

调查培训 实用教材

— 上册 —

DIAOCHAFANGFA YU SHIWU

调查方法与实务

主编/于斌 主审/谢邦昌

 中国统计出版社
China Statistics Press

调查培训 实用教材

— 上册 —

DIAOCHAFANGFA YU SHIWU

调查方法与实务

主编/于斌 主审/谢邦昌

 中国统计出版社
China Statistics Press

(京)新登字 041 号

图书在版编目(CIP)数据

调查培训实用教材(上):调查方法与实务/于斌主编. —北京:
中国统计出版社, 2011. 8

ISBN 978-7-5037-6268-0

I. ①调… II. ①于… ②王… III. ①统计调查—在职教
育—教材 IV. ①C811

中国版本图书馆 CIP 数据核字(2011)第 143762 号

调查方法与实务

作 者/于 斌

责任编辑/陈悟朝

装帧设计/杨 超

出版发行/中国统计出版社

通信地址/北京市西城区月坛南街 57 号 邮政编码/100826

办公地址/北京市丰台区西三环南路甲 6 号

网 址/www.stats.gov.cn/tjshujia

电 话/邮购(010)63376907 书店(010)68783172

印 刷/河北天普润印刷厂

经 销/新华书店

开 本/710×1000mm 1/18

字 数/260 千字

印 张/19.125

印 数/1—3000 册

版 别/2011 年 8 月第 1 版

版 次/2011 年 8 月第 1 次印刷

书 号/ISBN 978-7-5037-6268-0/C. 2500

定 价/70.00 元(上、下册)

中国统计版图书,版权所有。侵权必究。

中国统计版图书,如有印装错误,本社发行部负责调换。

调查培训实用教材编委单位及编委会

编委单位 国家统计局统计教育培训中心

西安财经学院

编委会 (按姓氏笔画排序)

丁远忠 王立元 王佐仁 王国梁 王莉霞

毛有丰 田鲁生 杨学义 李纲 李萍

严建辉 张仲梁 罗兰 胡健 阎岭

编写人员 《上册·调查方法与实务》

主 编:于斌

编写人员:于斌 李林梅 张维群

主 审:谢邦昌

《下册·统计数据分析方法》

主 编:王佐仁

编写人员:张文红 朱钰 孙学英 李俊瑞

谢邦昌

主 审:谢邦昌

编写说明

在知识经济时代,培训——已经成为组织、机构,乃至个人不断提高素质,走向成功的一大法宝。

本教材是为国家统计局调查系统进行“调查技能与实务”培训而专门编写的。

教材上册根据调查的流程和步骤,按照调查方案设计、抽样设计、问卷设计、数据收集方法、调查质量控制等内容,结合调查实际案例进行了系统的介绍。教材下册对国内外目前常用的调查数据分析技术进行了介绍,由浅入深地分别介绍了数据描述、统计推断、多变量数据分析、数据挖掘等内容,并结合实际调查案例进行了系统的演示。

本教材由国家统计局统计教育培训中心与西安财经学院共同组织策划,最后由西安财经学院组织教师编写完成。

教材上册为《调查方法与实务》,主要介绍了与调查相关的知识与技能,具体包括:调查基本理论、基本概念、调查方案设计、抽样设计、问卷设计、数据收集方法、非概率调查方法、概率抽样方法、抽样估计的其它方法、非抽样误差、调查质量控制、中国的抽样调查实践等内容。第一、二、三、四、五、六章由于斌编写;第七、八、九、十四章由李林梅编写;第十、十一、十二、十三章由张维群编写。

教材下册为《统计数据分析方法》,主要介绍了数据描述、统计推

断、多变量数据分析、数据挖掘方法等内容。教材第一章由朱钰编写；第二章由孙学英、李佼瑞编写；第三章由张文红、孙学英编写；第四章由谢邦昌编写。

教材编写得到了国家统计局统计教育培训中心、国家统计局财务司、国家统计局城市司、国家统计局农村司、国家统计局陕西调查总队有关领导和专家的关心和指导；台湾辅仁大学谢邦昌教授亲自参加了教材编写并担任教材主审；西安财经学院王国梁教授、王佐仁教授亲自带领编写老师进行讨论。在此一并表示感谢！

