

<http://www.phei.com.cn>

SPSS

在社会调查中的应用

◎ 杜智敏 樊文强 编著

◎ 郭宜斌 审校



SPSS

统计 分析 教材



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

SPSS

在社会调查中的应用

◎ 杜智敏 樊文强 编著
◎ 郭宜斌 审校



统计分析教材



電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书是为读者在开展社会调查过程中正确使用 SPSS 而写的,也可以作为《社会调查方法与实践》的下册。SPSS 作为社会统计分析的强有力工具,在社会调查中主要应用于抽样调查,因此本书不以 SPSS 的功能模块为序,不单纯介绍具体操作,而是立足于初学者,结合具体的调查案例,将统计学的基本知识与 SPSS 的运用融为一体,以抽样调查过程中对问卷的统计分析 workflow 为主线而展开。全书共 11 章,第 1 章概述抽样调查的全过程;第 2 章介绍在问卷回收之后,如何进行数据的净化、编码、数据文件的建立,以及在分析之前所做的统计预处理;第 3、4 章介绍如何通过统计表与统计图对样本数据进行频数分析和估计总体的分布特征;第 5~9 章介绍根据不同的变量类型如何进行不同群体差异的比较,怎样分析调查项目(即变量)之间的相关关系和不确定性因果关系;第 10 章说明对调查对象的分类;第 11 章则讲明问卷的信度与效度分析。读者可登录华信教育资源网 www.hxedu.com.cn 下载本书配套数据文件及电子教案。

本书既可以作为大学生、研究生的教材和教师的教学参考书,也可以作为实际工作者开展调查研究时的指导手册。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

SPSS 在社会调查中的应用/杜智敏,樊文强编著. —北京:电子工业出版社,2015.1

统计分析教材

ISBN 978-7-121-25016-3

I. ①S… II. ①杜… ②樊… III. ①社会调查-统计分析-软件包-高等学校-教材 IV. ①C915-39

中国版本图书馆 CIP 数据核字(2014)第 282012 号

策划编辑:秦淑灵

责任编辑:郝黎明

印 刷:涿州市京南印刷厂

装 订:涿州市京南印刷厂

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:787×1092 1/16 印张:28.25 字数:723.2 千字

版 次:2015 年 1 月第 1 版

印 次:2015 年 1 月第 1 次印刷

印 数:3000 册 定价:59.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zlt@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

前 言

随着抽样理论、统计分析技术、计算机技术和统计软件的发展,抽样调查作为社会调查的一种重要方法及获取统计资料的重要手段,日益受到政府各部门、企业、学术界与社会公众的重视,在学术研究、行政管理、民意调查和市场调查等领域,得到了广泛的应用。对调查问卷进行统计分析是抽样调查研究的关键环节,是保证调查研究质量的重要基础。鉴于 SPSS(Statistical Product and Service Solutions,统计产品和服务解决方案)具有界面友好、统计功能强大、易学易用等优点,越来越多的调查研究者希望掌握或正在使用 SPSS,将其作为工具对调查数据作深度挖掘。但是,随之而来的也出现了某些乱用统计方法和软件的现象,貌似用数据说话、貌似分析严谨,但实则错误百出,其危害比不用统计分析方法更为严重。因此,对于抽样调查,目前人们最需要解决的问题是对调查数据如何进行深入的、科学的分析,具体地说,在使用 SPSS 时需要能够正确解决下面 4 个问题:

第一,针对调查研究的课题,需要用哪些方法来进行统计分析?

第二,结合调查数据的类型和条件,能不能用所选择的统计分析方法?

第三,会不会操作 SPSS?

第四,对给出的各种统计图表,能否看得懂?能得出哪些结论?这些结论在实际中的意义是什么?

显然,要达到这样的目的,只有“用到哪里学哪里”,读者才能真正地体会到如何用 SPSS 来完成自己的统计分析工作。正是从这一需求和现实出发,本书没有以 SPSS 软件的模块体系为顺序展开,而是以对一份具体的调查问卷进行统计分析的过程为主线,随着对调查数据分析的不断深入,对所涉及的统计学概念、理论和 SPSS 中的相关功能做出详略有度的介绍。

本书共分 11 章。第 1 章概述抽样调查的全过程,抽样调查是一个完整的过程,没有好的研究设计方案、高质量的问卷、科学的抽样方法和数据采集,统计分析工作将变得毫无意义,因此需要读者从总体上把握好抽样调查;第 2~10 章按照对问卷统计分析的工作流程展开,即第 2 章的内容是在采集数据之后,如何进行数据的净化、编码、数据文件的建立以及在分析之前需要做哪些统计预处理;第 3、4 章介绍如何通过统计表与统计图对样本数据进行频数分析,如何估计总体的分布特征;第 5~9 章介绍根据不同的变量类型如何进行不同群体差异的比较,怎样分析调查项目(即变量)之间的相关关系和不确定性因果关系;第 10 章说明对调查对象如何进行分类;第 11 章则讲明对问卷的信度与效度分析,作为结构效度分析的基础,对主成分分析与因子分析给予了比较详细的介绍。

