 等对常态分布的研究与应用推广,促使统计学的发展向前迈进了一大步,而 Karl Pearson 与 Sir Francis Galton 等人发展出许多以概率论为基础的

统计推论方法,如相关分析法、回归分析法、卡方检验法等。由于众多学者在此领域的卓越贡献开启了现代统计学之门,也使得统计学作为一种方法被逐步应用到许多科学的研究之中。

20世纪30年代,学者 F. Wilcoxon 等人广泛开展了对于统计方法的研究,使得统计方法拓展到了更为宽广的应用领域。1935年 R. Fisher 提出了实验设计、变异指标分析、最大似然估计等方法,使得统计学在理论上出现了一个较大的飞跃。紧接着 J. Neyman 与 E. Pearson(K. Pearson 之子)在假设检验方面作出了卓越的贡献,20世纪60年代以后许多统计学家如 C. R. Rao, D. Blackwell 等人对统计估计理论的研究,促进了统计学的进一步快速发展。

在统计学的应用方面,19世纪中叶之后,由于古典概率的引入,使得统计的发展有了质的飞跃。20世纪30年代以后,世界各国已普遍采用随机抽样的方法对国家经济、人口、行业和部门等进行统计。虽然在那时现代统计理论已成雏形,随机抽样的方法也得到应用,但是它的更深层次的应用并不十分普遍与深入。这不仅是因为统计数据的储存累积及分析非常费时,而且因为统计计算过程也非常复杂。20世纪40年代由于电子计算机的普遍使用,大大提高了统计数据处理效率和准确性,更为数据的存储、更新、检索、计算及其一体化的自动化创造了良好的物质技术基础。尤其是开发了成套的统计软件程序以后,统计方法的应用更为便捷。这导致统计学在社会实践中的应用大为普及。20世纪50年代,世界各国普遍开始了以国民经济体系为整体的统计研究方法的推广应用,这自然大大提高了统计的认知能力。20世纪80年代由于个人计算机的快速发展与普及,大大促进了统计的发展与应用。现在统计学的方法和观念已不知不觉地渗透到人们的日常生活之中。

统计作为一种科学方法,应用的范围遍及自然科学及社会科学的整个领域中,大凡农业、工业、商业、教育、医药、政治、社会、经济等许多问题无不适合采用统计方法处理。现代统计学传入我国已有相当时日,不但政府机关较为重视,工商企业、事业单位等各行各业近年来也渐渐讲求科学管理,越来越多的应用统计方法。

2. 统计学派

在统计发展的近百年来,形成了许多学派,其中以数理统计学派和社会统计学派最为著名。数理统计学派的创始人是比利时的 A. 凯特勒,其最大的贡献就是将古典概率引入统计学,用纯数学的方法对社会现象进行研究;社会统计学派的首倡者是德国的 K. 克尼斯,他认为统计研究的对象是社会现象,研究方法为大量观察法。在近代统计学的发展过程中,这两个学派的矛盾是比较大的。我国大部分学者认为:社会统计和数理统计是两门对象不同、性质不同的统计学。二者既彼此独立,又相互联系,它们之间不可以相互替代,却是可以相互借鉴的。

三、统计的基本任务及其重要作用

1. 统计的任务

《中华人民共和国统计法》第二条写道,“统计的基本任务是对国民经济和社会发展情况进行统计调查、统计分析,提供统计资料和统计咨询意见,实行统计监督”。由此可

以看出,统计的任务可以概括为两部分:统计服务与统计监督。统计服务就是通过对国民经济和社会发展情况进行统计调查和统计分析,提供统计资料,为人们的决策提供服务。统计监督就是对国民经济的各个部门和社会生活的各个领域实施监督,及时发现问题,从而采取措施解决问题,保证社会经济的稳定协调发展。

2. 统计的重要作用

《中华人民共和国统计法》第一条写道,“发挥统计在了解国情国力、指导国民经济和社会发展中的重要作用,促进社会主义现代化建设事业的顺利发展”。可见统计的作用主要有两点,即认识作用和指导监督作用。

1) 认识作用

我们知道任何事物都是质与量的统一。社会经济现象也是如此。统计是利用数字去反映社会经济现象在具体时间、地点条件下的数量表现、关系和变化,反映社会经济的规模、水平、速度、结构和比例关系等。人们通过统计调查,可以增强对社会经济现象的感性认识,再经过统计整理分析,可以认识社会经济现象的规律,从而为驾驭规律、利用规律有效地为社会经济服务奠定基础。

