

应用统计

(第四版)

贾怀勤 主编

对外经济贸易大学出版社

应用统计

(第四版)

贾怀勤 主编

对外经济贸易大学出版社

(京)新登字 182 号

图书在版编目(CIP)数据

应用统计/贾怀勤主编. —4 版. —北京:对外经济贸易大学出版社,2006
ISBN 7-81078-774-8

I. 应... II. 贾... III. 应用统计学 IV. C8

中国版本图书馆 CIP 数据核字(2006)第 131570 号

© 2006 年 对外经济贸易大学出版社出版发行

版权所有 翻印必究

应用统计(第 4 版)

贾怀勤 主编

责任编辑:宋志红

对外经济贸易大学出版社

北京市朝阳区惠新东街 10 号 邮政编码:100029

邮购电话:010-64492338 发行部电话:010-64492342

网址:<http://www.uibep.com> E-mail:uibep@126.com

唐山市润丰印务有限公司印装 新华书店北京发行所发行

成品尺寸:185mm×230mm 25 印张 486 千字

2006 年 11 月北京第 4 版 2006 年 11 月第 1 次印刷

ISBN 7-81078-774-8

印数:0 001-5 000 册 定价:39.00 元

第四版前言

本书定位于经济和管理类非统计专业核心课教材,自1994年面市以来,承蒙诸多院校教师和学生选用,得以在贡献教学和回报社会方面连创佳绩。为适应不断发展的统计学教学需要,我们本着每隔数年修订一版的方针,于1998年出版了第二版,2002年出版了第三版。第二版主要是在内容架构上进行调整;第三版一方面压缩了原有文字,另一方面加写了“卡方检验和方差分析”;这次修订的第四版,除了内容安排上有一定的调整外,主要是对例题和习题进行了较大幅度的新编和修改,以使其更能反映社会、经济和企业管理的现实情况。此外在体例上,各章正文前编写了“教学目的和要求”。

这次修订的分工如下:贾怀勤负责绪论和第七章,刘春英负责第一、二章,杜学孔负责第三、四、五章,赵军负责第六章,丁岚负责第八、九章,王玉荣负责第十、十一章。由贾怀勤负责总纂。

我们将一如既往地欢迎全国各用户学校的老师和同学们提出宝贵的批评和建议。

编者

2006年8月

第一版前言

20 世纪 80 年代初,对外经济贸易大学在创办国际企业教育时,即将“应用统计”(Applied Statistics)定为一门主干课程。十余年来,本着以国际企业需求为导向、中西融通、文理渗透的方针,经老、中、青三代统计学教师不断探索,已建立起具有经贸特色的统计学教学体系,包括一整套教学大纲、教科书、教学参考资料、计算机应用软件和相应的教学方法,并有幸获得北京市优秀教学成果一等奖(1993 年)。本课程所用教材,原为李志伟先生等编著的《统计分析概论》(由对外贸易教育出版社 1984 年出版,1989 年修订再版)。

建设社会主义市场经济必然对经贸高等教育,当然也包括对“应用统计”的教学提出新的要求,特别是要求经贸工作者提高在国内外市场的复杂环境中运用统计方法和工具对经营管理信息进行分析,并作出科学决策的能力。为此,我们决定编写这部《应用统计》,作为各经贸专业本科“应用统计”课程的教科书,并力求使之更好地适应新形势下培养经贸人才的要求。这部教材对“应用统计”的知识框架重新作了安排,在内容的广度和深度上都较原用教材有所发展。

本教材由贾怀勤主编。其中,绪论、第一章、第五章、第十三章和第十四章由贾怀勤执笔;第八章、第九章、第十章、第十一章和第十二章由张杰执笔;第二章、第三章和第六章由朱雅华执笔;第四章由杜学孔执笔;第七章由王玉荣执笔。

