

经·济·与·管·理·类·统·计·学·系·列·教·材



普通高等教育“十一五”国家级规划教材

教育部高等学校统计学类专业
教学指导委员会推荐用书

应用抽样技术

(第三版)

李金昌 主编



科学出版社



经济与管理类统计学系列教材
普通高等教育“十一五”国家级规划教材

应用抽样技术

(第三版)

李金昌 主编

科学出版社

北京

内 容 简 介

本书全面介绍了抽样的一般原理、方法与技术。在对抽样的含义、种类、产生历史和作用等进行阐述的基础上,对简单随机抽样、分层抽样、整群抽样、系统抽样和多阶段抽样等抽样方式,比率估计和回归估计等估计方法,不等概率抽样、样本轮换、双重抽样、随机化装置和交叉子样本等常用抽样技术,分别进行了阐述和讨论。同时,还对非抽样误差问题作了专门分析。每章后均附有思考与练习题,便于教师组织教学和学生进行学习。

本书适合作为经济管理类统计学专业的本科生教材,亦可用作经济管理类其他相关专业的本科生教材。

图书在版编目(CIP)数据

应用抽样技术/李金昌主编. —3 版 —北京:科学出版社, 2015. 1

经济与管理类统计学系列教材 普通高等教育“十一五”国家级规划教材

ISBN 978-7-03-042994-0

I. ①应… II. ①李… III. ①抽样调查-高等学校-教材 IV. ①C811

中国版本图书馆 CIP 数据核字 (2015) 第 005942 号

责任编辑: 兰 鹏/责任校对: 鲁 素

责任印制: 李 利/封面设计: 蓝正设计

科 学 出 版 社 出 版

北京东黄城根北街16号

邮政编码: 100717

<http://www.sciencep.com>

北京华正印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2007 年 8 月第 一 版 开本: 787×1092 1/16

2010 年 8 月第 二 版 印张: 15 1/2

2015 年 1 月第 三 版 字数: 368 000

2015 年 1 月第十次印刷

定价: 36.00 元

(如有印装质量问题, 我社负责调换)

全国经济与管理类统计学系列教材编委会

编委会主任：

曾五一 教育部高等学校统计学类专业教学指导委员会主任委员、厦门大学教授、博导

编委：(以姓氏笔画排序)

王振龙 全国统计职业教育教学指导委员会副主任委员、陕西广播电视大学教授

王艳明 教育部高等学校统计学类专业教学指导委员会委员、山东工商学院教授

刘建平 暨南大学教授、博导

刘洪 中南财经政法大学教授、博导

朱建平 教育部高等学校统计学类专业教学指导委员会委员、厦门大学教授、博导

许鹏 湖南大学教授、博导

张润楚 南开大学教授、博导

李宝瑜 山西财经大学教授、博导

李金昌 教育部高等学校经济学类专业教学指导委员会委员、浙江工商大学教授、博导

杨灿 厦门大学教授、博导

肖红叶 国家级教学名师、天津财经大学教授、博导

周恒彤 天津财经大学教授、博导

庞皓 国家级教学名师、西南财经大学教授、博导

杭斌 山西财经大学教授、博导

罗良清 教育部高等学校经济学类专业教学指导委员会委员、江西财经大学教授、博导

茆诗松 华东师范大学教授、博导

郑明 教育部高等学校统计学类专业教学指导委员会委员、复旦大学教授、博导

徐国祥 教育部高等学校统计学类专业教学指导委员会副主任委员、上海财经大学教授、博导

蒋萍 东北财经大学教授、博导

雷钦礼 暨南大学教授、博导

黎实 西南财经大学教授、博导



总序

统计学是有关如何测定、收集和分析反映客观总体数量的数据，以便给出正确认识的方法论科学。随着社会经济的发展和科学技术的进步，统计应用的领域越来越广，统计已经成为人们认识世界不可或缺的重要工具。

