

统计学

王爱莲 史晓燕 编



西安交通大学出版社
XI'AN JIAOTONG UNIVERSITY PRESS

统计学

王爱莲 史晓燕 编



西安交通大学出版社
XI'AN JIAOTONG UNIVERSITY PRESS

内容提要

本书共分为七章,即总论、统计调查与统计整理、统计指标分析法、动态数列分析、统计指数、抽样推断和相关关系分析。本书在介绍统计学的基本理论与方法的同时,补充、增加了一些教学体系中的空缺点,删减了与概率论与数理统计重复的内容,每章后均附有覆盖各个知识点的思考与练习题。

本书通俗易懂、简明实用、新颖系统,可作为各大专院校经济学类、管理学类及其他人文社科类相关专业的专业基础课教材,也可作为从事统计工作及其相关人员的参考书。

图书在版编目(CIP)数据

统计学/王爱莲,史晓燕编著. —西安:西安交通大学出版社,2010.8
ISBN 978-7-5605-3610-1

I. ①统… II. ①王… ②史… III. ①统计学 IV. ①C8

中国版本图书馆 CIP 数据核字(2010)第 118619 号

书 名 统计学
编 著 王爱莲 史晓燕
责任编辑 屈晓燕

出版发行 西安交通大学出版社
(西安市兴庆南路 10 号 邮政编码 710049)
网 址 <http://www.xjtupress.com>
电 话 (029)82668357 82667874(发行中心)
(029)82668315 82669096(总编办)
传 真 (029)82668280
印 刷 陕西元盛印务有限公司

开 本 787mm×1092mm 1/16 印张 17 字数 415 千字
版次印次 2010 年 8 月第 1 版 2010 年 8 月第 1 次印刷
书 号 ISBN 978-7-5605-3610-1/C·96
定 价 29.00 元

读者购书、书店添货,如发现印装质量问题,请与本社发行中心联系、调换。

订购热线:(029)82665248 (029)82665249

投稿热线:(029)82664954

读者信箱:jdjgy@yahoo.cn

版权所有 侵权必究

前 言

统计学是一门搜集、整理、分析研究统计数据的方法论科学,是对客观现象进行定量分析的重要工具,是经济、管理类专业必修的学科大类基础课。统计学课程教学的最终目的是让学生真正掌握收集、整理和分析研究统计数据的一套科学方法,并能正确地使用统计分析方法,以满足对各类问题的实证研究、科学决策和经济管理的需要。

本书遵循理论联系实际、学以致用原则,在总结多年来的教学经验、并参阅大量的国内外相关资料的基础上编写而成。主要有以下几方面的特点:第一,删减了统计学和概率论与数理统计重复的内容,将数理统计方法有机地融合于统计学中。第二,补充、增加了一些讲授过程中的空缺点及实际中应用但理论上未体现的内容。第三,教材将重心放在统计方法的应用上,在全面系统介绍统计理论方法的同时,侧重阐明统计方法背后隐含的统计思想,以及这些方法在实际领域中的应用场合。每种方法、每个应用场合都选择典型的实例运用与分析,以便使学员正确掌握或使用统计分析方法,探索出统计数据所隐含的真正的规律性。第四,教材的每章后面附有相应的思考与练习题,练习题力求覆盖到每一个需要动手练习的基本方法;同时,为了便于学生对小知识点的思考与掌握,从各个角度编制配备了相应的判断、选择与简答题。

本书简明扼要、深入浅出、实用性强,可以作为高等院校经济类、管理类及其他人文社科类相关专业的专业基础课教材,也可作为教师及从事统计工作人员的参考书。

全书第一、四、五、六章由史晓燕编写;第二、三、七章由王爱莲编写。最后由王爱莲负责全书的总纂和定稿工作。

本书在编写和出版过程中得到西安石油大学经管院领导、教务处及西安交通大学出版社的大力支持,在此我们表示衷心的感谢。

由于编者水平有限,加之时间仓促,书中难免有疏漏或不妥之处,恳请同行专家和读者批评指正,以便我们修改与完善。

作者

2010年5月

目 录

第一章 总论	(1)
第一节 统计学的产生与发展	(1)
第二节 统计学的基本问题	(4)
第三节 统计学中的基本概念	(8)
思考与练习	(12)
第二章 统计调查与统计整理	(14)
第一节 统计调查	(14)
第二节 统计资料的整理	(26)
思考与练习	(39)
第三章 统计指标分析法	(43)
第一节 总量指标	(43)
第二节 相对指标	(46)
第三节 平均指标	(52)
第四节 标志变异指标	(67)
第五节 偏态与峰度的测定	(73)
思考与练习	(75)
第四章 动态数列分析	(84)
第一节 动态数列	(84)
第二节 动态分析指标	(86)
第三节 动态趋势分析	(101)
思考与练习	(120)
第五章 统计指数	(127)
第一节 统计指数概述	(127)
第二节 综合指数的编制与应用	(129)
第三节 平均指数的编制与应用	(139)
第四节 平均指标对比指数的编制	(147)
第五节 指数体系与指数因素分析	(150)
第六节 指数数列	(155)
思考与练习	(157)