恳请读者对本书不当和疏漏之处提出宝贵意见。

编者

1993 年 5 月

目 录

绪论	(1)
第1章 统计数据的采集和整理	(6)
1.1 统计数据的类别及其来源	(6)
1.2 数据的采集	(7)
1.3 统计数据的整理	(15)
1.4 统计表与统计图	(23)
习题	(36)
第2章 量别数据的描述性分析	(41)
2.1 集中趋势指标	(42)
2.2 离散趋势指标	(54)
2.3 偏态与峰度	(60)
2.4 位次指标	(63)
习题	(74)
第3章 概率与概率分布	(78)
3.1 概率的概念	(78)
3.2 概率计算规则	(83)
3.3 随机变量与概率分布	(98)
3.4 常用的离散型概率分布	(108)
3.5 正态分布	(124)
习题	(133)
第4章 抽样和抽样分布	(139)
4.1 抽样的基本概念	(139)
4.2 抽样分布的基本原理	(147)
习题	(161)

第 5 章 参数估计	(164)
5.1 点估计与区间估计	(164)
5.2 区间估计的应用	(170)
5.3 样本容量的确定	(178)
习题	(183)
第 6 章 假设检验的基本原理和程序	(186)
6.1 假设检验的有关概念	(186)
6.2 单一样本的假设检验	(199)
习题	(213)
第 7 章 假设检验的扩展	(216)
7.1 双样本假设检验	(216)
7.2 卡方检验	(223)
7.3 方差分析	(229)
习题	(240)
第 8 章 简单相关与回归分析	(245)
8.1 简单相关与回归	(246)
8.2 关于因变量的估计与推算	(261)
8.3 关于回归和相关参数的推断	(270)
8.4 判定系数	(273)
习题	(282)
第 9 章 多元回归和多重相关性分析	(288)
9.1 多元线性回归关系的描述和推断	(289)
9.2 多重相关分析	(298)
9.3 多元回归分析中的 F 检验	(305)
9.4 关于多元回归模型假定条件的讨论	(309)
9.5 多元回归和多重相关分析的综合运用	(313)
习题	(322)
第 10 章 指数	(328)
10.1 指数的编制	(328)
10.2 指数公式的再讨论	(335)
10.3 指数数列	(338)
10.4 指数的应用	(344)

习题	(346)
第 11 章 时间数列分析	(348)
11.1 时间数列解析概说	(348)
11.2 长期趋势的测定	(351)
11.3 季节指数的测定	(360)
11.4 循环变动的测定	(365)
习题	(367)
统计数表	(370)

绪 论

人类的统计活动,尤其是调查和整理社会经济总合现象数据的活动,有着悠久的历史。据一些统计史专家考证,关于人口、土地数字的搜集和整理活动,在中国可以上溯到殷、周时代,在外国可以追及古埃及、古希腊和古罗马。然而,由统计实务上升到理论和系统的方法的研究,最早也只能从 17 世纪算起。至今,统计学的发展大体上经历了三个阶段。

从 17 世纪中叶到 19 世纪中叶,为统计学的初创阶段。具体地讲,可以把英国的威廉·配第(William Petty)于 1671~1676 年写就的《政治算术》一书作为统计实务上升到理论和方法的标志。或者说,初创阶段的统计学,始于“政治算术”这门有统计学之实、却无统计学之名的学问。马克思说威廉·配第“在某种程度上也可以说是统计学的创始人”^①。在此阶段,统计学主要用来描述和比较各个国家的综合国力及人口研究,从研究对象上说是实质而不是方法论,从方法上看是经验研究多于理论研究,尚未进入系统的理论研究。

从 19 世纪中叶到 20 世纪 40 年代,是统计学发展的第二阶段——近代统计学阶段。1869 年,凯特勒(L. A. J. Quetelet)发表了《社会物理学》,揭开了近代统计学发展的序幕。凯特勒把作为数学分支的概率论引入了统计研究,解决了客观总合现象在数量变化上的规律性问题。而后,统计理论和方法又取得了一系列进步,如高尔登(F. Galton)的回归理论,戈塞特(R. A. Gosset)的小样本理论,费舍(R. A. Fisher)的 F 分布理论和方差分析方法,皮尔逊(E. S. Pearson)的区间估计方法等等,都极大地推动了统计理论和方法的发展和应用,特别是在自然科学和工程技术中的应用。至此,统计学已经建立起系统的理论和方法,并且实现了由实质性科学向方法论科学的转变。