教材编写人员虽然尽了很大努力，力求结合调查案例，介绍和反映调查的理论和方法。但由于水平有限，加之时间仓促，不足在所难免，敬请广大调查人士、专家、学者批评指正！

编者

2011年6月于西安

序

“十二五”时期是加快统计工作科学发展的关键时期,是深化统计改革、推进统计建设的攻坚阶段。“十二五”时期统计发展改革规划纲要明确指出,要紧紧围绕“提高统计能力、统计数据质量和政府统计公信力”这个中心,进一步推进统计工作规范统一、改革创新、公开透明,着力建设基本单位名录库、企业一套表制度、数据采集处理软件、联网直报等四大工程,加快推进统计现代化,为“十二五”时期的统计工作指明了方向。

建设四大工程是统计调查流程的系统再造,是统计数据生产方式的深刻变革。这一深刻变革必然带来统计管理模式、工作机制的改革,必然要求统计组织机构、人员结构、特别是统计人员的知识结构与之相适应。这一深刻变革必然对各级政府统计人员和企事业单位统计人员的业务素质提出新的更高的要求。马建堂局长在全国统计工作会议上指出“高素质的人才 是统计部门的第一生产力,是统计事业科学发展最可宝贵的财富”,把人才队伍建设的重要性提升到了前所未有的高度。正因为如此,国家统计局始终把统计培训工作作为事关统计科学发展的一项重要基础性工作,给予高度重视和支持。

我们高兴地看到,国家统计局统计教育培训中心和西安财经学院共同合作,在多年培训工作的基础上,形成了一套有效针对统计调查人员在职培训的教材。这套教材集理论性、实务性、操作性于一体,专门针对政府统计调查系统的在职人员工作需要进行了精心设

计。教材分上下两册,内容涵盖了统计调查设计(整体方案设计、问卷设计、抽样设计、数据挖掘设计)、统计调查方法、抽样技术、数据质量评估与管理、数据挖掘分析方法等全过程。教材结合我国统计调查实际案例进行分析和展示,可以说,是对我国过去、现在统计抽样调查、统计分析工作经验的总结,教材还对我国统计调查未来的发展做了前瞻性的预研,对我国统计调查实际工作将起到重要的促进作用。

我衷心地希望各级政府统计工作者认真分析、深入研究这本教材,从中得到有益的启示,进而将我国统计调查提高到一个新的水平。

国家统计局

目 录

第一章 调查基本理论	(1)
第一节 统计与调查	(1)
第二节 调查分类及其特点	(5)
第三节 抽样的科学性	(13)
第四节 抽样理论及其发展	(21)
第二章 基本概念	(24)
第一节 总体与样本	(24)
第二节 抽样框与抽样单元	(27)
第三节 参数与统计量	(29)
第四节 抽样误差与非抽样误差	(31)
第五节 精度与费用	(35)
第三章 调查方案设计	(38)
第一节 调查流程	(39)
第二节 调查方案设计	(43)
第三节 调查方案的完善	(47)

第四章 抽样设计	(55)
第一节 抽样的流程	(55)
第二节 抽样方案设计	(57)
第三节 抽样中的几个问题	(60)
第四节 抽样方案的完善	(66)
第五章 问卷设计	(83)
第一节 问卷设计的程序	(83)
第二节 问卷设计的技巧	(85)
第三节 问卷中问句态度的度量	(92)
第四节 问卷示例	(94)
第六章 数据搜集方法	(112)
第一节 面访调查法	(113)
第二节 其他调查方法介绍	(121)
第七章 非概率调查方法	(126)
第一节 重点调查与典型调查	(126)
第二节 网络调查	(128)
第三节 非概率抽样	(131)
第八章 概率抽样方法(一)	(137)
第一节 简单随机抽样	(137)
第二节 分层抽样	(146)

第九章 概率抽样方法(二)	(156)
第一节 系统抽样	(156)
第二节 整群抽样	(161)
第三节 多阶段抽样	(167)
第十章 概率抽样方法(三)	(174)
第一节 不等概率抽样	(174)
第二节 二重抽样	(185)
第三节 其他抽样	(193)
第十一章 抽样估计的其他方法	(198)
第一节 辅助变量的应用	(198)
第二节 比估计、回归估计与差值估计	(203)
第三节 其他估计方法	(216)
第十二章 非抽样误差	(220)
第一节 抽样框误差	(220)
第二节 无回答误差	(227)
第三节 计量误差	(236)
第十三章 调查质量控制	(241)
第一节 引言	(241)
第二节 设计阶段控制	(244)
第三节 调查过程控制	(247)
第四节 数据处理过程控制	(251)

第五节 数据质量评估	(253)
第十四章 中国的抽样调查实践	(261)
第一节 住户调查	(261)
第二节 农产量调查	(276)
第三节 景气调查	(282)
参考文献	(292)