现代统计学可以分为两大类：一类是以抽象的数量为研究对象，研究一般的收集数据、分析数据方法的理论统计学；另一类是以各个不同领域的具体数量为研究对象的应用统计学。前一类统计学具有通用方法论的理学性质，其特点是计量不计质；后一类统计学则与各不同领域的实质性学科有着非常密切的联系，是有具体对象的方法论，因而具有复合性学科和边缘学科的性质。所谓应用，既包括一般统计方法的应用，也包括各自领域实质性科学理论的应用。经济与管理统计学是以社会经济数量为对象的应用统计学。要在经济和管理领域应用统计方法，必须解决如何科学地测定经济现象即如何科学地设置指标的问题，这就离不开对有关经济现象的质的研究。要对经济和管理问题进行统计分析，也必须以有关经济和管理理论为指导。因此，经济与管理统计学的特点是在质与量的紧密联系中，研究事物的数量特征和数量表现。不仅如此，由于社会经济现象所具有的复杂性和特殊性，经济与管理统计学除了要应用一般的统计方法外，还需要研究自己独特的方法，如核算的方法、综合评价的方法等。

从历史和现状看，我国统计学专业的办学也有两种模式：一是强调各类统计学所具有的共性。这种模式主要培养学生掌握通用的统计方法和理论。它肯定统计学的“理学性质”，按照理学类学科的特点设置课程，概率论和各种数理统计方法等通用的统计方法论在课程中占有较大分量。其培养目标是有良好的数学基础、熟练掌握统计学基本理论与各种方法，同时有一定的专门领域的知识，能够适应各个不同领域的统计工作和统计研究的统计人才。二是强调各类统计学的个性，对于经济与管理类统计学来说，就是强调其与经济学和管理学等其他学科的密切联系，按照经济与管理类学科的特点设置课程，除统计学本身的专业课外，经济管理类的课程占相当大的比重。其培养目标是所谓的“复合型人才”，即具有坚实的经济与管理理论功底，既懂数理统计方法又懂经济统计方法，并能熟练掌握现代计算手段的经济与管理统计人才。这种人才既是统计人才又

是经济管理人才,不仅能胜任基层企事业单位和政府部门的日常统计业务,而且能从事市场调查、经济预测、信息分析和其他经济管理工作。上述两种办学模式,各有特色,同时也各有一定的社会需求。从我国的国情看,现阶段后一种模式培养的人才市场需求要更大一些。应该根据“百花齐放,百家争鸣”的方针,允许多种办学模式同时并存,由各院校根据自己的特色和市场对有关人才需求的大小,自主选择合适的办学模式。

为了更好地满足新世纪对统计人才的需要,无论是理学类统计学专业还是经济管理类统计学专业都有一个如何面向未来、面向世界、加强自身建设、更好地与国际接轨的问题。但是,这两类专业的培养目标不同,知识体系也有相当大的差异,难以完全统一或互相取代。2003年11月,教育部高等学校统计学专业教学指导分委员会在厦门召开年会,会上各方面的专家达成共识,为了促进统计学的学科建设和发展,有必要按授予学位的不同,分别制定指导性的教学规范。2004年1月,全国经济与管理类统计学专业的部分专家和学者在天津财经学院讨论了《统计学专业教学规范(授经济学学士学位)》征求意见稿,对初稿进行修改与补充,又经过教育部高等学校统计学专业教学指导分委员会研究和审定,最终形成了正式的教学规范(以下简称新规范),并已上报国家教育部。

根据新规范的设计,经济管理类统计学专业应开设的统计学专业主干课程包括以下10门:①统计学导论;②数理统计学;③应用多元统计分析;④应用时间序列分析;⑤应用抽样技术;⑥计量经济学;⑦国民经济统计学;⑧企业经营统计学;⑨证券投资分析;⑩货币与金融统计学。为了进一步提高统计教材的质量,更好地满足新世纪培养经济管理类统计人才的需要,我们成立了经济与管理类统计学系列教材编委会,组织全国高校同行分工协作,根据新规范设计的课程体系和教学内容,编辑出版一套经济与管理类统计学系列教材。本系列教材比较适合作为高等院校经济与管理类统计学专业的教材,其中《统计学导论》和《计量经济学》还可作为一般经济与管理类专业的核心课程教材。