第六章 抽样推断	(163)
第一节 抽样推断的基本问题	(163)
第二节 抽样误差	(171)
第三节 抽样估计	(184)
第四节 样本容量的确定	(190)
第五节 假设检验	(193)
思考与练习	(199)
第七章 相关关系分析法	(205)
第一节 相关关系分析概述	(205)
第二节 相关分析	(207)
第三节 简单直线回归分析	(219)
第四节 多元线性回归与曲线回归分析	(230)
思考与练习	(235)
附录 统计学常用表	(241)
附表 1 标准正态分布概率表	(241)
附表 2 t 分布临界值表	(243)
附表 3 χ^2 分布临界值表	(244)
附表 4 F 分布临界值表	(246)
附表 5 相关系数显著性检验表($n-2$ 为自由度)	(256)
附表 6 Spearman 等级相关系数检验表	(257)
附表 7 随机数码表(两位数)	(258)
附表 8 常用对数表	(259)
附表 9 平均增长速度累计法查对表(摘选)	(262)
主要参考文献	(266)

第一章 总论

第一节 统计学的产生与发展

以史为鉴,学史使人明智,了解统计实践活动及统计学的发展历史,对真正掌握统计学基本理论和基本方法是一种必备的基础。

一、统计实践活动的产生与发展

统计实践活动的产生距今已有几千年的历史,在统计实践活动发展到一定阶段,人们通过总结统计实践活动的经验并进行理论概括形成了统计学。为此,要了解统计学的产生,有必要先了解统计实践活动的产生和发展。统计实践活动是随着社会生产的发展和适应国家管理的需要而产生和发展起来的。原始社会,人类最初的一般计数活动蕴藏着统计实践活动的萌芽。随着奴隶制国家的产生,统治阶级为了对内统治和对外战争,需要征兵、赋税,很早就开始了人口、土地、财产和天文等统计工作。封建社会初期,社会经济统计已略具规模,我国在公元前三百多年前已经有了全国规模的人口调查登记制度以及国民经济调查中的各种数量对比分析。到了封建社会末期,随着生产力的迅猛发展,社会分工越来越细,这时,不仅政府需要有关国情国力等方面的统计信息,各类企业主和商人为了经营管理和争夺市场,也需要各种商业情报和市场信息,统计逐步扩展到各行各业。在西方,古希腊罗马时代就已经开始了人口和居民财产统计。古代巴比伦,埃及和罗马帝国政府都有人口和资源的详细记录。十五世纪,由于都市人口的急剧增加,瘟疫流行,英国、法国都开展生命统计(定期登记死亡出生、结婚人数、公布人口变动数字)。十六、十七世纪,欧洲的英国、法国、荷兰等资本主义国家经济迅速发展,适应国内外竞争和扩张的需要,使人口、税收、土地、商业,航运、外贸和工农业等许多方面的统计数字记录和传播达到了空前的规模,积累了大量的统计工作经验和资料。十七世纪以后,随着统计实践的发展,客观上要求总结丰富的实践经验,使之上升为理论,并进一步指导实践,于是开始了统计科学的研究,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的一个独立的部门和专业,正是在这样的历史条件下,欧洲出现了一些统计理论著作,标志着统计学的产生。所以,统计学的产生已有三百多年的历史。在后来的发展过程中统计学逐步形成了不同的学派。

二、统计学的产生与发展

(一) 古典统计学时期

统计学创始于十七世纪七十年代,从其产生之日起就有两个不同的统计学派:政治算术学派和国势学派(记述学派)。十七世纪中叶,在英国,威廉·配弟《政治算术》一书的问世,标志着古典政治经济学的诞生,同时也标志着统计学的诞生。在德国,差不多同时,产生了与“政治算

术”并存的国势学,它们在统计学史上形成了长期并存的两大学派。

1. 政治算术学派

这一学派的创始人是英国的威廉·配第(Willian Petty,1623—1687)和英国的约翰·格兰特(John Graunt,1620—1674)。

十七世纪中叶,英国资本主义已经有了很大发展,新兴的资产阶级为了管理国家,发展经济,同荷兰、法国争夺世界霸权,迫切需要了解本国和外国的经济状况,解决国内外社会经济问题,正是在这种历史条件下,英国产生了政治算术学派。配第著有《政治算术》一书,政治算术与政治算术学派的名称由此产生,奠定了统计学基础。《政治算术》一书是以数字资料为基础,用计算和对比的方法对英、法、荷的实力进行了比较,论证了英国称雄世界的条件和地位,尤为重要的是他所采用的独特的方法,主张一切都用数字、重量和尺度来说明,提出了用图表形式概括数字资料的理论和方法。这种理论和方法对后来统计学的形成与发展产生了深远的影响,马克思称他为“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。政治算术学派在英国兴起后,很快发展到了欧洲大陆,涌现出一批政治算术学家。但该学派一直没有采用“统计学”这一名称,可谓有统计学之实,而无统计学之名。

政治算术学派依据古典的资产阶级政治经济学为理论基础,利用数字、重量和尺度来研究社会经济现象的规律,并应用了一些统计方法,如大量观察法、分组法、综合指标的计算、统计表的编制(生命表)、估计推算法(人口推算)等,后来逐渐吸收了概率论的观点。政治算术学派为新兴的资产阶级服务的,符合当时社会发展的规律。某些计算虽不完善或存在一些错误,但就当时的科学水平来说,仍应给予较高的评价。