从 20 世纪 50 年代起,统计学进入了它的第三阶段——现代统计学阶段。这个阶段带有三个明显的特点:其一是统计理论和方法的应用有了广泛的发展,不仅在自然科学研究方面大量应用统计方法,就是社会和人文科学也越来越广泛地应用统计方法,特别是在

^① 《马克思恩格斯全集》第 23 卷,第 302 页,北京:人民出版社,1972

经济和工商管理领域尤为如此;其二是进一步开发出一系列新的统计方法,如统计预测的新方法、多元统计方法和探索性数据分析等等;其三是统计研究与电子计算机应用密切结合,开发了一系列功能较为齐全,数据处理能力较强的统计专用软件(如 SAS, SPSS, STATISTICA 等),使得检索和处理大规模数据,以便从中导出对决策有用的信息成为可能,数据挖掘(Data Mining)技术从新的视角促进了现代统计学的发展。这三个特点也是相互联系的。统计学的应用扩展到自然科学、工程技术、心理学、经济和企业管理、社会、人口乃至语言文学等各个学科领域,极大地推动了这些学科的发展。反过来,统计学在各个实质性学科的应用又促进了统计理论和方法的发展。这既指一般统计理论和方法的进步,又指适用于专门领域的统计方法的开发。而计算机的应用使得现代统计方法的应用有了计算手段的保障,同时又为新的统计方法的开发提供了方便。

1. 统计学的定义、性质和研究对象

统计学是在质与量的辩证统一中,通过大量观察和平均方法,对客观总合现象(Aggregate phenomena)的数量特征进行研究的方法论学科。统计学离不开数。统计学涉及的数,不是随意的或单纯的数,而是能表明客观现象某一方面特征的数,如气温数、水压数、钢产量数、入境人数、不合格产品数,等等。关于客观现象特征的记载,有文字型资料(Alphabetic material)和数据型资料(Numerical data)两类。只有后者才是统计研究的素材。有的客观现象特征不适合用数字描述。另外一些客观现象特征既可以用文字描述,又可以用数字描述。对于后者,使用数字描述往往更确切。比如说“黑龙江省北部地区冬天非常冷”,就不如说那里“一月份平均气温在 -30°C 以下”更确切。统计资料(Statistics)是反映一定空间、时间条件下客观总合现象数量特征的数据型资料。统计资料又称为统计数据或统计数字(Statistical data)。它有数量性、客观性、具体性和总合性四个特点。在一定的意义上讲,统计学就是处理数据的科学——它以数据为研究对象,研究数据如何获取、组织、总结、表达,如何根据数据的分析结果寻求和揭示其内在规律性。

统计学是一门通用的方法论科学,按照认识世界和研究世界的方法的适用面,各种学科可以划分成三个层次:首先是最一般的世界观和方法论的科学——哲学,其次是一般的方法论科学,再次是实质性科学。统计学属于第二层次。正如形式逻辑不研究思维的具体内容的特殊规律,而研究思维的逻辑形式及其规律一样,统计学不是阐明社会领域和自然领域万千事物本身的实质性规律,而是阐明适用于研究各类总合现象数量特征的一般理论和方法。当然,也正如形式逻辑在研究思维的逻辑形式及其规律时不能抛开思维的具体内容一样,统计学在阐述通用的总合现象数量特征的研究方法时,也不能脱离作为其研究素材的各种总合现象数量特征的统计资料。

2. 大量观察法和平均法是统计学的基本方法

统计研究的是客观总合现象的数量特征,因而须采用大量观察法(Method of mass observation)。所谓大量观察,就是对同质的客观现象所组成的集团中的大量个体逐一观察,由此获得反映所有个体现象某些特征的一组数据,作为加工、分析的原始数据(Raw data)。这里的同质客观现象个体所组成的集团,在统计学上称为总体(Population)。组成总体的个体成员,称为总体单位(Population element)。每一个体都有一系列的属性(Characteristic),如体现在一个人身上的属性有性别、年龄、婚姻状态、受教育程度、收入水平等,体现在一个企业上的属性有注册地、资产额、销售额、主营行业等。总体是在某一属性相同基础上许多个体的集合,它们在其他属性上则不尽相同。这种不同或表现为数量的不等,或表现为状态的差异。表示总体单位属性的名词称为标志(Attribute)。凡表现为数量上不等的称为数量标志(Quantitative attribute),表现为状态上差异的称为品质标志(Qualitative attribute)。对数量标志和品质标志的观察分别产生量别数据(Numerical data)和质别数据(Categorical data)。