参加本系列教材建设的有厦门大学、西南财经大学、天津财经大学、上海财经大学、浙江工商大学、山西财经大学、湖南大学、西安财经学院、南开大学、东北财经大学、中南财经政法大学、暨南大学、华东师范大学、江西财经大学、山东工商学院、复旦大学等10多所院校的教师。本系列教材实行主编负责制,担任主编和主审的老师都是曾经主编和出版过相关统计教材的国内知名教授,不仅具有一线教学的经验,而且对相关学科的发展趋势和学科前沿也比较熟悉。本系列教材的编写,力求体现以下特点:

(一) 与时俱进,构建与培养目标相适应的教学内容体系。

教材建设的关键在于构建与培养目标相适应的教学内容体系。为此,要根据时代的发展,不断补充和引进新的教学内容。作为新世纪经济管理类统计专业的教材,不能只是简单地向理学类统计专业的教学内容靠拢,而应该根据自身的特点,努力贯彻“少而精”和“学以致用”的原则。在大胆吸收国外优秀教材特点的基础上,对原有的体系重新进行整理和完善,既适当增加一些在经济社会分析中有良好应用前景的数理统计理论与方法的内容,又适当增补经济社会统计方面的最新进展。同时删除过时的和不再适用的内容。尽可能做到既反映本门学科的先进水平,又比较简明易懂,便于教学。

(二) 统筹兼顾, 防止低水平重复, 发挥系列教材的整体功能。

适应未来需要的经济管理类统计人才必须掌握多方面的知识和能力, 各种知识是相互联系的, 各门课程在内容上难免有所交叉。为了提高学习的效率, 更好地发挥系列教材的整体功能, 在编写本系列教材的过程中, 我们作了必要的协调和适当的分工, 尽可能做到统筹兼顾, 防止低水平重复。同时, 本系列教材采用相同的版式、体例和统一规范的学术用语。

(三) 与计算机结合, 培养学生的动手能力。

为了提高学生运用统计方法解决实际问题的能力, 本系列教材的编写注意与计算机的紧密结合。本系列教材中统计方法类的教材均根据教材的内容, 结合常用的计算机统计软件, 并给出相应的案例和数据。从而使学生不仅可以从中学习统计学理论和方法, 而且可以实际上机操作, 培养实际动手的能力。

(四) 编写体例新颖, 提高学生学习的兴趣和效率。

为了便于师生教学互动, 提高学生学习的兴趣和学习效率, 本系列教材在编写体例上也作了一些新的尝试。各章开篇有内容要点和教学要求提示, 章末附有小结, 对有关教学内容和计算公式作扼要的总结。教材中尽可能使用本国的真实数据作为案例。各教材的“思考与练习”部分不设一般常见的名词解释型的简答题和论述题, 而是通过判断题、选择题、计算题和有趣味的思考题, 来帮助学生掌握有关概念和计算方法。为便于学生自学, 同时又为其留有独立思考和独立完成作业的余地, 各教材均给出编号为奇数的习题的详解。

(五) 配套出版教学课件和光盘, 便于教师组织教学。

本系列教材在出版纸质出版物的同时配套出版相应的光盘。内容包括: 用 Power Point 制作的教学课件, 教材中有关案例的数据, 常用的统计表, 编号为奇数的习题的详解等。

经济与管理类统计学专业系列教材的建设是一个复杂艰巨的系统工程, 完成这一工程需要全国统计教育工作者的共同努力。感谢参与本系列教材编写的全国各高校的专家和学者, 感谢为本系列教材的出版提供帮助的科学出版社的领导和编辑。衷心祝愿大家的辛勤劳动能够结出丰硕的果实, 能够为我国统计学的普及和提高做出更大的贡献。

曾五一

2006 年 1 月于厦门



第三版前言

作为普通高等教育“十一五”国家级规划教材,《应用抽样技术》于2007年8月由科学出版社首次出版发行,于2010年8月再版,以其“内容全面、体系完整、条理清晰、简明扼要、易教易学”的特点而深受全国众多高校选用者的欢迎和肯定。为了更好地满足教学需要,现对第二版教材进行了适当的修订并补充了习题,作为第三版予以发行。