第一章 调查基本理论

统计是关于数据搜集、整理、分析,并得出研究结论的过程。在统计工作中,统计调查有着基础性的地位和作用。在规模数以百万计、千万计、亿万计的研究对象中,只抽取几千个、甚至于只有几百个单元,用以推断前者,其结论可靠吗?本章就回答这些疑问。

第一节 统计与调查

一、统计与调查

1. 什么是统计

统计是关于数据的一门学科。人们对统计学下的定义很多,比较规范的定义有:“统计学是收集、整理、分析、表述和解释数据的科学”;“统计学是一系列方法,用以设计、获取数据、在数据基础上概括、展示、解释、分析,并得到结论的科学”。

总体上说,不管如何定义统计学,统计学总是离不开数据,统计是用数据表达其思想和工作的成果的。因此我们认为,统计学是关于数据收集、处理、分析研究、解释数据并从中得出结论的科学。它提供的是数据收集方法、有效处理方法、分析解释方法并从数据中得出正确结论的方法的思想体系和方法体系。

统计学不仅仅是一门科学,它也是一种思维方式。它要求我们以数据、以情况为依据去思维,任何时候、任何条件下都要以数字为准绳,实事求是。

统计学不仅是一种科学的思维方式,它更是一种工作方法。任何人只有掌握了统计科学,掌握了统计学的思维,掌握了统计学的方法,他就多了比常人更有效的表达思想的方法和工具——不仅可以用语言表达思想,而且可以用数字、用图表表达思想。哲学家培根曾经说过:“历史使人明智,诗歌使人巧慧,数学使人精细,博物使人深沉,伦理使人庄重,逻辑与修辞使人善辩。”我国著名统计学家邱东说“统计可以使人豁达”。

2. 什么是调查

什么是调查?所谓调,就是计算、算度的意思。《汉书·晁错传》里说:“调立城邑,毋下千家。”颜师古注曰:“调,谓算度之也。”“查”指寻检、查究、查核、考查之意。显然,调查是一种有目的的、有意识的自觉认识活动。

人们总是借助于调查,实现对所调查对象的考察、了解、分析和研究,达到对其发展变化规律的认识。它包含了以下几层含义:

(1)调查是一种自觉的有目的的认识活动。任何调查都是有目的、有意识进行的,是一种自觉的认识活动。它与人们日常生活中对现象的一般观察和了解有严格的区别。没有调查就没有发言权,在某种程度上说明了调查的重要性和必要性。

(2)调查需要科学的方法。没有科学的调查方法,即使进行了调查,也不可能得出正确的结论和认识。从这一点上说,缺乏科学的调查方法,有了调查也未必就有发言权。

(3)调查能力是现代人必须具备的一种能力和素养。在现代社会条件下,无论是在经济生活、社会生活等各个方面,都要求社会成员具备基本的调查分析的能力。缺乏调查分析能力,就只能人云亦云,失去独立的判断。

(4)调查是获取信息、数据的最重要、最基本的方式。调查获取的信息可以表现为定性信息,也可以表现为定量信息;可以是文字信息,也可以是数据信息。调查的过程也表现为从定性认识出发、通过定量、数据信息的收集、整理、分析,最终上升为定性的结果。可以说,调查过程完全的体现了认识论和方法论的统一。

调查是认识客观世界的手段,要正确认识社会并对其进行有效管理,必须掌握真实的、可靠的、准确的、完整的信息。调查作为收集、处理、分析研究客观世界信息的方法,对于了解客观世界现状、探索未来发展方向具有重要的意义。

3. 调查与统计的关系

调查作为一门科学,它与统计学、统计工作有着密切的联系。统计实践告诉我们,一个完整的统计工作过程大致可分为以下四个阶段,即设计、调查、整理、分析。调查是统计工作的基础性环节,没有统计调查,没有统计调查获取的基本数据,就没有统计加工、统计整理、统计分析,统计工作也就无法开展,更无法得出正确的结论,所谓“皮之不存,毛将焉附”就是这个道理。因此,从这一点上讲,调查是统计科学、统计工作的基础,调查能力是统计人员的基本功。

从另一个方面看,统计学又为调查科学、调查实践提供正确的方法。在调查科学中、在调查实践中所需要的方法,都是统计学方法在相应学科的具体体现。如社会调查,就是统计学方法在社会调查领域的具体应用。因此,调查活动、调查科学没有统计学所提供的方法,也无法获取调查数据。