2. 国势学派(记述学派)

在政治算术学派形成的同时,在西欧封建制的德国产生了国势学派。这一学派盛行于十七、十八世纪,创始人是海尔门·康令(Hermann Conring,1606—1681),代表人物(继承人)为阿亨瓦尔(Gottfried Achenwall,1719—1772)。康令是大学教授,他的代表作是《国情论》,坚持国势学派是“实际政治家所必须的知识”。他们在大学中开设了一门新课程,最初叫做“国势学”。他们所做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。阿亨瓦尔是国势学派的主要继承人,他在1749年出版的《近代欧洲各国国势学论》中首先使用了“统计学”这个名称,用统计学代替了国势学,确定了统计学这门学科的名称,但是内容、研究对象及研究方法与以前的国势学并没有什么本质的区别。

当时正是欧洲封建制解体的时期,由于新大陆的发现,资本主义的萌芽,海外殖民地的发展,各国的统治者不仅要了解本国的政治情况,而且要了解和研究外国的政治经济情况,以便比较各国的实力,权衡轻重得失,以决定国策,在这种历史条件下,在西欧封建制的德国产生了国势学派。国势学派的研究对象是社会经济现象的一般情况,研究的方法是记述的方法,即以文字记述国家的显著事项。其支派表式学派则是以计量为主,用列表的方法表示国家的显著事项,开始体现出统计学的特点,并逐渐发展为政府统计,因而成为统计学的源流之一。这个学派广泛应用了“统计数字资料”、“数字对比”等统计术语和对比的方法,为后人所继承,沿用至今。国势学派虽然提出了“统计学”这个名称,但它有其名而无其实,因为它既不研究社会经济现象的规律性,不属于实质性科学,同时它又不研究如何记述或列表的原理、原则和方法,所以也不属于方法论科学,只能说他们是把统计学当作一门社会科学来对待的。

政治算术学派与国势学派都是以社会经济现象作为研究对象,均以社会经济现象的实际调

查资料作为立论的基础,均认为自己这一门科学是阐明国情国力的社会学科。他们的不同点在于是否把数量方面的研究作为这一门科学的基本特征。正是由于这样的共性和个性,使得两个学派共同发展和相互争论达两百年之久,两个学派在欧洲各国产生了很大的影响。直至十九世纪中叶,随着社会科学的发展和分工,统计学作为一门对社会经济现象进行数量对比分析的方法论科学已为社会公认。德国一位经济学家和统计学家克尼斯(1821—1898)于1850年发表了论文《独立科学的统计学》,提出了用“国家论”作为国势学的科学命名,“统计学”作为政治算术的科学命名。在统计学史上,以这篇论文的发表,体现了两派争论的结束,两个学派也就同时完成了它们的历史使命。在新的历史条件下,在统计理论的发展中,又产生了新的学派。

(二) 近代统计学时期

近代统计学时期是十八世纪末至十九世纪末这一时期,主要有数理统计学派和社会统计学派。

1. 社会统计学派

这一学派的先驱者是德国的克尼斯(K·G·A·Knies),他把用数值研究社会经济规律的政治算术作为唯一的统计学,把用文字记述的国势学恢复其名称,并排斥于统计学之外,为德国统计学界指出了统计学的发展方向,德国的统计学正是在这种统计思想的指导下发展为社会统计学派。德国的梅尔(Georg Von Wagr, 1841—1925)和恩格尔(C·L·E·Engel, 1821—1896)是这一学派的代表人物。

十九世纪中叶(1871年)德意志实现了统一,德国的资产阶级为了发展资本主义,与英法资产阶级争夺海外市场和殖民地,迫切需要掌握国内外大量完整的国民经济资料,并要求进行综合整理,以揭示社会经济现象的规律。在这种历史条件下,德国资产阶级统计学界在英国政治算术学派的影响下,以统计学是一门“具有政治算术内容的社会科学”为主张从而产生了社会统计学派。社会统计学派产生于十九世纪后半叶的德国,认为统计学的研究对象是社会现象的规律性,研究方法是大量观察法,主张统计学是一门独立的实质性的科学。十九世纪后半叶到二十世纪初,当数理统计学派尚未充分发展的时期,社会统计学派在欧洲大陆占有优势,并向世界各国传播,对于美国、苏联、日本、中国等国家产生了不同程度的影响。

2. 数理统计学派

数理统计学派产生于十九世纪中叶,创始人是比利时的阿道夫·凯特勒(A·Quetelet, 1796—1874),他把概率论方法引进统计学,使统计学进入一个新的阶段。十九世纪末和二十世纪初,资本主义进入了帝国主义阶段。自然科学中有了进化论、细胞学和能量守恒定律三大发现,这样的历史条件促使英美统计学界产生了数理统计学派。凯特勒在概率论已有相当发展的基础上,主张用研究自然科学的方法研究社会现象,把概率论与统计学结合起来,为后来数理统计学的形成与发展奠定了基础。在学科性质问题上,他明确地认为统计学是一门既研究社会现象又研究自然现象的独立的方法论科学。随着社会生产的发展和自然科学技术的进步,统计研究的领域不断扩大,概率方法的应用也日渐增多并渐趋成熟。到十九世纪末,数理统计学就从统计学中分离出来自成一派,由于它主要是在英美等国发展起来的,故又称英美数理统计学派。数理统计学在自然科学中的应用得到突飞猛进的发展,在社会科学中的应用则较少,二者没有有机地融合。

