



“十二五”普通高等教育本科国家级规划教材



Statistics

21世纪统计学系列教材

Sampling Technique

抽样技术 (第四版)

金勇进 杜子芳 蒋 妍 编著

中国人民大学出版社



“十二五”普通高等教育本科国家级规划教材



Statistics 21 世纪统计学系列教材

Sampling Technique

抽样技术 (第四版)

金勇进 杜子芳 蒋 妍 编著

中国人民大学出版社

· 北京 ·

图书在版编目 (CIP) 数据

抽样技术/金勇进等编著. —4 版. —北京: 中国人民大学出版社, 2015. 7
21 世纪统计学系列教材
ISBN 978-7-300-21622-5

I. ①抽… II. ①金… III. ①抽样调查统计-高等学校-教材 IV. ①C811

中国版本图书馆 CIP 数据核字 (2015) 第 160660 号

“十二五”普通高等教育本科国家级规划教材

21 世纪统计学系列教材

抽样技术 (第四版)

金勇进 杜子芳 蒋 妍 编著

Chouyang Jishu

出版发行	中国人民大学出版社	
社 址	北京中关村大街 31 号	邮政编码 100080
电 话	010-62511242 (总编室)	010-62511770 (质管部)
	010-82501766 (邮购部)	010-62514148 (门市部)
	010-62515195 (发行公司)	010-62515275 (盗版举报)
网 址	http://www.crup.com.cn	
	http://www.ttrnet.com (人大教研网)	
经 销	新华书店	
印 刷	北京鑫丰华彩印有限公司	版 次 2002 年 6 月第 1 版
规 格	185 mm×260 mm 16 开本	2015 年 7 月第 4 版
印 张	19.5 插页 1	印 次 2015 年 7 月第 1 次印刷
字 数	424 000	定 价 39.00 元

版权所有 侵权必究

印装差错 负责调换

总 序

改革开放以来,高等统计教育有了很大的发展。随着课程设置的不断调整,有不少教材出版,同时也翻译引进了一些国外优秀教材。作为培养我国统计专门人才的摇篮,中国人民大学统计学系自1952年创建以来,走过了风风雨雨,一直坚持着理论与应用相结合的办学方向,培养能够理论联系实际、解决实际问题的高层次人才。随着新知识经济和网络时代的到来,我们在教学科研的实践中,深切地感受到,无论是自然科学领域、社会科学领域的研究,还是国家宏观管理和企业生产经营管理,甚至人们的日常生活,信息需求量日益增多,信息处理技术更加复杂,作为信息技术支柱的统计方法,越来越广泛地应用于各个领域。

面对新的形势,我们一直在思索,课程设置、教材选择、教学方式等怎样才能使学生适应社会经济发展的客观需要。在反复酝酿、不断尝试的基础上,我们决定与统计学界的同仁,共同编写、出版一套面向21世纪的统计学系列教材。

这套系列教材聘请了中科院院士、中国科技大学陈希孺教授,上海财经大学数量经济研究院张尧庭教授,中国科学院数学与系统科学研究所冯士雍研究员等作为编委。他们长期任中国人民大学的兼职教授,一直关心、支持着统计学系的学科建设和应用统计的发展。中国人民大学应用统计科学研究中心2000年已成为国家级研究基地,这些专家是首批专职或兼职研究人员。这一开放性研究基地的运作,将有利于提升我国应用统计科学研究的水平,也必将进一步促进高等统计教育的发展。

这套教材是我们奉献给新世纪的,希望它能促进应用统计教育水平的提高。这套教材力求体现以下特点:

第一,在教材选择上,主要面向经济类统计学专业。选材既包括统计教材也包括风险管理与精算方面的教材。尽管名为统计学系列教材,但并不求大、求全,而是力求精选。对于目前已有的内容较为成熟、适合教学需要、公认的较好的教材,并未列入本次出版计划。