大量观察法的“观察”是一个广义的概念,泛指获取总体单位标志数据的行为。“大量”的意思是足够多,即只有对足够多的总体单位进行观察才能正确认识总体的特征。如果情况允许对全部总体单位进行观察,可以考虑实施普查(Census)。但通常只观察部分总体单位,即进行抽样(Sampling)。这部分被观察的总体单位就构成了样本(Sample)。样本的成员即为样本单位。对于总体,我们强调其具有同质性、变异性和大量性。构成总体的总体单位,如不至少在某一属性上相同,则该总体没有确定的范围。变异性指总体单位在许多感兴趣的属性上呈现差别。倘总体单位感兴趣的属性趋同,则实无统计之必要。总体应涵盖大量单位,非如此不能大量观察,也就不可能认识总合现象数量特征。

大量观察法在哲学上的根据就是必然性与偶然性的对立统一。恩格斯说:“被断定为必然的东西,是由纯粹的偶然性构成的,而所谓偶然的东西,是一种有必然性隐藏在里边的形式。”^①恩格斯又指出:“在表面上是偶然性起作用的地方,这种偶然性始终是受内部隐藏着的规律性支配着,而问题只是在于发现这些规律。”^②大量观察法的意义在于,在同质总体的基础上,经过大量观察,把个别的、偶然的差异性相互抵消,而将总合的、必然的规律性显示出来。例如对新生儿性别比例进行观察,如果只观察某10个新生儿,其性别比例可能是7:3,或5:5,或其他数值。但是经过对新生儿性别的大量观察,发现这一性别比例稳定在105:100上,从而显现出新生儿男女性别大体平衡而男略多于女的自然规

① 《马克思恩格斯全集》第4卷,第240页,北京:人民出版社,1972

② 《马克思恩格斯全集》第23卷,第243页,北京:人民出版社,1972

律。又就一个城市的各家庭人均收入而言,各户高低不一,甚至有些户之间相差悬殊。但是通过大量观察,可以找到全市人均收入的一般水平。这个一般收入水平,是由该市经济发展水平、就业情况和工资政策等多种宏观因素决定的。尽管各户收入不等,但是上述宏观因素决定了全市的一般收入水平必然如此;反过来,这个一般收入水平又体现在千差万别的各户收入之中。

客观总合现象数量方面的必然性与偶然性的对立统一,体现在总体的集中趋势与离散趋势的辩证关系上。而集中趋势和离散趋势的测度又无不通过平均法(Method of Mean)实现。保利(A. L. Bowley)说“统计学就是平均法科学”,这种说法在一定意义上是正确和透彻的。平均法是统计学的基本分析方法。总体的同质性、变异性和大量性是平均法的应用前提。在同质的基础上,通过大量观察,使用平均法,可以将总体单位标志表现的差异抽象掉,得出总体集中趋势的测度值。离散趋势的测度值也建立在总体单位标志表现差异平均化的基础上。其他许多统计分析方法,如动态现象趋势分析,指数法,期望值标准决策等等,也都是以平均法为基本分析方法的。平均法对于认识社会经济规律是非常重要的。马克思在《资本论》研究中,对于平均社会必要劳动量、平均利润率、平均工资、平均地租、平均资本构成、平均市场价格等等,无不以平均数作为规律的表现形式。那些牵动着全球大小小每一家公司的中枢神经、可以引发全球经济“大地震”的著名股票价格指数,如纽约道一琼斯指数,东京日经指数,那一轮轮原关税贸易总协定和世界贸易组织框架下的多边谈判旨在不断削减的关税税率,其计算也都是平均法的运用。当然,社会现象在量上的差异,必须以总体的同质性为前提。例如社会必要劳动量,指“在现有社会正常的生产条件下,在社会平均熟练程度和劳动强度下制造某种使用价值所需要的劳动时间。”^①这个同质前提就是生产同样的使用价值。马克思指出:“劳动的纯粹量的差别是以它们的质的统一或等同为前提的。”^②这句话透彻地说明了这个道理。