抽样技术(或者称为抽样调查、抽样法等)是本科统计学专业的核心课程之一,也是经济管理类专业的重要课程之一,目前开设这门课程的高校越来越多。然而,抽样技术课程由于内容丰富、概念繁多、数学基础要求较高而有点“令人生畏”,其教学难度显然比其他统计学课程大,因此学生普遍反映比较难学,有些教师也觉得难教。因此,如何编写一部易教易学的教材成了问题的关键。我们认为,一部好的教材一定要有清晰的架构,紧扣要点,层次分明,全而不乱。正是基于这样一个目的,我们按照“一个主题概念、两条基本主线、三个主要关系、四方面关键技术和五点补充内容”的架构来安排《应用抽样技术》的内容。

所谓一个主题概念就是“什么是抽样技术”,包括抽样技术和抽样调查的定义、分类、特点和作用等,重点是抽样调查与其他统计调查、概率抽样与非概率抽样的区别。所谓两条基本主线就是如何获取样本和如何估计总体,前者突出各种抽样方法和抽样方式的特点,后者强调估计量的性质和选择原则。所谓三个主要关系就是总体、样本与单位的关系,调查变量与辅助变量的关系,抽样精度与概率保证程度的关系,理清这三方面的关系对于科学设计抽样方案十分重要。所谓四方面关键技术就是抽样设计技术、误差分析技术、估计量构造技术和方差估计技术,它们构成了抽样技术的基础,也成为了教材的主要内容。所谓五点补充内容就是交叉子样技术、双重抽样技术、样本轮换技术、敏感性调查技术和非抽样误差分析,它们是对抽样技术的进一步拓展和深化。可以说,本教材的十一章内容比较好地体现了上述意图,通过有限的篇幅准确、清晰地阐述

了抽样技术的完整内容。

为了增加教材的趣味性和可读性，本教材每一章后面都对某一位世界著名的抽样专家作了简介，读者也可以从中更好地了解抽样技术的发展历程。

李金昌

2014 年 10 月



第一章

抽样技术概述	1
第一节 什么是抽样技术.....	1
第二节 抽样技术的产生与发展.....	7
第三节 抽样技术的应用	12
本章小结	14
思考与练习	15
著名抽样专家简介	17

第二章

抽样技术基本概念.....	18
第一节 总体与样本	18
第二节 估计量与抽样分布	23
第三节 抽样误差与置信区间	27
第四节 样本设计	31
本章小结	33
思考与练习	34
著名抽样专家简介	36

第三章

简单随机抽样	37
第一节 抽样方式	37
第二节 总体均值与总体总值的简单估计	42
第三节 总体比例的简单估计	47
第四节 样本量的确定	50
第五节 子总体估计	55
本章小结	58
思考与练习	59
著名抽样专家简介	61

第四章

分层抽样	62
第一节 抽样方式	62
第二节 简单估计量及其性质	64
第三节 样本量的分配	68
第四节 样本量的确定	72
第五节 分层抽样设计效果分析	77
第六节 进一步讨论的问题	78
本章小结	85
思考与练习	86
著名抽样专家简介	89

第五章

比率估计与回归估计	90
第一节 问题的提出	90
第二节 比率估计	91
第三节 回归估计	99
第四节 分层比率估计与分层回归估计	103
本章小结	109
思考与练习	110
著名抽样专家简介	112

第六章

不等概率抽样	113
第一节 问题的提出	113
第二节 放回不等概率抽样	115
第三节 不放回不等概率抽样	120
本章小结	132
思考与练习	132
著名抽样专家简介	134

第七章

整群抽样	135
第一节 抽样方式	135
第二节 群大小相等的整群抽样	137
第三节 群大小不等的整群抽样	142
第四节 估计总体比例的整群抽样	147
本章小结	150
思考与练习	150
著名抽样专家简介	153

第八章

系统抽样	154
第一节 抽样方式	154
第二节 等概率系统抽样的估计量及其方差	161
第三节 估计量方差的样本估计	164
第四节 进一步讨论的问题	167
本章小结	169
思考与练习	169
著名抽样专家简介	171

第九章

多阶段抽样	172
第一节 抽样方式	172

第二节 初级单位大小相等的二阶段抽样.....	174
第三节 初级单位大小不等的二阶段抽样.....	180
第四节 进一步讨论的问题.....	183
本章小结.....	185
思考与练习.....	186
著名抽样专家简介.....	187