第二,每部教材的内容和写作,注意广泛吸收国内外优秀教材的成果。教材力求简明



易懂、内容系统和实用，注重对统计方法思想的阐述，并结合大量实际数据和实例说明统计方法的特点及应用条件。

第三，强调与计算机的结合。为着力提高学生运用统计方法分析解决问题的能力，教材所涉及的统计计算，要求运用目前已有的统计软件。根据教材内容，选择使用 SAS、SPSS、TSP、STATISTICA、EViews、MINITAB、Excel 等。

感谢中国人民大学出版社的同志们，他们怀着发展我国应用统计科学的热情和提高统计教育水平的愿望，经过反复论证，使这套教材得以出版。感谢参与教材编写的同行专家、统计学系的教师。愿大家的辛勤劳动能够结出丰硕的果实。我们期待着与统计学界的同仁，共同创造应用统计辉煌的明天。

易丹辉

前言

抽样调查是统计学专业的专业基础课,对非统计专业的学生而言,了解与掌握有关抽样调查的知识和技能也非常重要。抽样调查在国际上已有很长的发展历史,它是政府部门、各社会团体、企事业单位了解情况和收集信息的最主要方式。近年来,抽样调查在我国得到了广泛应用。实践证明,抽样调查是收集统计信息的最主要方法和手段。

传统的抽样技术理论都是在有限总体假定下,基于抽样随机性的统计推断。抽样设计的核心内容为抽样方法和估计量的选择,具体方法主要包括简单随机抽样、分层随机抽样、整群抽样、系统抽样、不等概抽样、多阶段抽样、比率估计、回归估计等。本书除了系统介绍传统抽样方法,还介绍了复杂样本的设计、方差估计和数据分析,讨论了几种主要的非抽样误差产生的原因、非抽样误差的测定模型以及对由于无回答造成缺失数据进行调整的方法等。

本书是在调研了读者意见之后,在第三版的基础上修改完成的,在内容上主要进行了以下修订:

(1) 对第2,3章的内容进行了删减,目的是压缩篇幅,使讨论的问题更集中、更有针对性;

(2) 对第9章的内容进行压缩,并降低了难度,在介绍复杂样本方差估计的基本思想基础上,重点剖析其中一两种方法的原理,侧重方法的应用;

(3) 增加抽样方法在R软件中的使用案例,以突出抽样方法的应用性和可操作性;

(4) 增加习题答案详解,方便读者使用;

(5) 对全书的公式进行了仔细核查,改正上一版中存在的错误。

总之,本书具有以下特色:

(1) 强调抽样技术的实际应用。抽样技术具有很强的理论性,但我们仍把它视为一门应用性课程,在论述中侧重于方法的应用,如不同方法的应用场合、应用条件及不同方法的特点比较等。为了与全书的基调和风格一致,本书没有拘泥于理论推导,略去这些推导并不妨碍对书中核心内容的理解。本书中的部分例题和习题以我们所从事过的实际调查项



目为背景。

(2) 书中第8~10章的内容在其他教材中并不多见,却非常实用。现实中的抽样设计往往是多种抽样方法的结合,第8章和第9章侧重介绍复杂样本的设计、方差估计和数据分析。第9章虽然内容有些复杂,但符合现代抽样技术的发展趋势,计算机技术的发展也为复杂样本的方差估计提供了方便。事实上,许多方差估计软件中的算法就是取自其中的。第10章介绍调查中的非抽样误差。大量抽样调查的实践表明,非抽样误差正在成为影响调查数据质量的一个十分重要的因素。该章讨论了几种主要的非抽样误差产生的原因、非抽样误差的测定模型、控制非抽样误差的方法以及对由于无回答造成缺失数据进行调整的方法。

(3) 加强案例分析。本书选取美国人口状况调查(Current Population Survey, CPS)作为案例,用一整章的篇幅予以介绍和剖析。之所以选择CPS,是因为它是国际上最著名的大型居民入户抽样调查项目之一,有70多年的发展历史,集世界调查统计学家之精华,其设计科学、巧妙,是抽样调查中的经典之作。我们从CPS的设计与方法中,可以得到许多有益的思考与借鉴。