社会经济统计学,是马克思主义奠基人在广泛使用统计学方法对资本主义进行经济分析的过程中,逐步建立和发展起来的。列宁强调指出,社会经济统计是认识社会的有力武器。毛泽东在《党委会的工作方法》中写道:“胸中有‘数’。这是说,对情况和问题一定要注意到它们的数量方面,要有基本的数量的分析。任何质量都表现为一定的数量,没有数量也就没有质量。我们有许多同志至今不懂得注意事物的数量方面,不懂得注意基本的统计、主要的百分比,不懂得注意决定事物质量的数量界限,一切都是胸中无‘数’,结果就不能不犯错误。”(《毛泽东选集》第4卷,第1380~1381页)多年来的实践经验表明,统计已成为解决社会经济问题的必不可少的工具和手段。

2) 指导作用和监督作用

统计在指导国民经济和社会发展中有着极其重要的作用。

列宁在《现代农业的资本主义制度》中写道:“一般的社会统计,特别是经济统计,最近二三十年来做出了巨大的成绩。有许多问题,而且是涉及现代国家的经济制度和这种制度的发展的最根本的问题,过去是根据一般的估计和大致的材料加以解决的,现在如果不根据某一个一定的纲要收集并经统计专家综合的关于某一国家全国情况的浩繁材料,就无法加以比较认真地研究。尤其是争论最多的农业经济问题,更加要求根据精确的和大量的材料作出回答,况且在欧美各国所有农户进行定期调查,已经愈来愈成为一种惯例。”(《列宁全集》第一版,第16卷,第420页)

1984年1月6日《国务院关于加强统计工作的决定》指出,“经济越发展,越需要加强统计。经济越搞活,越需要发挥统计的监督作用”。

总之,统计有着极其重要的作用。著名学者马寅初先生曾说过:“……学者不能离开统计而研学,政治家不能离开统计而施政,事业家不能离开统计而执业。”现如今统计已完全融入人们事无巨细的日常生活中,影响着人们的行为及观念,改变着人们对诸多事物的认知态度。

3. 统计的基本特点

统计学在研究目标、研究对象、研究内容、统计分析及学科性质等方面,表现出不同的特征,这些特征反映出统计的基本特点。

1) 数量性

统计的认识对象是社会经济及其相关现象。社会经济相关的问题可概括为质与量两个方面。统计的认识对象主要是指社会经济现象的数量方面,包括数量规模和数量关系。具体地说,统计是用规模、水平、速度、结构和比例的关系去描述、分析和揭示客观事物的数量表现、数量关系和数量变化,依此来揭示事物的本质,反映事物发展的规律,并对事物的发展前景作出预测。

2) 具体性

统计的数量性反映的是客观社会经济现象的规模水平,有具体的内容,表现为明显的时空等具体性特点。统计不能离开事物的质去孤立地研究事物的量,研究事物量的目的也在于认识事物的质及其发展变化的规律。

3) 社会性

统计所研究的内容是社会经济的演化数据,并从中探究这种规律的内在关联关系。基于此它就必然受到社会经济发展规律的支配。统计是在人们的社会生产实践的过程中产生和发展起来的,统计的数量总是反映人们社会生产生活的条件、过程和结果。统计研究的是通过数量特征和数量关系反映生产关系,以及其他社会关系的特点和实质。例如,国民收入的统计,失业人数的统计、货币供需的估计、商品供需的统计与预测、商品品质的管制、客户的调查统计与需求分析、商品市场占有率的估计、政府施政满意度的调查、民意调查、经济景气的调查分析、产业分析、股票价格的涨跌变化分析等,这些都是统计社会性的具体表现。

4) 总体性

统计学研究的目的在于探究复杂社会经济现象的总体数量特征。统计与其他具有数量性特点的学科的主要区别在于统计具有明显的总体性特点。社会经济统计首先要统计个别现象的数量,但是对个别量的观察只是一个过程,而不是目的。统计是通过对一个个具体单位的大量观察和综合分析,运用统计知识得出能够反映总体的数量特征。

第2节 统计学的研究对象、基本概念及其内容

一、统计学的研究对象

统计是认识社会经济现象的重要工具。它主要通过社会经济现象的数量方面的研究,揭示出社会经济系统的现状、内在本质及其演变的规律。

统计学所研究的社会经济现象的数量方面,是在质与量密切结合中的数量。任何现象都是质与量的辩证统一,任何质量都表现为一定数量,没有数量就没有质量。因此,必须在质与量的统一中,研究社会经济现象的数量方面,才能正确反映出经济现象