(三) 现代统计学时期

统计学的现代期是二十世纪初至今这一时期,主要有数理统计学派和社会经济统计学派两

大学派。不内门部。将学会并前式国制国博图县学样门一查与自改人似,脚基的引立式书陈查查

1. 数理统计学派

这一时期,数理统计学由于同自然科学,工程技术学等紧密结合,被广泛应用而获得迅速发展,进入了鼎盛时期,建立了推断统计、假设理论、决策理论等,拓宽了统计学的范围。20 世纪 60 年代后,数理统计学越来越广泛地应用数学方法,出现了许多以数理统计为基础的边缘科学(经济计量学、非参数统计等),数理统计学的应用日益广泛深入,在国际统计学术领域中地位大大提高,因此成为现代统计学的主流学派,主要以欧美为代表。

2. 社会经济统计学

社会经济统计学是从俄国十月革命后逐步建立起来的。受德国社会统计学派的影响,认为统计学是一门研究社会经济规律的实质性社会科学,其特点是利用大量观察法调查社会现象的事实,以发现其中的规律。社会经济统计学产生后,实践上曾经为社会主义国家高度集中的计划经济服务,在理论上如分组理论、指数理论等也有不少建树,被认为是统计学史上又一次质的飞跃。但是,由于二次世界大战后国际上的冷战局面和意识形态上的对立,使前苏联等社会主义国家的社会经济统计学不能及时汲取世界各国统计学发展的最新成果,按照科学自身发展的规律不断进步,特别是 1954 年前苏联统计科学会议以决议的武断方式来解决学术上的争议,绝对地排斥数理统计,否定统计学的方法论意义,因此扼杀了社会经济统计学发展的生机。

新中国的社会经济统计学是直接引进前苏联的,也曾经在高度集中的计划经济体制下发挥了重要的作用。但由于受到 1954 年前苏联统计科学会议的影响也缺乏生机,进步迟缓。进入八十年代以后,随着中国向社会主义市场经济体制转轨,统计也进入了全面改革的新时期,在统计理论界提出了“大统计”的思想。“大统计”是一种理念,一种思想,是顺应国际统计学发展趋势改变中国统计学现状的一种创举,它强调统计学“大家庭”的成员应相互学习,相互借鉴,共同促进统计学的发展。广义的统计学包括数理统计学和统计学原理,人们通常所说的统计学指的是统计学原理。数理统计学侧重于统计思想、原理、方法的研究,属于一般性科学的范畴。统计学则侧重于各领域数据处理与分析的研究,属于较为具体的科学范畴。统计学的这两个组成部分并不矛盾,在“大统计”的理念下,走向统一,共同发展。目前,“大统计”思想的宗旨仍在于不断发展和完善统计方法和技术,并在各个领域推广应用。

第二节 统计学的基本问题

一、统计学的含义

统计一词,源于中世纪拉丁语,意思是各种现象的状态和状况,由这一词根组成的意大利语表示国家的概念。德国学者阿亨瓦尔用德语统计来命名国势学,直到十八世纪末,德语统计学的译文传入英国,并赋予统计一词新的内容,即用数字表述事实。在中国古代,统计一词仅仅有数字总计的意思。现代含义的统计大约是在二十世纪初传到中国来的,今天,“统计”一词已被人们赋予多种含义,在不同的场合有三种不同的理解,即统计工作、统计资料和统计学。为此,要掌握统计学的含义,必须首先理解与之有密切联系的统计工作和统计资料的含义。

统计工作:即统计实践活动,是指统计人员利用科学的方法对客观现象的数量方面进行搜集、整理和分析研究工作过程的总称。例如,某统计师说我是干统计的,这里所说的统计指的就是统计工作。统计工作的基本任务是通过统计人员一系列的工作提供个人、公司和企业以及国

家所需要的各方面的统计资料,并进行统计咨询,实行统计监督。对统计工作的基本要求是准确、及时、系统地提供统计资料。

统计资料:是指统计工作过程所取得的各项数字资料以及与之相联系的其他资料的总称,如统计年鉴、统计图表、统计公报、统计手册、分析资料等都是统计工作所取得的资料。例如,电视台所说的“统计显示:”中的统计指的就是统计资料。统计资料包括原始的调查资料以及经过加工整理、分析而成的系统的统计资料。它是统计工作的成果或“产品”。统计工作的优劣直接影响统计资料的数量和质量。

统计学即统计科学,它是研究如何搜集、整理和分析客观现象数量方面的原理、原则和方式的科学。例如,某大学生说我是学统计的,这里所说的统计指的就是统计学。统计学是统计实践经验的科学总结,它来源于实践,又高于实践,反过来又用其理论和方法指导统计工作,推动统计工作的不断提高。随着统计工作的进一步发展,统计学不断地充实和提高,两者是理论和实践的关系。