(4) 增加抽样方法在R软件中的使用说明。R软件在统计工作和科研领域中的应用越来越广泛。应读者要求,本书在附录部分结合案例数据详细介绍简单随机抽样、分层抽样、系统抽样、PPS抽样,整群抽样和分阶段抽样在R软件的实现方法。

本书可以作为统计学专业学生抽样调查方面课程的教材,也可以用作非统计专业学生和各类人员学习抽样技术的教材或参考书。本书涉及内容较多,学习中可以根据不同的需求,有所取舍。

本书由金勇进教授、杜子芳教授和蒋妍副教授共同编写。金勇进编写第1, 8, 9, 10, 11章,并负责全书修订大纲的设计、书稿的组织和最后的统稿;杜子芳编写第2, 3, 6章;蒋妍编写第4, 5, 7章。本书中的部分习题选自所列的参考文献,恕不一一列举。大多数例题来自编著者所做项目的实际案例,或借鉴其他参考书中的例题进行设计,个别典型的例题数据取自其他书籍,在引用处均有注明。在此特向有关作者和出版社表示谢意。

为了方便老师教学和学生学学习,我们将本书的PPT、部分附录和部分习题答案放在中国人民大学出版社工商管理分社的网站(<http://www.rdjg.com.cn>)上,读者可以通过访问网站获得。

在本书出版过程中,中国人民大学出版社给予了大力支持和协助,在此我们表示深深的感谢。

尽管我们尽了最大的努力,但书中仍会有一些缺憾。对于书中的不足,恳请各位专家和读者提出宝贵意见。

金勇进 杜子芳 蒋妍
中国人民大学统计学院

教师教学服务说明

中国人民大学出版社工商管理分社以出版经典、高品质的工商管理、财务会计、统计、市场营销、人力资源管理、运营管理、物流管理、旅游管理等领域的各层次教材为宗旨。

为了更好地为一线教师服务，近年来工商管理分社着力建设了一批数字化、立体化的网络教学资源。教师可以通过以下方式获得免费下载教学资源的权限：

在“人大经管图书在线”（www.rdjg.com.cn）注册，下载“教师服务登记表”，或直接填写下面的“教师服务登记表”，加盖院系公章，然后邮寄或传真给我们。我们收到表格后将在一个工作日内为您开通相关资源的下载权限。

如您需要帮助，请随时与我们联系：

中国人民大学出版社工商管理分社

联系电话：010-62515735，62515749，62515987

传 真：010-62515732，62514775

电子邮箱：rdcbjsjg@crup.com.cn

通讯地址：北京市海淀区中关村大街甲 59 号文化大厦 1501 室（100872）

教师服务登记表

姓 名	☐ 先生 ☐ 女士		职 称		
座机/手机			电子邮箱		
通讯地址			邮 编		
任教学校			所在院系		
所授课程	课程名称	现用教材名称	出版社	对象（本科生/研究生/MBA/其他）	学生人数
需要哪本教材的配套资源					
人大经管图书在线用户名					
院/系领导（签字）： 院/系办公室盖章					

目 录

第 1 章 绪论	1
1.1 调查与抽样调查	1
1.2 基本概念	8
1.3 几种基本的抽样方法	13
1.4 抽样调查步骤	15
第 2 章 简单随机抽样	20
2.1 概述	21
2.2 简单估计量及其性质	27
2.3 比率估计量及其性质	37
2.4 回归估计量及其性质	46
2.5 简单随机抽样的实施	52
第 3 章 分层随机抽样	59
3.1 概述	59
3.2 简单估计量及其性质	61
3.3 比率估计量及其性质	66
3.4 回归估计量及其性质	71
3.5 各层样本量的分配	83
3.6 总样本量的确定	91
3.7 分层抽样的其他方面	96
第 4 章 等概率整群抽样和多阶段抽样	107
4.1 概述	107
4.2 等概率整群抽样	111
4.3 等概率两阶段抽样	120
4.4 等概率两阶段抽样设计	128