第十章

其他抽样方法技术	188
第一节 样本轮换.....	188
第二节 双重抽样.....	192
第三节 随机化装置.....	195
第四节 交叉子样本.....	200
本章小结.....	202
思考与练习.....	203
著名抽样专家简介.....	205

第十一章

非抽样误差	206
第一节 非抽样误差构成.....	206
第二节 抽样框误差分析.....	208
第三节 无回答误差分析.....	218
第四节 计量误差分析.....	224
本章小结.....	229
思考与练习.....	229
著名抽样专家简介.....	231

主要参考文献.....	232
-------------	-----



第一章

抽样技术概述

本章教学目的与要求

本章对抽样技术的含义、作用、产生历史和实际应用等作简要介绍,为以后各章的学习奠定基础。具体要求:①正确理解抽样技术的科学含义、基本分类和特点,对抽样调查的基本程序和作用有初步的认识;②对抽样技术产生与发展的历史有一般的了解;③对抽样技术的实际应用有大致的认识。

第一节 什么是抽样技术

一、抽样技术的含义

众所周知,统计认识活动的一个重要环节是通过统计调查获取统计数据,然后再据以进行科学的统计分析。统计调查的技术有多种,其中应用最为广泛的技术就是抽样技术(sampling technique)。面对复杂的统计调查对象,抽样技术已广泛应用于社会、经济、科技、自然等各个领域,成为获取统计数据的最重要手段。抽样技术也早已成为现代统计学的重要组成部分,是20世纪人类最伟大的科学成就之一。

那么什么是抽样技术呢?最通俗的理解就是从统计调查总体(population)中抽取样本(sample)进行调查,获取数据,然后对总体数量特征作出推断的技术。很显然,抽样技术是一种非全面统计调查的技术。运用抽样技术所进行的调查称为抽样调查(survey sampling),从广义上说,一切非全面的统计调查都是抽样调查。例如,顾客买米时抓一把看看以判断是否干燥、饱满、洁白,学校通过召开部分同学的座谈会来了解教学情况,统计部门通过部分工业企业的产值资料来估计整个地区的工业增加值等,都属于抽样调查。在抽样调查中,抽样技术的运用主要有两个方面:抽取样本和估计总体。

根据样本抽取的方式不同,抽样可以分为两类:非概率抽样(nonprobability sampling)和概率抽样(probability sampling)。

(一) 非概率抽样

非概率抽样没有严格的定义,也称非随机抽样,是一种不按照随机原则、总体中各单位被抽中概率事先未知或难以确定的抽样,样本的抽取主要根据人们的主观判断或简便性原则来进行。非概率抽样主要有以下几种形式:

(1) 随意抽样 (haphazard sampling),也称任意抽样,即抽样者(通常是遵循简便性原则)随意地或任意地从总体中抽取样本。例如,实验人员从笼子里抓取最靠近笼门的小白鼠(而不是先对所有的小白鼠编号,然后随机抽取若干号码,再抓取相同编号的小白鼠)做试验,节目主持人从放在玻璃缸里的观众来信中信手抽取若干来信(而不是先对所有的观众来信编号,然后随机抽取若干号码,再挑出相同编号的来信)以选定获奖观众等,都属于随意抽样。利用已有的、但不完整的名册或号簿确定调查单位,在街口向过往行人进行调查,样本由自愿参加调查的人组成等,也属于这类情况。这种抽样的优点是简便易行、成本低,缺点是调查范围有时难以真正涵盖所有总体单位,样本的偏差有时会比较大会比较大,调查或推断的结论不具有一般意义。

(2) 判断抽样 (judgment sampling),也称有目的抽样 (purposive sampling) 或有代表性抽样 (representative sampling),即抽样者根据自己的知识、经验和判断从总体中挑选出“典型的”或“有代表性”的单位来组成样本。例如,从所在区域抽取几家“规模中等、生产经营比较稳定”的企业来了解其能源消费状况,从社区中抽取几户“收入中等”的住户来了解其对某种产品的购买意向等,都属于判断抽样。这种抽样的优点是可以发挥有关人员的主观能动性和利用已有的有关信息,缺点是对于什么是选择有代表性单位的最好办法或如何决定哪些单位最具有代表性,没有统一的客观标准。