3. 统计学的分科

(1) 描述统计学和推断统计学

统计学对数据研究的程序大致包括数据采集、数据整理和数据分析三个阶段。如果使用普查方式采集数据,可直接得出对总体数量特征的认识;如果通过抽样方式采集数据,需要在对样本数量特征认识的基础上推断总体数量特征。描述统计学(Descriptive statistics)是阐述数据采集、数据整理和对数据做概略分析的原理和方法。推断统计学(Inferential statistics)是阐述依据样本数量特征推断总体数量特征的原理和方法。为了阐

① 《马克思恩格斯全集》第23卷,第52页,北京:人民出版社,1972

② 《马克思恩格斯全集》第23卷,第97页,北京:人民出版社,1972

明推断统计学,还需要有概率和分布知识做基础。

(2) 理论统计学和应用统计学

理论统计学(Theoretical statistics)指的是统计学的数学性原理,也称数理统计学,具有通用方法论的理学性质。应用统计学(Applied statistics)指的是理论统计学的基本原理应用到各个领域的数据处理方法、分析方法和推断方法。各个领域在理论统计学的基础上形成了方法各有侧重和特色的应用统计学,这些应用统计学的基本原理和基本方法是相同的,但都有特定的方法,比如医药统计学中的生存解析法,经济统计学中的指数分析法,人口统计学中的平均寿命计算法等。本书主要阐述涉及统计学在经济和管理领域应用的原理和方法。

(3) 应用统计学与其他学科的关系

各领域的应用统计学是理论统计学与该领域实质性学科交叉发展的结果。统计学是借助于观察和分析具体的总合现象来研究观察和分析总合现象数量特征的一般方法。统计学不研究各种现象自身的发展规律。各种自然现象和社会现象都有其自身发展规律,对这些规律的研究只能由各门实质性科学承担。统计学提供的一系列研究方法,使得人们有可能透过对客观总合现象数量特征的观察与分析觉察这种现象自身发展规律的存在,或者加深对这种规律的理解。但是,对这样的实质性规律的论证和解释要由研究特定现象的实质性科学作出。我国第五次全国人口普查资料表明:零岁人口男女性比例为112:100,由此往上,一直到69岁人口的性比例的102:100,都是男性多于女性。自70岁人口起,性比例开始逆转,为96:100,随着年岁增大,性比例急剧下降,85岁以上人口的性比例为52:100,其他国家的人口统计数字也有类似分布。这些统计资料向人们揭示出一条人口规律,初生儿男性比例略大于女性,在较低年龄段和中年龄段,大体仍保持这一结构,然而由于女性寿命平均高于男性,到了高年龄段则呈现女性比例越来越超过男性的趋势。至于如何解释这条规律,则不属于统计学研究的任务,恐怕要由关于人的生理规律及心理规律的学科来回答。同样,统计方法也帮助分析经济和管理领域事物和现象数量的规律。但是要准确地、自如地运用统计方法解决理论问题或实际问题,必须熟悉相关学科的知识,如经济学、金融学、管理学、营销学、会计学、财务管理学等学科知识,同时还要求具备相关领域的实务经验。

第 1 章 统计数据的 采集和整理

教学目的和要求

本章介绍统计数据采集和整理的理论和方法。通过本章的学习,要求:

1. 了解初级资料的采集方法;
2. 了解普查、抽样调查及统计报表等调查方式的特点及适用环境;
3. 了解统计调查方案的基本内容;
4. 了解统计数列的种类;
5. 掌握统计分组的方法;
6. 掌握频数分布表的编制方法;
7. 了解统计表的结构,掌握统计表的编制方法和统计图的绘制方法。