第 5 章 不等概抽样	133
5.1 不等概抽样	133
5.2 放回不等概抽样	137
5.3 多阶段放回不等概抽样	143
5.4 不放回不等概抽样	148
第 6 章 系统抽样	157
6.1 定义与实施方法	158
6.2 等概率情形: 估计量及其性质	160
6.3 不等概率系统抽样	171
第 7 章 其他抽样方法	175
7.1 二重抽样	175
7.2 捕获再捕获抽样	184
7.3 电话调查抽样	187
第 8 章 调查中的复杂样本	195
8.1 调查中的辅助信息	195
8.2 设计效应	200
8.3 样本的权数	204
8.4 复杂样本的数据分析	208
第 9 章 复杂样本的方差估计	219
9.1 概述	220
9.2 直接推导法	220
9.3 随机组方法	221
9.4 刀切法	226
第 10 章 调查中的非抽样误差	234
10.1 概述	234
10.2 抽样框误差	236
10.3 无回答误差	241
10.4 计量误差	249
10.5 离群值的检测和处理	255
第 11 章 设计与方法 —— 美国 CPS 案例	260
11.1 概述	260
11.2 CPS 的抽样设计	264
11.3 CPS 的目标量估计	268
11.4 CPS 的方差估计	272
11.5 非抽样误差及控制	277
附录 R 软件中的抽样方法应用	283
参考文献	304

本书主要讲授抽样理论，这些理论为实践中的不同抽样方法奠定了良好的基础。而本章是全书的第 1 章，客观上起着导言的作用。本章首先对有关调查的一些知识性问题做简单的讨论，这些讨论集中在抽样调查方面，如抽样调查的类型、抽样调查与普查的关系等。接下来对抽样调查中的一些重要概念做稍微深入的分析，这些概念包括：目标总体与抽样总体，抽样框与抽样单元，总体特征与估计量，方差、偏差与均方误差，抽样误差与非抽样误差，精度与费用等。本章还对一些常用的抽样方法进行了简单介绍，目的是使读者在接触抽样理论之前，对其应用背景有大致的了解。本章最后介绍了抽样调查的基本步骤。

1.1 调查与抽样调查

1.1.1 调查

调查是获取信息的方式，信息可以有多种，信息的类型或对信息的兴趣点不同，对调查的定义就可能不同。本书关注的是数据信息，所以这里把调查定义为：通过使用明确的概念、方法和程序，依据专门设计的调查方案指导的方式，从一个总体全部或部分单元中收集感兴趣的指标信息，并将这些信息综合整理成数据系列的有关活动。

上面的定义表述了调查的两个重要特征：首先，调查是一种有策划、有方法、有程序的活动；其次，调查的结果表现为收集到的数据。这里还需要指出，调查并不是获取数据信息的唯一方式。

数据有两种类型：实验数据和调查数据。收集数据的方法之一是通过实验，这



是指在实验方案的指导下,通过控制一个或多个变量,在有控制的条件下得到观测的结果。所以,实验数据是在实验中控制实验对象,依据实验设计的指导完成的。在可控条件下得到的数据是实验数据的重要特征。例如,检验一个小麦的新品种是否具有更高的生产率,实验人员将小麦不同的品种按某种设计方式分种在不同的地块中,控制某情景的所有相关方面,操纵少数感兴趣的变量,然后观察实验结果。其他的调查还包括检验化肥的增产效果、防治病虫最有效的杀虫剂、某种化学变化合适的温度等。这种类型的数据往往和实验的条件有关,改变进行实验的控制因素,结果就会发生变化。实验的次数可以是无限的。