二、统计学的研究对象和性质

(一) 统计学的研究对象

由统计学的发展史可知,统计学是从研究社会经济现象数量开始的,随着统计方法的不断完善,统计学得以不断的发展,统计学发展到今天不论是自然领域,还是社会经济领域的客观现象总体的数量方面,都是统计学所要分析和研究的,所以统计学的研究对象是自然、社会客观现象总体的数量方面。

(二) 统计学的性质

从统计学的发展史来看,统计学最初是作为一门实质性科学建立起来的,它从数量上研究具体的社会经济规律。但是,随着统计学研究范围的不断扩大以及统计方法在社会领域和自然领域的有效应用,加之统计方法体系本身的不断发展和完善,促使统计学从研究现象发展规律的实质性科学中分离出来,逐渐发展成为研究统计方法的方法论科学。所以,统计学是研究自然现象和社会经济现象总体数量特征和数量关系的一门独立的方法论性质的科学。

三、统计研究的过程

(一) 统计设计

统计设计是统计研究的首要阶段,是根据统计研究的目的,结合研究对象的性质、特点,对统计工作各个方面和各个环节所作出的整体规划。包括统计指标和统计指标体系、统计分类和分组、统计调查方式和统计方法、统计分析方法、统计工作组织与协调等方面的设计。统计设计的结果表现为各种设计方案,如统计指标体系、统计分类目录、统计报表制度、统计调查方案、统计汇总或整理方案以及统计分析方案等。统计设计方案是否科学合理、是否具有可操作性,直接影响到统计研究的各个阶段,也会影响统计研究目的实现。

(二) 统计调查

统计调查是根据统计设计的要求,采取切实可行的科学调查方式,具体地搜集统计资料的过程。对于大多数自然科学和工程技术研究来说,有可能通过有控制的科学实验去取得数据。对于社会经济现象来说,一般无法进行重复实验,要取得有关数据就必须到社会总体中选取一定单位进行调查。统计调查在整个统计工作过程中担负着提供基础资料的任务。

（三）统计整理

统计整理,就是根据统计研究的任务,对统计调查所搜集到的原始资料进行科学的加工整理,使之条理化、系统化,并编制成统计图、统计表的过程。由于统计调查所取得的原始资料是不系统、分散的,还可能带有一定的片面性,因此需要进行分类、汇总、加工。统计资料的整理,属于统计研究的第三阶段,在统计研究中起着承上启下的作用,既是统计调查阶段的继续,又是统计分析的基础和前提。

（四）统计分析

统计分析是指根据统计研究的目的,在加工整理统计资料基础上,计算各种统计指标,并运用各种统计方法和所计算出来的统计指标,对所掌握的统计资料进行分析研究,以认识事物的本质和规律性,并据此对未来的发展作出科学的预测。统计分析是统计研究过程中的最终环节。

综上所述,一般来说,一个完整的统计研究过程由统计设计、统计调查、统计整理和统计分析四个密切联系的阶段构成。统计研究从以定性研究为核心,构建整体研究框架的统计设计出发,经过侧重对个体事物量的认识的统计调查阶段和通过对个体事物量的综合,得到事物总体数量的描述性认识的统计整理阶段,最终到达以事实描述为基础,从定性和定量结合的角度,对现象总体认识的统计分析阶段。经过定性认识→定量认识→定性认识与定量认识相结合这一完整过程,从而达到对统计总体认识的升华。需要说明的是,在统计设计、调查及整理阶段也需要分析,在统计分析阶段还存在进一步调查、整理的问题,所以不能把四个阶段截然分开。

四、统计研究的基本方法

统计是研究客观现象总体数量方面的一门方法论性质的学科,已经形成了一些专门的研究方法,这些方法相互联系、互相影响,构成了统计研究的方法体系。在此仅就对统计研究全过程起重要作用的基本研究方法予以介绍。

（一）大量观察法

大量观察法是指从客观现象的总体出发,对其全部单位或足够多数单位进行数量观察的统计方法。它是统计研究的一种基本方法。统计研究的对象是客观现象总体的数量方面,只有通过总体中足够多的单位进行观察,才能发现其规律性。然而,客观现象的发展是在诸多因素错综复杂的作用下形成的。总体内的各个单位,由于各自的具体条件不同,既受到共同起作用因素的支配,也受着某些特殊的、暂时的因素的影响,使得它们的数量变化带有一定程度的偶然性和随机性。因此,统计研究不能任意抽取个别或少数单位进行观察,而要调查研究总体的足够多数单位,消除偶然性,才能揭示客观现象的特征和规律性。例如,为了研究城乡人民物质生活的提高程度,就要观察足够多数的职工、农民家庭的收支情况,才能做出正确的结论。大量观察法的数学依据是大数定律。大数定律是随机现象的基本规律,大数定律的一般概念是:在观察过程中,每次取得的结果不同,这是由偶然性所致的,但大量、重复观察结果的平均值却几乎接近确定的数值。大数定律的本质意义在于经过大量观察,把个别的、偶然的差异性相互抵消,而总体的、必然的规律性便显示出来。“大量”是一个相对的概念,只要所观察的单位能够说明现象总体时,就可以称为“大量”。有些现象总体十分庞大,但在实际中往往只能搜集到总体中一部分单位的资料,只要这一部分单位构成的样本能够代表总体,就可以称为大量观察。同时,统计研究在防止任意抽选个别单位进行观察的同时,并不排斥在某些特定问题的研究中,从总