1.1 统计数据的类别及其来源

1.1.1 初级数据与次级数据

获取、分析统计数据以形成统计信息的这样一个完整过程可以称为统计活动。统计活动涉及四个方面:① 原始数据的载体或提供者;② 统计数据的获取和编辑者;③ 统计数据分析者;④ 统计信息的使用者,以统计数据分析者为活动主体。这一活动的完整过程应包含在统计活动总体方案的指导下依次完成的数据获取、数据整理、数据分析、信息提供等工作。如果数据分析与数据获取和编辑置于同一整体方案指导下,从原始数据的载体或提供者那里获取数据,继而进行分析,就称统计分析使用的是初级数据(Primary data)。反之,如果进行分析时数据来自本项统计活动之外的数据采集和编辑结果,就称

统计分析使用的是次级数据 (Secondary data)。获取初级资料的行为称作统计资料的采集,而搜集泛指初级资料和次级资料的获取。

各种机构在从事其业务活动的同时都留下了记录,如公司的生产记录、销售记录、会计记录及政府的海关进出口记录、交通安全事故记录。出于工作需要和其他目的,一些机构也会将其某些业务记录加工整理,形成以应用统计表和统计图形式展示的统计数据汇编。当人们要研究某一课题做统计分析需要数据时,就可以去查阅这些业务记录及其派生的统计数据汇编。这些资料对某一使用统计数据开展研究的研究者来说都是次级资料。

然而业务记录并不能满足统计活动对数据的全部需求,有时需要直接组织进行以原始数据的载体或提供者作为工作对象的数据采集。对这些原始数据分析加工得到初级资料。

一次采集的原始数据,可以多次被利用。即是说,次级数据还可以进一步派生出下一轮次级数据。在实践中,引用别人发布出来的数据,乃至再引用别人研究成果所引用的数据为自己课题服务,屡见不鲜。

1.1.2 内部数据与外部数据

从一个企业的角度看,统计数据可以划分为内部数据和外部数据。凡是取自本企业内部的数据,不论是营销部门、生产部门、财务部门等职能部门所做的业务记录和由记录生成的数据表,还是在本企业内为完成当前课题而组织的职工调查所取得的数据,都属于内部数据。凡取自本企业以外的数据,无论是来自政府部门、中介机构、数据公司和咨询公司、高校和科研院所、新闻媒体、互联网等,还是直接组织的针对社会成员的抽样调查,都属于外部数据。

1.2 数据的采集

1.2.1 统计调查与实验

统计数据的采集分为统计调查 (Survey) 和实验 (Experiment) 两种基本方式,每种基本方式都可以采取观察 (Observation) 和询问 (Communication) 两条途径。派人员到所研究现象发生的现场去注视现象的即时状态和演化进程,并作好数据记录,即为观察。通过调查人员与被调查者的问答取得数据,即为询问。显然观察既适用于人,也适用于物,询问

只适用于人。统计调查是研究者对要研究的人或物不加任何影响的纯客观的观察或询问。实验是在控制环境和条件的前提下,对被研究的人和物施加影响,关注其演化进程和结局,以取得所需数据的科学方法。实验数据采集主要靠观察。如果以人为实验对象,也不排除通过询问途径采集数据。

1.2.1 统计调查

统计调查是指有计划、有组织地收集原始数据的方式,它不介入被调查现象的自身变化,只是客观地记录被调查现象的静态或动态的状况、规模。各种调查主体或出于纯粹向需求者提供统计数据的目的,或直接使用统计数据开展研究,他们进行的初级资料采集活动都是统计调查。其中,统计报表调查是调查主体和调查程序都具有独特性的调查方式。一般统计调查的组织方式是任何旨在纯粹向需求者提供统计数据或者直接使用统计数据开展研究的调查者都使用的,这些调查者有政府、企业、研究机构和大专院校、非政府组织等。统计报表调查则是政府通过其设立的综合统计机关和部门统计机关开展的定期采集反映国民经济和社会发展基本情况数据的特有调查方式。