二、数据

从某种意义上讲,调查的过程就是收集数据的过程。数据反映了客观对象的数量特征,由于数据的来源不同,我们将其划分为两大类:调查数据和实验数据。

1. 调查数据

调查数据是指通过向被调查对象进行观察或者询问所获取的数据。它可能是显在的,也可能是潜在,但都需要调查组织者进行观察或者询问才能获得。如一个国家或者地区的国内生产总值,就是这个国家和地区在一定的时间范围内、地域范围内所生产出的全部最终产品和劳务的总和。只要在这个时间范围、空间范围,只要是属于这个指标核算的内容范围,我们都需要加以核算统计。显然,这些数据在客观上是存在的、显在的,只需要我们去观察、询问就可以获得。又比如顾客或者社会公众对某些问题的态度或意向,也是客观存在的、但未必是显在的,很可能是潜在的,需要调查人员采用一定的技巧和方法才能获得数据。

调查数据大多情况下与时间有关,数据展示的都是一定时间状态下的特

征。时间不同,即使同一研究对象,其数据就会发生变化,因此,调查数据需要不断更新,才能反映研究对象的最新发展。因此,按照不同的时间周期,对有些调查对象、调查现象进行研究,可以反映其发展变化的趋势和规律。

调查数据可以分为:定类数据、定序数据、数值型数据(定距数据、定比数据)。

定类数据用于区分被调查对象中每个个体的类型特征。如性别、籍贯、宗教等。

定序数据用于区别同一类别个体等级次序。定序数据能决定次序,即按数据的值把研究对象排列高低或大小。

数值型数据用于反映单元或个体之间数值差异。所有的定量数据都属于数值型数据,但数值型数据也包含了上述两类数据的变化。

(1)真值

调查从本质上讲,就是一种测量活动。调查的最高境界就是获取到了被调查对象的真实值。真实值也称真值,它是指被调查对象在一定的条件下的量值。从理论上说,除非碰巧,否则,被调查对象的数值型数据其真实值是不可能得到的。调查人员都试图获得调查对象的真值,但是,它是一个理论上存在的量,是我们永远不可能知道的值。因为我们如果知道真值,也就用不着调查了。事实上,即使我们获得了真值,我们也不知道它是真值!这不能不说是一件非常遗憾的事情。

(2)测量值

测量值也就是调查值。是调查人员从被调查对象那里所获得的调查结果。在调查实践中,我们往往能获得的就是测量值了。我们总是以获得的测量值为基础,再加上可能的误差,对调查对象进行推算和认识。

凡是有测量活动或调查活动,都必然会产生误差。由于真值是未知的,因此,在调查研究中,我们不可能计算实际的误差。在调查理论中,我们往往是计算理论的或者叫平均误差,以平均误差为基础对真值进行推算。

2. 试验数据

试验数据是指通过试验研究所获得的反映试验对象特征的数据,它通常与自然科学的研究相联系,但近年来,其研究方法在社会科学、经济科学中也得到

了广泛的应用。

试验数据在试验前不可能获取,只有在进行试验,通过观察或记录才可能获取。试验调查一般都需要事前进行试验设计,在试验控制条件下进行。一般来讲,试验调查过程是可重复的,如化肥增产效果测试、不同商品包装销售效果试验、药品疗效测试等。由于试验条件不同,试验数据就可能不同,因此,试验调查中对实验条件的控制是关键。

需要说明的是,作为统计科学来讲,主要研究的是调查数据,而不是实验数据。因此,本书所进行的研究,是以调查数据为对象进行的。

第二节 调查分类及其特点

对调查可以采用多个标志进行不同的分类,但是最基本的是按照调查范围进行的分类。

全面调查,就是对调查对象中的所有个体逐一进行的调查,最典型的就是普查。全面调查具有全面性、周期长、费用高、动用调查人员多等特点。

非全面调查,就是对调查对象中的部分个体进行的调查,包括典型调查、重点调查、抽样调查等,最常采用的是抽样调查。非全面调查具有经济、快捷等特点,对某些调查具有独特的优势,如:无限总体的研究、破坏性实验等。

抽样调查有狭义和广义之分,广义的抽样调查包括概率抽样和非概率抽样;狭义的抽样调查仅仅指概率抽样。在统计学研究中,讲到抽样调查,大多都指的是狭义的抽样调查。

一、全面调查

全面调查就是对调查对象中的全部个体进行的调查。在我国,全面调查主要表现为普查,统计报表制度也具有全面调查的特性。在调查实践中,除非统计部门,社会其他部门都无力承担起全面调查所需要的人力、物力、财力的压力。

(一) 普查

普查是专门组织的一次性调查,一般用来调查属于一定时点上调查对象各个方面的特征,它属于全面调查。说是一次性,事实上都是按照一定的时间间