体中选择少数具有代表性的单位进行调查。

(二) 统计分组法

统计分组法是根据研究目的和现象的总体特征,按照一定的标志把总体划分为不同性质或不同类型的几个组成部分的一种统计方法。统计是研究事物的量的,但它研究的不是脱离现象性质的量,而是同质现象的量。因此,观察与分析现象的数量方面必须以区分现象的不同性质为前提,这就需要运用统计分组法。统计分组法是统计研究的基本方法之一,它贯穿于整个统计研究过程中。通过分组可以把相同的部分归并在一起,把组与组明显区别开来。它是对总体进一步研究、区分内部差异的方法。通过分组可以研究总体的构成以及各个组成部分之间的比例关系。如在研究国内生产总值的构成时,把所有经济单位按三次产业划分,可以研究国内生产总值中第一、第二、第三产业增加值分别占多大比重,以及第一、第二、第三产业增加值之间的比例关系,等等。

(三) 统计指标法

统计指标法是统计分析的主要方法,是指运用各种统计指标来反映和研究总体现象的一般数量特征和数量关系的方法。通过计算各类统计指标可以从静态和动态综合反映总体现象在具体时间、地点条件下的总规模、内部结构、各组成部分之间的比例关系和依存关系、发展水平和发展速度以及总体的一般水平和变异程度等数量关系与数量特征。

(四) 统计模型法

统计模型法是统计分析过程中的重要方法,它是根据一定的经济理论和假定条件,以实际资料及一套相互联系的统计指标为基础,配合适当的数学模型对总体及其运动过程作出比较完整的、近似的反映或描述,以反映客观现象的数量特征和数量关系,揭示其运动规律的分析方法,是系统理论与统计相结合的产物,是在较高层次上认识事物的一种方式。也可以说,是大量观察法、统计分组法和统计指标法的进一步综合化、系统化,能够较为严谨地表现出总体的结构和功能。

(五) 统计推断法

统计推断法是根据部分总体单位组成的样本的数量特征去推断总体的数量特征。统计在研究客观现象的总体数量关系时,需要了解的总体往往很大,有时甚至是无限的,以致于不能进行全面调查研究。另外,由于客观现象之间的联系性和相似性,在很多情况下也不需要进行全面的统计调查。因此,统计推断已经成为现代统计学的基本方法,在统计研究中得到了极为广泛的应用,它既可以用于对总体参数的估计,也可以用作对总体某些分布特征的假设检验。

五、统计学与其他学科的关系

(一) 统计学与数学的关系

统计学与数学既有密切的联系,又有本质的区别。统计方法和数学方法一样,并不能独立地直接研究和探索自然现象和社会现象的本质规律,而是给其他学科提供了一种研究和探索其客观规律的数量方法。同时从学科层次上,数学是基础的数量方法,它为统计理论和统计方法的研究和应用提供了数学基础。

研究理论统计学需要较深的数学知识,应用统计方法也要具备良好的数学基础,这就造成了一种错觉似乎统计学是数学的一个分支,实际上统计学与数学有着本质的区别。虽然统计学与数学都是研究数量的,但数学研究的是可以没有量纲或计量单位的抽象的量,而统计学研究

的是有计量单位的具体的数量。同时,统计学与数学研究中所使用的逻辑方法不同,数学研究中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)

(二) 统计学和其他学科的关系

统计学是研究客观现象数量方面的。由于几乎所有的学科都要研究和分析数据,因而统计学与这些学科领域都有或多或少的联系。这种联系表现在统计方法作为一种有用的定量分析工具可以帮助其他学科探索其内在的数量规律性,而对这种规律性的解释只能由各学科自己的研究来完成。例如,利用统计方法对吸烟和不吸烟者患肺癌的数据进行分析,得出吸烟是导致患肺癌的重要原因的结论,而吸烟为什么会患肺癌,仅仅依靠统计学是无法说明的,而必须由医学作出解释。统计学发展的趋势之一就是与其他学科结合,用统计方法解决其他学科的具体问题。可以说,统计学的发展动力,越来越多地来自于其他各个学科不断给统计学“出难题”。

第三节 统计学中的基本概念

统计学在论述统计理论与统计方法时,要运用一些专门的概念,有些是基本的、常用的概念,有些是在论述专门问题时使用的局部的概念。本节就几个基本的、常用的概念加以阐述。

一、总体与总体单位

(一) 总体

1. 总体的含义

由客观存在的并至少具有某种相同性质的许多个别事物或个别单位组成的整体叫统计总体,简称总体。总体就是我们研究对象的全体。所以,明确了研究目的,才能确定相应的总体。例如,当我们要了解我国原油生产及加工情况时,全国所有的油田企业就构成了一个总体,首先它是客观存在的,其次它所包含的每个油田企业的经济职能是相同的,都是进行原油生产和加工活动的生产单位。