的数量规律。

社会经济现象是极其复杂的,受到众多因素的影响。有的是必然因素,有的是偶然因素。只有研究大量的数量,使得偶然因素所造成的数量离差相互抵消,呈现出现象的必然性,才能得到正确的结论。在统计的实际工作过程中,先从个别现象的具体事实开始进行调查、搜集资料,然后加以汇总和统计分析,从而可以揭示总体的规律。因为社会经济现象的数量是随着时间、地点、条件的变化而变化的,所以统计研究的数量必须是在一定时间、地点、条件下的数量。只有这样,数量才能具有意义,其结果才能符合客观实际,才能实现统计的服务作用和指导、监督作用。正是由于统计学的研究对象有这样的特点,《中华人民共和国统计法》第一条明确规定:必须保障统计资料的及时性和准确性。

二、统计学的基本概念

1. 总体与总体单位

客观存在的具有相同性质的多数单位所组成的集体,叫做统计总体,简称总体。它可分为有限总体和无限总体。

总体单位就是构成总体的每一个事物。总体和总体单位是相对概念,要根据统计研究目的对它们进行科学、准确的选择与界定。

例如,在研究某个学校的学生状况时,该学校的学生就是一个总体,构成这个总体的每一个学生就是总体单位。但要对该市各个学校的学生状况进行调查研究时,全市所有的学校就组成一个总体,该市的每一个学校就是总体单位。在界定统计总体时必须周密考虑总体的几个基本特征。

第一,组成总体的各单位必须在某一方面有共同的性质,也称为同质性。

第二,组成统计总体的各个总体单位在某一方面是同质的,但是,在其他一些方面又是有差别的,即差异性。如果统计总体的总体单位不存在差别,统计工作也就不必要了。

第三,组成总体的总体单位应该是足够多的、大量的,即大量性。因为组成统计总体的总体单位的特征具有变异性,为了能准确地反映总体的特征,必须要求组成总体的总体单位是大量的,只有这样才能抵消因个体的变异所带来的离差,才能反映总体的特征。

例 1.1 在调查某小型企业员工个人基本情况时得到表 1-1。那么该企业的所有员工就构成一个总体,其中每个员工就是一个总体单位。他们的共性在于他们都是该单位的员工,同时他们又表现出诸如性别、年龄、身高、体重、民族、公司服务年限和受教育年限等方面的差异。

表 1-1 某小型企业员工个人资料

姓名	性别	年龄	身高(cm)	体重(kg)	民族	公司服务年限	受教育年限
甲	男	33	185	65	汉	3	18
乙	女	25	165	55	回	2	16
丙	男	26	172	60	满	1	15
丁	女	35	160	53	回	4	16
戊	男	32	183	68	汉	2	19

2. 标志与标志数值

标志是指总体单位所具有的属性和特征。例如,就学生这个总体而言,每个学生的性别、年龄、籍贯、身高、体重、学习成绩等特征,可以说明每个学生的具体情况。这些属性和特征,就叫做标志。

标志可分为品质标志和数量标志。品质标志表明总体单位属性方面的特征,其标志只能用文字来表现,称为标志值。

例 1.1 中的性别、籍贯就是品质标志。

数量标志表明总体单位数量方面的特征,例如例 1.1 中的年龄、身高、体重、学习成绩所反映出来的数值就是数量标志。数量标志表现可以用数值表示,这个数值就是标志数值,有时也简称为标志值。

3. 变量与变异

1) 变量

可变的数量标志又称为变量。变量的具体表现值称为变量值。例如,工人的工龄、工资等标志又称变量。这些变量的具体表现值,如工龄 5 年、10 年;月工资 2 600 元、3 000 元等,就是变量值。有时为了研究的需要,可变的品质标志也可以用变量来描述,称为定性变量。这时,为了区别起见,将可变的数值变量称为定量变量。例如,表 1-1 中,年龄、身高、体重、公司服务年限、受教育年限是定量变量,性别、民族为定性变量。数量标志总是数值型的,而品质标志可能是数值型的,也可能是非数值型的。例如,为了方便数据的采集,将数据输入到计算机的数据库中,这样就可以使用数字代码作为性别变量:用 1 表示男性,用 2 表示女性;对于民族变量也可以做类似处理。在这种情况下,数值型数值就是用于标志性别和民族的标记或代码。即使显示的数据是数值型的数值,该变量仍是定性的。