(3) 定额抽样 (quota sampling),也称配额抽样,即抽样者按照规定的定额获得一个在某些特征上与总体结构大致成比例的样本。它是先对总体按一定标志分类,并按比例分配每类应调查单位的定额,然后由抽样者在每类进行判断抽样。这种抽样最先由美国盖洛普咨询公司发明使用,目的是增进判断抽样的样本代表性。由于这种抽样有一定的科学性,因而在市场调查、民意测验等方面有较广泛的应用,但判断抽样存在的缺点它也同样难以克服。

(4) 流动总体抽样 (sampling of mobile population),也称“捕获—标记—再捕获”(capture-tag-recapture) 抽样,即抽样者先从总体(如水库中的鱼)获取部分单位(如300尾鱼),加以标记后放回总体,过一段时间后再获取部分单位(如1000尾鱼),然后根据再获取单位中有标记单位的比例来推算总体的数量。这种抽样适用于事先对总体缺乏认识而调查单位又处于流动状态的研究对象。

如果运用得当,非概率抽样能发挥积极的作用。但非概率抽样有一个共同的致命缺点,那就是难以计算和控制抽样误差,难以保证推断的准确性和可靠性。

(二) 概率抽样

概率抽样也称随机抽样 (random sampling),是一种以概率论和随机原则为依据来抽取样本的抽样,是使总体中的每个单位都有一个事先已知的非零概率 (nonzero

probability) 被抽中的抽样。总体各单位被抽中的概率可以通过样本设计 (sample design) 来规定, 通过某种随机化 (randomization) 操作来实现。通常所说的、狭义上的抽样就是指概率抽样, 狭义上的抽样技术就是指概率抽样技术。常用的概率抽样方式有简单随机抽样 (simple random sampling)、分层抽样 (stratified sampling)、系统抽样 (systematic sampling)、整群抽样 (cluster sampling)、多阶段抽样 (multistage sampling), 将在以后各章分别介绍。

与非概率抽样相比, 概率抽样有以下一些特点:

(1) 样本的抽取遵循随机原则。所谓随机原则 (random principle) 就是样本的抽取排除了人的主观随意性或目的性, 调查对象总体中的每个单位都按照一定的、事先已知的概率被抽选, 也就是说总体中的任何一个单位都有机会被抽中。因此, 随机抽样的根本特点是总体中每个单位的入样概率事先已知或可以计算。它具有这样一些性质: 一是当把抽样方法应用于一个具体总体时, 能够被抽取的各个不同的样本集合是可以加以确定的。例如, 用简单随机抽样方式从由 A、B、C 三个单位组成的总体中抽取容量为 2 的样本, 三个可能的备选样本是 (A、B), (A、C) 和 (B、C); 二是每个可能被抽取的样本都被确定一个已知的被抽中的概率; 三是用一种随机的方法抽选所有可能备选样本中的一个时, 每个备选样本都有一个适当的被抽中的概率; 四是根据样本来估计总体时, 必须选择合适的估计量并加以说明, 对于任何一个具体的样本只能得到唯一的估计值。这里需要强调的是随机与随意 (或随便) 的区别, 随机有科学的含义, “随机”的结果可以用概率来描述, 而随意则更多地带有人的主观性, “随意”的结果难以用概率来表示, 两者的根本区别在于能否确保总体中的每个单位有事先可以计算和确定的非零概率被抽中, 这也是随机抽样 (概率抽样) 与非随机抽样 (非概率抽样) 的根本区别所在。例如在前述的抽取观众来信的例子中, 若事先对所有的观众来信进行编号, 然后通过随机化操作产生的号码来确定抽中的观众来信, 就属于随机抽样, 而由主持人直接从玻璃缸中抽取观众来信就属于随意抽样, 因为信封的大小、质感和在玻璃缸中的位置, 会影响到主持人的判断或偏好, 从而事实上不能使每封来信都有均等被抽中的机会, 每封信被抽中的概率也难以确定。