2. 总体的种类

总体按其所包含的单位数多少分为有限总体和无限总体。总体中所包含的单位数是有限的,称为有限总体,如全国所有人口所构成的整体。对有限总体根据研究目的既可以进行全面调查,也可以进行非全面调查。总体中所包含的单位数是无限的,称为无限总体,如连续生产的某种产品所构成的整体。对无限总体只能抽取一部分单位进行非全面调查,据以对总体进行推断。

3. 总体的特征

统计研究就是对于根据研究目的所确定的总体进行分析研究的,统计所研究的总体必须同时具有大量性、同质性、差异性三个特征。同质性是指构成总体的各个单位至少具有某一种相同的性质,是形成统计总体的一个必要条件。但总体的同质性并不要求构成总体的各个单位在所有方面都具有共同性,而只是当统计研究的目的确定后,构成总体的各个单位在特定点上应具有共同性。如研究工业行业总体的发展情况时,只要是工业企业就应该包括在内,而不管是钢铁工业还是石油天然气开采工业。当具体研究石油天然气开采工业的发展时,就只有是从事石油天然气开采的企业才能包括在内。大量性是指总体是由许多单位组成的。这是由统计研

究的目的所决定的,只有大量的总体单位,才能消除个别单位偶然因素的影响,呈现出统计研究所要达到的总体的一般特征和规律性。差异性是指构成总体的单位在一些方面是同质的,但在另一些方面必须是有差异的。全国所有油田企业构成的总体,虽然经济职能相同,但各油田的规模大小、经济效益、职工人数等是不同的。差异性使统计研究成为必要,如果总体内的各个事物之间不存在差异,那么统计研究就变成了毫无意义的活动。所以,统计研究是在大量性和同质性的基础上研究总体的差异性。

(二) 总体单位

总体单位是构成统计总体的个别事物或个别单位,也就是构成统计总体的最基本的单位,各项统计数据的原始承担者。如全国工业普查时,总体是全国所有工业企业,总体单位是全国每一个工业企业。

(三) 总体与总体单位具有相对性

总体的范围是根据研究目的来确定的,随着研究任务的改变,总体与总体单位可以相互转化。例如,要比较全国所有高校之间职工的工资收入差异情况,那么全国所有高校构成的整体是总体,各个高校(由该校所有职工构成的整体)是总体单位。如果旨在了解某个大学职工之间的工资收入差异情况,则该大学所有职工构成的整体就成了总体,每位职工就是总体单位了。可见,总体与总体单位不是固定不变的,而是具有相对性。

二、标志与指标

(一) 标志

1. 标志和标志表现

标志是说明总体单位的属性或特征的名称。在标志名称之后所具体表明的文字或数值是标志的具体表现。例如,我们要研究某企业工人的情况,那么每一位工人是总体单位,每一位工人的工种、性别、工龄、工资等是说明总体单位特征的称为标志。某工人的性别是女,“女”就是“性别”的具体表现,她的工资为2 500元,2 500元就是“工资”的具体表现。

2. 标志的分类

标志按变异情况可分为不变标志和可变标志。当一个标志在各个单位的具体表现都相同时,这个标志称为不变标志;当一个标志在各个单位的具体表现有可能不同时,这个标志称为可变标志。如人口普查时,在作为调查对象的人口总体中,国籍就是不变标志,而性别、年龄、民族、职业等则是可变标志。统计在对总体研究时,不变标志和可变标志分别发挥着重要的作用,不变标志是构成统计总体的基础,因为至少必须有一个不变标志将各总体单位联结在一起,才能构成一个总体,如果没有不变标志,总体也就不存在。如果总体中各单位在所有方面都相同即没有可变标志,就不需要进行统计研究,所以可变标志是统计研究的主要对象。

标志按其性质可以分为品质标志和数量标志。品质标志说明总体单位质的特性,是不能用数值表示的,只能用文字来描述。如工人的性别、民族、工种等。数量标志说明总体单位量的特性,是用数值表示的,如工人的年龄、工资、工龄等。数量标志的具体表现称为标志值。品质标志主要用于分组,将性质不相同的总体单位划分开来,便于计算各组的总体单位数,计算结构和比例指标。数量标志既可用于分组,也可用于计算标志总量以及其他各种质量指标。

(二) 指标

1. 指标的含义

对于指标一般有两种不同的理解并在两种不同的场合使用。一种理解是指反映总体现象数量特征的概念。如人口数、工资总额、工业增加值等,它单纯是指标的名称或范畴。另一种理解是指反映总体现象数量特征的概念和具体数值。如2008年我国原油产量为1.9亿吨,2008年全国国内生产总值为300 670亿元等。它不仅包括指标名称还包括了指标的时间限制、空间限制和具体数值,同一指标名称因统计的时间、地点、范围不同而有不同的数值,指标名称反映总体在某一方面的质的规定性,指标数值反映总体在一定时间、地点下量的具体表现。关于指标的这两种理解都是成立的,它们分别用于统计研究的不同环节或不同场合。在做一般性统计理论研究和统计设计时,只需要设计指标的名称。而在用经过搜集、整理、加工计算所得到的指标具体说明总体现象的实际数量状况及其发展变化情况时,就需要有具体的数值。所以后一种理解经常在统计实践中应用。

2. 指标的特点

(1) 数量性。指标是反映客观现象的数量特征的,所以所有的统计指标都是可以用数值来表现的,这是统计指标最基本的特点。正因为统计指标具有数量性的特点,才使统计研究运用数学方法和现代计算技术成为可能。