就统计分析来说,定性变量和数值变量最显著的区别就是:对于通常的算术运算来说,只有定量变量才有意义。例如,对于定量变量,将变量值相加,然后再除以变量的个数,就可以计算变量的平均值。这个平均值有具体的实际意义,而且容易解释。但是,当定性变量被记录为数值型的数值时,这种算术运算的结果就没有任何意义。

定量变量按其变量值的连续性分为连续变量和离散变量。连续的或相邻两值之间可无限分割的变量称为连续变量,如身高、体重等都是连续变量。连续变量的数值要通过测量或计算的方法取得。各变量值之间是以整数断开的变量称为离散变量,如人数、

机器数、废品个数等。离散型变量可以按大小顺序排列起来,得到变量数列。

2) 变异

指标具体表现之间的差别称为变异。变异有属性的差别,也有数量上的差别。如性别标志的表现分为男、女,又如年龄标志的表现为 25 岁、35 岁等,这些表现的差异称为变异。

4. 指标

1) 统计指标的概念

统计指标是指表明统计所研究的社会经济现象的数量方面的科学范畴。社会经济现象及统计总体的特征,都具有复杂多样性,不同特征形成不同指标,同一总体的各个指标,分别从不同的角度反映了总体的特征。例如,人口数量、土地面积、国民生产总值、成本、利润、国民收入等概念,当它们是用于反映一定总体的数量时,就是指标。

统计指标总是要通过一定的数值加以说明的,这样的数值称为指标数值。指标数值表示社会经济现象在一定的时间、地点和条件下所达到的规模和水平。指标与指标数值既有联系又有区别。在通常情况下,反映某种现象的统计指标只有一个,而它的指标数值则可以有若干个。指标总是要通过数值来说明,而数值若离开指标也就没有任何实际意义。

2) 统计指标的分类

按指标的作用不同,统计指标分为数量指标和质量指标。

数量指标是指用来反映社会经济现象规模大小或总量多少的指标。一般用绝对数表示。例如工业总产值 11 000 万元、职工人数 2 300 人,等等。数量指标是计算质量指标和进行统计分析的基础。质量指标是指用来反映社会经济现象相对水平或工作质量的指标。它是由两个有联系的指标对比求得的,如废品率、平均工资、利润率等。它一般用相对数或平均数表示。

例 1.2 根据表 1-2 中的数据,对于该私营企业来说,数量指标包括以下内容:

$$\begin{aligned} \text{每月总薪金} &= 2\,000 \times 2 + 3\,000 \times 8 + 4\,000 \times 6 + 10\,000 \times 3 + 100\,000 \times 1 \\ &= 182\,000 (\text{元}) \end{aligned}$$

$$\text{总人数} = 2 + 8 + 6 + 3 + 1 = 20$$

质量指标包括的内容:

$$\text{平均薪金} = \text{每月总薪金} \div \text{总人数} = 182\,000 \div 20 = 9\,100 (\text{元})$$

表 1-2 某私营企业每月薪金表

职别	薪金(单位:元)	人数
工友	2 000	2
办事员	3 000	8
店职员	4 000	6
经理	10 000	3
老板	100 000	1

3) 统计指标和统计标志的区别与联系

统计指标和统计标志是一对既有明显区别又有密切联系的概念。二者区别是:指标是说明总体特征的,标志是说明总体单位特征的;指标具有可度量性,无论是数量指标还是质量指标,都能用数值表示,而标志不一定都能用数值来表示。数量标志具有可度量性,品质标志不具有可度量性。标志和指标的主要联系表现在:指标值往往由数量标志值汇总而来;在一定条件下,数量标志和指标之间可以互相表示。

三、统计学的内容

统计学的内容包括哪些呢?有些人会以为统计学只不过是將人类的行为与自然现象的结果加以记录、整理,并以一些数字、一些统计图表將数据或事物的特质表示出来而已。其实这只是统计学中较为粗浅和常见的部分,是所谓的叙述统计学。叙述统计学包括搜集、整理、表现、分析与解释资料。它是讨论如何搜集调查资料,以及將所搜集到的资料加以整理,表现为解释与分析。叙述统计学在一般人们日常生活当中最为常见与常用。新闻报道、电视、杂志,公司及机关学校,乃至家庭个人的资料等,大都可利用叙述统计学的方法將资料表现出来。事实上,统计学除了上述探讨资料的搜集、整理、表现外,还进一步采用科学的方法,利用已获得的样本数据去推论未知的总体,以获得较客观可靠的结果。这一部分称为推论统计学。推论统计学是將叙述统计中由样本数据所获得的结果,从其一般化推论至总体,或是由样本统计量去推论总体参数的方法,又称为归纳统计学。推论统计学基本上是在讨论总体与样本间的关系,从已知推论到未知,从局部推论至全体。统计推论是一个工具,在情况不明或不详的情况下,帮助决策者作出最佳的选择与判断。