(2) 可以运用概率估计的方法对总体数量特征进行推断。抽样的目的是想通过样本提供的信息对总体数量特征 (或总体目标量) 进行估计或做出假设检验, 虽然样本的抽取是随机的, 样本是非唯一确定的, 以样本估计总体时并不存在确定的函数关系, 但样本与总体的关系可以通过抽样分布规律来描述。例如, 只要样本容量足够大, 样本均值就会充分接近总体均值, 样本均值的频率分布就会充分接近正态分布, 这就是通常所说的大数定律 (law of large numbers) 和中心极限定理 (central limit theorem)。因此, 我们可以根据样本与总体的这种内在联系关系, 运用概率估计的方法, 对总体数量特征做出具有一定概率保证程度的推断。例如, 我们经常以 95% 的概率保证对某市居民人均年收入的范围做出估计。当然, 这种推断是有风险的, 即有一定的可能性 (如 5% 的概率) 所做的推断是错误的。

(3) 抽样误差 (sampling error) 可以计算并加以控制。以样本统计值去估计总体数量特征, 不可能做到分毫不差, 必然会存在着估计误差, 即抽样误差, 这是由抽样调

查的非全面性和样本抽取的随机性所引起的。抽样误差的大小,反映了抽样调查效果的好坏,因此,我们总是希望能够知道抽样误差的大小并根据需要加以控制。概率抽样完全具备这一特点,因为运用同一方法反复从总体中抽取同样容量的样本,就能计算出由此产生的样本统计值的频率分布即抽样分布,就可以以该频率分布的方差或标准差作为衡量抽样误差大小的依据,就可以在以一定的概率保证估计总体数量特征的同时也可以控制抽样误差的范围。这一点是非概率抽样难以具备的,因为非概率抽样虽然也可以在一定条件下对总体进行估计,但样本与总体的内在关系难以描述,抽样误差当然也就难以计算和控制了。

还需要说明的是,概率抽样不同于等概率抽样(sampling with equal probability)。如前所述,概率抽样是指总体中的各单位都有非零概率被抽中,各单位被抽中的概率可以相等,也可以不相等。在实践中,绝大多数情况都以采用等概率抽样为主,当抽样单位大小不等时可能会采用不等概率抽样(sampling with unequal probability),要根据具体情况而定。

二、抽样调查的基本程序

抽样调查作为统计调查活动,有其一般的操作程序。一个完整的抽样调查过程,大致包括以下七个基本步骤:

(1) 设计抽样方案。这是抽样调查的开始阶段,也是影响抽样调查全局的重要阶段。抽样方案是抽样调查的指导纲领和总体思路,必须认真对待,精心设计。在抽样方案中,首先需要明确调查的背景、主题和目的,然后对调查对象总体与调查单位、调查内容、抽样方式、估计方法、调查时间、样本资料收集方法、样本容量、精度要求、概率保证程度、经费预算、人员安排、实施方法等逐一加以明确规定,需要对抽样调查的各个方面进行统一协调和总体考虑。

(2) 编制抽样框(sampling frame)和设计调查表(inquiry form)。编制抽样框和设计调查表(或问卷)(questionnaire)是抽样调查得以开展的两个必不可少的前提,是把抽样方案转化为具体实施的两项基础工作。一个完善的抽样框是保证抽样的随机性和对总体推断有效性的重要条件,并且能有效预防非抽样误差(nonsampling error)中的抽样框偏差,因此根据调查目的,把调查对象总体转化为一个合适的、完善的抽样框,需要做认真细致的工作。调查表(或问卷)则是把调查内容具体化的、用以收集个体资料的重要工具,一份科学的调查表(或问卷)能使被调查者正确理解调查的内容和要求,能够愿意并正确填写调查表或回答所提的问题,能够便于进行统计处理(计算机处理)和有效地利用调查资料进行统计分析。调查表(问卷)的设计既是科学也是艺术,需要有与调查内容有关的专业知识,以及统计学、社会学、心理学和计算机科学等知识,往往需要多方征求意见和反复修改。

(3) 试抽样调查。这是抽样方案设计后、正式抽样调查开始前的一个重要环节。试抽样调查就是根据抽样方案,先从调查对象总体(抽样框)中抽取一个小样本(通常比正式样本小很多)进行试调查,以便通过试抽样调查来发现各个方面可能存在的不完善之处,检验所设计抽样方案及其抽样框、调查表的科学性和可操作性,以免在正式调查