调查数据一般是指客观上已经存在,但需要通过观察或询问才能得到的数据。例如,社会现象规模、水平、相互关系和发展变化的资料。具体说,它有如下特点:首先,这类资料大多与时间有关,数据所展示的是特定时期内或时点上的结果,如一定时期内的生产量、一定时点上的人口数等。其次,这类资料会随着时间的变化而改变,因此定期收集非常重要,因为每次收集的结果不仅展示了研究对象当时的状态,而且把以往收集的资料汇集在一起,构成时间数列,可以据以分析事物之间的相互影响和发展变化,这就为信息的进一步开发提供了广阔的空间。最后,还会有一些数据,它们在短期内相对稳定、变化不大。最常见的就是一个国家或地区的地理和地质资料,如地形、气候条件、土壤类型、矿物储量等。这类数据的调查往往技术性强,需要相关专业人员使用专门的设备进行。虽然这类调查的成本较高,但所取得的这方面的资料相对比较稳定,不需要经常更新。与实验数据相比,调查数据有一个显著特点,即一般情况下存在一定的误差。获得高质量的调查数据十分困难,这是因为调查对象通常是以人为主体的,虽然有些调查对象是客观事物,如土地资源,但也与人的活动有关。这种情况下利益关系就会卷入其中,调查结果就更富于变化。所以调查是一门与人打交道的艺术,好的调查会将人为因素的干扰降到最低限度。这不仅仅是一个调查的纯技术性问题,而且需要根据具体的调查内容、调查对象、调查方式和调查手段,发挥艺术的创造力。一个接近完美的调查一定是科学与艺术的结合。而且,在调查中获取数据的基本方式是询问,调查问卷是询问的依据,也是信息的载体。在有些调查中,调查者与被调查者面对面交流,如入户调查;在另一些调查中,调查者与被调查者不是面对面的,如电话调查、网络调查、留置问卷等;还有一些制度性的调查,要求被调查方填报。无论哪种调查,问卷(报表)都是必需的,所以问卷设计的质量就会对调查结果产生影响。本书是从统计学的角度讨论抽样理论,问卷设计并不是抽样理论中的内容,所以不在本书讨论范围之内,但需要指出,问卷的设计十分重要,好的问卷是收集高质量数据的基础,问卷设计得如何也是调查是否具有艺术性的一个标志。

1.1.2 抽样调查

抽样调查是调查应用最常见的模式,是一种非全面的调查,它是指从研究对象的全体(总体)中抽取一部分单元作为样本,根据对所抽取的样本进行调查,获得

有关总体目标量的了解。这是广义的抽样调查的概念。

从总体抽取样本的方法看,可以分为两类抽样:一类是非概率抽样;一类是概率抽样。

1. 非概率抽样

非概率抽样并没有严格的定义,其最主要的特征是抽取样本时并不依据随机原则。这类抽样有许多不同的具体抽取样本的方法。下面列举一些。

(1) 判断选样。判断选样是指在抽取样本时,调查人员依据调查目的和对调查对象情况的了解,人为确定样本单元。实践中确定样本单元通常有几种情况,一种是选择“平均型”单元作为样本,选定的样本可以代表感兴趣的变量的平均水平,目的是了解总体平均水平的大体位置;另一种是“众数型”,即在调查总体中选择能够代表大多数单元情况的个体为样本;再一种是“特殊型”,如选择很好或很差的典型单元为样本,目的是“解剖麻雀”,分析研究造成这种异常的原因。

(2) 方便抽样。方便抽样是指在抽取样本时依据方便原则,以达到最大限度降低调查成本的目的。典型的形式是“拦截式”调查,如在街边或居民小区拦住行人实施调查,在商场门口或柜台向消费者进行调查等。方便抽样操作简单,能及时取得所需要的信息,节省调查费用。该方法的主要局限是样本信息无法说明总体状况,无法根据样本信息对总体进行数量特征的推断,因而样本不适合描述性研究和因果关系研究。它比较适合探索性研究,通过调查发现问题的,产生想法和假设。它也可以用于正式调查前的预调查。

(3) 自愿样本。自愿样本不是经过抽取,而是由自愿接受调查的单元所组成的样本。比较典型的是网上调查,调查人员将调查问卷贴在相关网页上,感兴趣的读者可以自行填答。将调查问卷刊登在报刊上,读者可以自愿参加等。自愿样本的特点是样本的结构具有独特性,接受调查的样本单元往往属于某些特定的群体,例如,他们必须有接触到调查问卷的条件,同时对这项调查感兴趣,愿意积极参与。参与调查的群体与没有参与的群体可能会存在很多差异,从调查结果无法推断总体。但这种调查操作简单方便,成本低廉,而且参与者对调查内容感兴趣,调查人员能够了解这个特定群体的观点和看法,这对于了解情况、分析问题也是必要的,特别是如果调查目标就是了解这个特定群体的信息,自愿样本是一个不错的选择。