按着在采集过程中实际调查对象是否与调查目标对象范畴基本一致,调查分为全面调查和非全面调查。前者旨在对调查目标对象范畴内的每个个体都进行调查,后者对调查目标对象范畴内的部分个体进行调查。无论是一般统计调查,还是统计报表调查,都有全面调查和非全面调查之分。

1.2.1.1 普查

一般统计调查中的全面调查组织方式称作普查(Census)。

普查是为某一特定目的而专门组织的一次性全面调查,它通常由政府组织设计和实施,主要用于采集处于某一时点状态下的社会经济现象总量,目的是掌握特定社会经济现象的基本全貌,为国家制定经济社会发展规划、政策提供依据。世界各国一般都定期进行各种普查,如人口普查、工业普查、农业普查等。

普查具有如下特点:

(1) 普查需要规定标准调查时点和统一的作业时间,以保证普查结果的准确性,避免数据的重复和遗漏。

(2) 由于普查的涉及面广、调查单位多,需要耗费大量人力、物力、财力和时间,只限于基本国情、国力调查,且不能频繁实施。

(3) 普查可以是一次性的,但也可以是周期性的。世界各国周期性的普查的间隔期一般为10年。非普查年份的数据可以用样本数据与相应的周期性普查数据的对应和比例关系进行推算得出。

1.2.1.2 抽样调查

抽样(Sampling)调查是重要的非全面调查。它是从调查对象中抽取一部分单位作为样本进行观察,然后根据样本数据了解调查对象的总体状况。

抽样调查有随机抽样调查和非随机抽样调查两种形式。如果自总体中抽取样本单位的程序符合随机原则——在每一个总体单位被抽取之前面临同等被抽取机会,这样的抽样称为随机抽样(Random sampling),亦称概率抽样(Probabilistic sampling);如果不按随机原则进行抽样,则属于非随机抽样(Non-random sampling)。

随机抽样具有经济性好、时效性强、适用面广,可由此推断总体特征等特点,在社会经济问题的研究和科学实验中发挥着重要的作用。本书第4章将详述随机抽样的具体组织技术。

非随机抽样主要有判断抽样(Judgment sampling)和方便抽样(Convenient sampling)两种具体方式。判断抽样是数据获取者将其认为有可能反映总体特征主流的个体抽选出来,如编制物价指数时有意识地选择“代表性商品”,做企业调查时根据企业销售额标准选择少数线上企业等。方便抽样在访员下现场前不规定具体对象,而让他们在调查现场随客观机遇和方便条件碰到那个符合总体同质标准的个体就向其做调查。这种方式简单易行,在对总体单位构成和框架不甚了了的情况下可以使用。

非随机抽样调查由于没有概率保障,不能用于推断总体。有的教科书将之称为“选样”而不称为“抽样”。本书以后所说“抽样”,凡不加“非随机”定语的,即是指随机抽样。

1.2.1.3 统计报表

统计报表是按照国家有关法规的规定,自上而下布置,自下而上逐级提供基本统计数据的一种调查方式。它以一定的原始记录为基础,按照统一的表式、统一的指标、统一的报送时间和报送程序进行填报,带有一次性布置、周期性填报的特点。按报送频度不同,统计报表分为日报、旬报、月报、季报、半年报和年报。报送频度大的所要求填报的指标项目少些,只限于最重要的几项指标;报送频度小的所要求填报的指标项目多些,除最重要的几项指标外还可以有其他次要指标。按报送内容和报送的范围不同,统计报表可划分为国家的、部门的和地方的。

统计报表按调查范围不同,分为全面统计报表和非全面统计报表。在计划经济时期,全面统计报表大行其道,成为政府采集社会经济统计数据的最主要形式。随着计划经济向社会主义市场经济转轨,经济社会活动形式多样化,行业门类增多和细化,特别是各种所有制形式的经济组织并存格局的出现,都使得建立在政府分部门机构直管企业体制之上的全面统计报表难以为继,政府更多地采用抽样调查等非全面调查方式采集社会经济统计数据。但是,这并没有排除在少数调查活动中采用全面统计报表的可能。由于统计