举例来说,依据过去十年来的统计,每年来我国旅游观光的人数、平均每人停留的日期、平均每人每天的花费、十年内哪一年创最高记录等都属于叙述统计的范围;但是,如果我们根据这些年所得的数据来预测来年可能的观光客人数就是统计推论的问题了。

大致说来,统计推论分为三大类,即估计、检定和分类与选择。

例 1.3 美国加州某议员欲竞选州长,他想估计一下可能有多少人会投票给他,于是以随机抽样的方式询问 100 位有投票权的市民的意见,而后根据所得结果推论可能全市有多少人会选他。这就是典型的统计估计问题。又如某家庭主妇想知道洁王牌洗衣粉的洗净力是否比爱王牌洗衣粉的洗净力强,便首先假设洁王牌比爱王牌好,然后通过试验来推断这个假说是否成立。在本例中,家庭主妇并不想估计任何参数,而只是想检验事先所作的假设是否成立和可靠性有多大。这就是检定问题。还有,某人想知道新制造的三种药品中有哪几种比目前所用的同类药品的疗效更好,这是选择的问题。

统计推论的实施一般包括五个步骤,如图 1-1 所示。

1. 确定问题

首先必须确定问题之所在,明确分析的目的、对象与范围。

2. 搜集资料

针对研究对象、目的进行数据搜集。搜集资料时应考虑是否有现成可用的资料,搜集的成本或费用如何,并考虑搜集的方式、数据涵盖的范围等。

3. 审核整理、呈现资料

对搜集到的资料应先审核其是否完整、正确、合理与一致,然后利用叙述统计学中所介绍的方法进行分类整理,并以文字、统计图、统计表及统计测量数据将所获得的结果呈现出来。

4. 分析解释资料

根据整理的结果加以分析研究,探讨各数值间的相互关系并加以分析比较。

5. 统计推论

根据步骤 3、4 所得到的结果,来推论总体指标并下结论或提出建议。

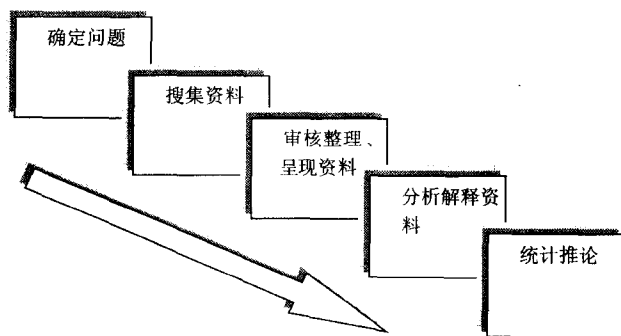


图 1-1 统计推论的实施步骤

统计的主要工作在于查询纷繁复杂的社会现象问题。为了达到这一目的,就要利用统计方法,找到适当的统计变量来解释这些现象。在叙述统计初期,范围只限于现象已经出现的事件上。这一时期的主要工作是整理统计数据使其简明化。如果这种社会现象是一个更庞大的事实,那么我们可以将这标志普遍化,以推论和预测它所能代表的这个庞大的事实的一些方面。这种情形比较复杂。我们要考虑这个庞大的事实规律性问题,如果可能控制,则进一步假设它服从某种概率模型;如果不可能控制,就不能设想这一控制模型的存在,只好根据它的一些经验结果来继续研究。目前研究得比较清楚和应用比较多的是,假设其服从某些已知的概率模型——分布函数,尤其是在数目大的时候。例如,我们可以从庞大的事实现象中抽取大量的有代表性的样本,也可以进一步考虑它是否服从常态分布函数。这就需要我们找寻到总体标志,来估计它的标志值,或推论及预测整个现象的趋向问题,甚至可以依照某些原则而做出某些决策。例如,利用防止最大损失的原则,或看重平均赚钱比较多这一类的原则,会引导我们对决策函数的判定。到了这一步,统计的结果不仅能叙述现象的一堆数据的特性,且能进一步建议人们采取某些行动了。

我们可以得出这样的结论:统计的目的在于先清楚地分析事实,然后通过数据和模