(4) 配额抽样。配额抽样是将总体中的各单元按一定标准划分为若干类型,将样本数额分配到各类型中,从各类型中抽取样本的方法则没有严格限制,一般采用方便抽样的方法抽取样本单元。

在配额抽样中,可以按单一变量控制,也可以按多变量控制。表1—1是单一变量控制的例子。在一个城市中欲采用配额方法抽出样本量 $n=500$ 的样本,控制变量有年龄和性别,配额按单变量分配,如各个年龄段的配额和不同性别的配额。单变量配额简单,但可能出现偏斜,如年龄小的女性多,年龄大的男性多。表1—2是两个变量控制的例子。



表 1—1

单变量配额分配表

年龄	配额
20~29 岁	150
30~39 岁	150
40~49 岁	100
50 岁及以上	100
合计	500
性别	配额
男	250
女	250
合计	500

表 1—2

交叉变量配额分配表

	男	女	合计
20~29 岁	70	80	150
30~39 岁	75	75	150
40~49 岁	55	45	100
50 岁及以上	50	50	500
合计	250	250	500

除了两个变量配额控制的例子,类似地还可以构造更多的控制变量。当然,选择的控制变量越多,在不同变量交叉的格子中所能分配到的样本量就越少。样本量过少,会显得缺乏说服力。

多变量配额控制可以使样本分布更为均匀,样本结构与总体结构更接近。在许多调查中,配额抽样有广泛应用。但在抽取样本时没有要求随机原则,所以配额抽样属于非概率抽样,同样具备非概率抽样的共有特征。

非概率抽样的应用背景体现在几个方面,一个重要的应用是调查结果可以用于了解情况、形成想法。有时调查的目标并不是对总体的特征进行推断,而是进行探索性研究,进而发现问题,寻找解决问题的途径,非概率抽样就是收集信息的重要途径。非概率抽样另一个重要的应用是充当预调查角色,作为开发概率抽样的初始步骤。由于概率抽样比较复杂、费时费力,为了保证效果,需要事先做预调查,对即将进行大规模调查的领域有所了解,对调查的问题有所检验,采用非概率抽样进行预调查也是不错的选择。有时,调查的目的是了解总体大致情况,并不要求对调查的精度进行评估,也不要求计算目标变量的置信区间,只需要了解其所在的大致范围,这种情况下也可以使用非概率抽样。非概率抽样的优点是操作简单,不需要抽样框,经济、快速,调查数据的处理也容易,所以有广阔的应用空间。非概率抽样的局限是不能计算抽样误差,不能从概率的意义上控制误差,样本数据不能对总体情况进行推断。同时由于抽取样本时具有较大的随意性,调查员通常选择那些容易接触的、比较友好的单元进行调查,从而导致被调查单元间存在系统性差异。

2. 概率抽样

概率抽样也称随机抽样,是指依据随机原则,按照某种事先设计的程序,从总体中抽取部分单元的抽样方法。它具有以下几个特点:

(1) 按一定的概率以随机原则抽取样本。所谓随机原则,就是在抽取样本时排除主观上有意识地抽取调查单元,使每个单元都有一定的机会被抽中。需要注意的是,随机不等于“随便”,随机有严格的科学含义,可以用概率来描述,而“随便”带有人为主观因素。例如,要在一栋楼内抽取 10 位居民作为样本,若采用随机原则,就需要事先将居住在该楼的居民按某种顺序编号,通过一定的随机化程序,如使用随机数表,抽取出样本,这样可以保证居住在该楼的每位居民都有一定的机会被选中。如果调查人员站在楼前,将最先走出楼外的 10 位居民选入样本,这就是随便而不是随机,这种方法不能使每个单元都有一定的机会被选中,已经在楼外的人不可能被选中,由于某些原因在调查时段不可能外出的人也没有机会被选中。随机与随便的本质区别就在于,是否按照给定的人样概率,通过一定的随机化程序抽取样本单元。