(2) 具体性。是指统计指标不是抽象的概念和数字,而是具体现象量的反映,是在质的基础上的量的集合。即指标总是总体某一方面质和量的具体统一,说明总体在具体时间、地点条件下的数量特征。

(3) 综合性。是指指标是通过总体单位调查登记并对其数量标志值加以汇总整理或进一步计算而得到的数值,说明的不是个别单位或部分单位的数量特征,而是许多个体现象数量特征综合的结果。如2008年全国工业增加值146 183亿元是当年所有工业企业增加值之和,2008年第三产业增加值在国内生产总值中所占比重40.1%是在对其数量标志值汇总基础上对比得到的,某企业职工平均工资是该企业所有职工工资总额和平均的结果。

3. 指标的分类

(1) 统计指标按其作用和表现形式的不同,可分为总量指标、相对指标和平均指标。这些统计指标的含义、内容、计算方法和作用各不相同,将在以后各章中叙述。

(2) 指标按其所说明总体现象内容的不同可分为数量指标和质量指标。

数量指标(总量指标)是反映总体现象的总规模或总水平的指标,它一般是由绝对数的形式表现出来的。如总人口数、原油产量、工资总额等。它的数值随总体范围的大小而增减。如某个省的总人口数肯定小于全国的,全国的原油产量肯定大于某个油田的。

质量指标是说明总体内部数量关系和总体一般水平的指标,它一般是以相对数或平均数的形式表现出来的,如人口密度、人均国内生产总值、平均工资等。它是数量指标的派生指标,其数值的大小与总体范围的大小无直接关系。

4. 指标体系

单个统计指标只能反映总体现象的一个侧面,客观现象是错综复杂的,要反映其全貌仅靠单个统计指标是不够的,需要把一系列相互联系、相互制约的统计指标集合起来建立指标体系进行考察。指标体系是由若干个相互联系、相互制约、相互补充的统计指标所构成的有机整体。

指标体系有两种表现形式,一种是通过数学公式表现出来的指标体系。如总成本、产量和单位成本构成的指标体系:总成本=产量×单位成本;再如原材料支出额、产品产量、单位产品原材料消耗量、单位原材料价格这四项指标也构成一个指标体系:原材料支出额=产品产量×单位产品原材料消耗量×单位原材料价格。另一种是指标之间仅仅存在一种间接的相互依存关系。例如,2008年科技部出台的《国家高新技术产业开发区评价指标体系》中规定,评价由知识创造和孕育创新能力、产业化和规模经济能力、国际化和参与全球竞争能力、高新区可持续发展能力4个一级指标构成,以及下设千人拥有研发人员数、单位面积营业总收入、高新技术产品出口额占高新区出口总额的比例、千人拥有的大专(含)学历以上从业人数等44个二级指标。这44个指标从不同的角度和侧面反映高新技术产业开发区的支撑、投入、产出及可持续发展能力。这些指标之间不存在、也没有必要采取数学关系式来反映它们之间的关系,而是存在的一种相互联系又相互补充的关系。

统计指标体系对于统计分析和研究具有重要的意义。一个设计科学且操作性强的统计指标体系,不仅可以描述现象的全貌和发展的全过程,而且还可以分析和研究影响现象的各种因素对现象总体变动结果影响的方向和程度,也可以对未来现象发展变化的趋势进行预测。

(三) 标志与指标之间的区别与联系

1. 标志与指标的区别

(1) 依附对象不同。标志是反映总体单位特征的,即标志依附于总体单位的;指标是反映总体数量特征的,所以指标是依附于总体的。没有总体单位和总体,标志和指标就是无本之木、无源之水。

(2) 具体表现不同。标志既可用数值表示(数量标志),又可用文字表示(品质标志);而指标只能用数值表示,不能用文字表示。

2. 标志与指标的联系

(1) 汇总关系:大多数指标数值是总体单位的数量标志值汇总而来的。如全国工业增加值(指标)是每个工业企业的增加值(数量标志值)加总得到的。

(2) 转换关系:由于研究目的或范围的变化,原来的总体(总体单位)变成总体单位(总体),相应的指标(标志)就变成标志(指标)。例如,如果所要研究的是全国工业企业的情况,则全国所有的工业企业构成一个总体,某个工业企业(甲公司)就是其中的一个总体单位,甲公司的增加值就是标志;如果所要研究的是甲公司的经营状况,那么甲公司就是总体,下属的分公司或二级单位就是总体单位,甲公司的增加值就变成了指标。

三、变量

(一) 变量和变量值

统计学上,把可变的数量标志称为变量。变量值是变量的具体表现,即可变的数量标志的不同取值。例如,如果研究的总体是所有的工业企业,每个工业企业就是总体单位,每个工业企业的年销售收入就是一个变量,甲工业企业年销售收入8 500万元,则8 500万元就是变量“年销售收入”的变量值。

(二) 变量的分类

变量按变量值是否连续分为连续型变量和离散型变量。连续型变量的数值是连续不断的,相邻两个数值之间可以作无限分割,可以用小数表示的变量。如零件的尺寸、产品重量、商品销