(2) 每个单元被抽中的概率是已知的,或是可以计算出来的。

(3) 当用样本对总体目标量进行估计时,要考虑到该样本(或每个样本单元)被抽中的概率。这就是说,估计量不仅与样本单元的观测值有关,也与其入样概率有关。

需要强调的是,概率抽样与等概率抽样是两个不同的概念。当谈到概率抽样时,是指总体中的每个单元都有一定的非零概率被抽中,单元之间被抽中的概率可以相等,也可以不等。若是前者,称为等概率抽样,若是后者,称为不等概率抽样。

概率抽样最主要的优点是,由于每个样本单元都是随机抽取的,而且能计算出每个单元的入样概率,所以能得到总体目标变量的估计值,并能计算出每个估计值的抽样误差,从而得到对总体目标量进行推断的可靠程度。从另一方面讲,也可以按照要求的精确度,计算必要的样本单元数目。所有这些都为调查方案的评估提供了有力的依据。当然,与非概率抽样相比,概率抽样也有一些不足,例如,概率抽样比较复杂,对调查人员的专业技术要求高;调查中需要抽样框,而构建和维护一个高质量的抽样框费用很高;随机抽中的调查单元可能非常分散,调查过程中不能轻易更换样本单元,这更增加了调查费用,延长了调查时间。虽然有这些不足,但其优点是其他调查方法所不具备的,所以概率抽样成为抽样调查中最主要的方式。

概率抽样调查可以视为狭义的抽样调查。本书后面讨论的抽样理论均是对不同的概率抽样方法而言。概率抽样是本书的主题。

1.1.3 抽样调查的作用

抽样调查的作用可以概括为以下几点:

(1) 节约费用。抽样调查能节约调查的人力、物力和财力,从而大大降低调查的费用。特别是当调查总体较大时,抽样调查的单元只占其总体的一小部分,因而



节约费用的特点表现得尤为突出。

(2) 时效性强。有些调查具有很强的时效性,要求在较短的时间内完成并提供调查数据。与全面调查相比,抽样调查所调查的单元少,数据采集和汇总整理的工作量较小,因而可以更快地提供调查结果。因此对于时效性要求比较强的调查,通常采用抽样调查的方法。

(3) 可以承担全面调查无法胜任的项目。对有些事物或客观现象,需要通过调查掌握其数据,但又不可能进行全面调查,这时必须采用抽样调查,如居民的家庭收支状况、电视节目的收视率,以及观察或测试具有破坏性的事物,如显像管的寿命、种子的发芽率等,这些项目的调查只能采用抽样的方法。

(4) 有助于提高调查数据的质量。虽然抽样调查只调查总体中的一小部分,用部分的调查结果推断总体,存在着抽样误差,但这只是问题的一个方面。抽样调查节约费用、时效性强,在某些情况下,还会得到比全面调查更准确的结果。这是因为一项调查的误差来自多个方面,全面调查由于参与的人员多、涉及范围大,因此虽然没有抽样误差,但在数据采集和数据汇总整理过程中却有产生其他误差的更大可能性。所以,调查规模并不是越大越好。与全面调查相比,抽样调查的工作量小,这就为使用素质较高的工作人员并对他们进行深入的培训创造了条件。此外,可以对调查过程实施更为细致的监督、检查和指导,使抽样调查所得到的数据质量比同样的全面调查数据质量更高,使调查的总误差更小。

1.1.4 抽样调查与普查

普查是一种全面调查的方法,它是指对研究总体中的所有单元逐一进行的调查。与全面调查相比,虽然抽样调查有许多特点和长处,但抽样调查不能代替普查,它们有各自的特殊作用。对于有关国计民生的重要现象,有时需要了解总体中每个单元的情况,就需要采用普查。普查不仅能够提供研究对象的总体情况,还可以提供各个区域、各种类别的统计信息,以满足各级政府行政管理的需要,而抽样调查在这些方面存在局限。普查资料还是构造抽样框的极好素材,所以抽样调查需要与普查相结合,互相补充。它们之间这种相辅相成的关系表现在以下几个方面:

(1) 抽样调查作为普查的补充。前面提到,对于有关国计民生的重要现象,需要采用普查的方法,了解总体中每个单元的基本情况,如我国进行过的普查有人口普查、全国经济普查、全国农业普查等。但每一次普查都需要投入很大的财力,不可能经常进行。这时可以在两次普查之间,采用抽样调查的方法,对该种现象的变化情况进行估计。例如,现在我国每10年进行一次全国性的人口普查,而中间的每年进行一次人口变动量的抽样调查,对当年的人口出生、死亡、迁移等情况进行估计,抽样调查对普查起到了补充的作用。

(2) 用抽样调查对全面统计资料进行评估和修正。例如,在一项普查结束后,通常采用抽样调查的方法,对随机抽取出一部分单位进行认真仔细的复核,对发

生错误的原因进行分析,计算误差率,作为对普查结果进行质量评估和数据修正的依据。

(3) 利用抽样调查进行深层次分析。由于普查的范围广,接受调查的单位多,因而调查项目不可能太多。在普查的基础上,根据研究的需要,可以针对某些问题,采用抽样调查的方法,获得更为详尽的资料,进行深层次的分析。

(4) 利用抽样调查,提前获得总体目标量的估计。普查所涉及的单位多,数据浩繁,整理汇总工作需要较长时间,为了尽快得到总体的某些特征的数据,可以采用抽样的方法,提前得到这些主要目标量的估计。

(5) 普查为抽样框提供资料。普查或其他全面调查的资料(例如某些统计报表)可以为抽样调查所需要的抽样框提供资料,提供辅助信息以提高抽样效率,同时也为样本轮换等提供基础资料。

1.1.5 抽样调查应用领域

近几十年来,抽样调查的理论和实践有了迅速的发展,抽样调查的应用越来越广泛。政府部门采用抽样调查的方法,收集统计信息,为制定政策、进行管理提供依据;学术机构、社会团体和企业也广泛采用抽样调查的方法收集数据,进行学术研究,了解社会情况和市场状况。可以说,凡是需要统计数据的领域,都有可能采用抽样调查。概括起来,抽样调查常常用于以下方面:

(1) 社会经济现象的调查。社会经济涵盖的范围十分广泛。目前我国政府统计部门制度化的抽样调查项目主要有:人口变动抽样调查;农产量抽样调查;城市居民住户抽样调查;农村经济抽样调查;小型工业企业生产情况调查;小型商业企业交易情况调查,等等。各部委根据本行业情况进行的抽样调查项目就更多。抽样调查已经成为政府部门获取统计信息的重要方式。

(2) 社会性的民意调查。民意调查在西方国家十分盛行,从总统选举到居民居住小区的改造,凡是人们关注的社会问题一出现,马上就伴随有相应的调查活动。频繁的调查活动培育出一批世界闻名的调查机构,如盖洛普。过去,我国的社会性民意调查基本上是一片处女地,随着改革开放进程的不断推进,民意调查越来越受重视,人们开始在这片沃土上耕耘。报刊上经常见到由各种学术单位和调查机构进行民意调查的调查报告,问题选择之精妙、涉及领域之宽广都是过去的调查无法比拟的。社会性的民意调查已成为调查业中一道靓丽的风景线,可以预言,人们一定会在这片沃土上取得更丰硕的果实。

(3) 市场调查。市场调查是企业获取市场信息的主要工具。市场经济越发展,竞争越激烈,市场信息就越重要。市场调查的对象通常是消费者,调查目的是了解不同消费者群体有关商品消费的事实、动机和意向。近年来,我国的市场调查发展很快,涌现出大批从事市场调查、咨询的专业性机构。我国人口众多,是一个巨大的消费市场,市场调查在我国有很大的发展潜力。