本书自始至终主要用的案例是对北京市大学生的学情调查。以一个实际的案例一以贯之,是希望读者能够看到,调查数据是一个富矿,只要一步一步深入地“挖掘”,就可以从杂乱无章的数据中看到事物的本质特征、看到背后的规律;还希望通过这个案例,使读者了解到对一份问卷的分析可能包括哪些方面的内容,体验对一个问题分析的不同视角和方法。事实上,数据本身的背景并不十分重要,重要的是要知道对于不同的数据类型如何做分析以及如何解释分析的结果。因此,本书没有将案例的作用停留在 SPSS 的操作与解释每个统计表

的行、列标题是什么上,而是兼顾理论与实用,针对研究的问题,从审核是否满足所选统计方法的条件、数据文件的格式到实际操作,从图表所表述的统计意义到反映在研究问题上的实际意义,都给予了尽可能详尽的说明,使本书具有较强的可操作性和实践性,有益于读者将 SPSS 的应用延伸到更为宽广的领域。当然,要真正做到这一点,还需要读者在阅读本书的过程中,重视概念的掌握,在用中加深理解。不要满足于对 SPSS 的简单操作上,要重视理解统计学概念、原理的内涵以及适用的条件;不要停留在“看”上,要带着课题读,边读边做,必须经过自己的实践,才能掌握统计分析方法与 SPSS 的真谛。

本书既可以作为实际工作者开展调查研究时的指导手册,也可以作为大学生的教材、教师的教学参考书。为适应不同读者的需要和正确理解 SPSS 的功能,作者采取了两项措施:一是在介绍 SPSS19.0 中文版的同时,也将相应的英文版加以标注。二是针对不同的读者群具备的基础有所不同,本书定位在为初学者的“用”而写。阅读本书时,不需要读者具备微积分、线性代数、概率论等相关的数学基础,不需要读者熟练地掌握计算机的操作,本书是自包含的,对所涉及的每个概念都尽可能用易于理解的方式加以阐明,对 SPSS 相关菜单尽可能给予全面的介绍。另外,除第 1、9、10 章,其他各章都给出了有关统计分析方法与应用 SPSS 的附表,以利于读者对每章内容的梳理与掌握。

本书是在原杜智敏编著的《抽样调查与 SPSS 应用》基础上进行改编的,杜智敏对全书的结构进行了总体再设计,樊文强撰写了第 9 章以及第 8、10、11 章中从 SPSS16.0 改版为 SPSS19.0 的工作,杜智敏撰写了第 1~7 章以及第 8、10、11 章的其他部分。樊文强与杜智敏从专业视角相互进行了审校与修改。

在本书的写作过程中,香港中文大学副校长、教育学院教育心理学讲座教授、国际应用心理学会教育心理部主席侯杰泰先生对原作给予了殊多的鼓励,并在应完善的内容方面提出了很中肯的建议;陈淑敏作为第一读者对第 8~11 章的修改提出了自己的看法;作者还拜读和参阅过许多专家学者的专著和论文,受益匪浅。本书完稿后,郭宜斌研究员对全书的文字仔细进行了最后的审校与修改。电子工业出版社秦淑灵编辑为本书的出版付出了许多心血,不但关注本书的进展还给予过多次指导。兹借本书出版之机,对以上所提各位深表谢忱。

编 者

目 录

第 1 章 抽样调查与 SPSS 概述	(1)
1.1 抽样调查概述	(1)
1.1.1 抽样调查的概念与特点	(1)
1.1.2 抽样调查的过程	(2)
1.1.3 对抽样调查的评价	(2)
1.2 调查问卷的一般问题	(3)
1.2.1 问卷的结构	(3)
1.2.2 问卷的类型	(4)
1.2.3 编制问卷的过程	(5)
1.3 测量与封闭式题目的类型	(7)
1.3.1 变量的测量水平	(7)
1.3.2 封闭式题目的类型	(10)
1.3.3 李克特量表	(12)
1.4 对问卷统计分析的基本内容	(13)
1.4.1 以正确的观念指导统计分析	(13)
1.4.2 选择统计分析内容与方法的依据	(13)
1.4.3 统计分析的主要内容	(14)
1.5 SPSS 及其在抽样调查中的应用	(16)
1.5.1 SPSS 公司与 SPSS 统计软件包	(16)
1.5.2 SPSS 的安装、启动与退出	(17)
1.5.3 SPSS 的运行方式	(18)
1.5.4 SPSS 的操作环境	(19)
1.5.5 对话框	(21)
1.5.6 中英文版本的转换与变量列表	(22)
1.5.7 SPSS 在抽样调查中的应用	(23)
附录 北京市大学生学情调查问卷	(24)
第 2 章 调查数据的预处理	(29)
2.1 对答卷的审核与编码	(29)
2.1.1 对答卷质量的审核	(29)
2.1.2 对问卷进行编码	(30)
2.2 建立 SPSS 格式的数据文件	(35)
2.2.1 利用数据编辑器窗口建立数据文件	(35)
2.2.2 Excel 格式数据文件的转换	(41)

2.2.3 数据文件的合并	(42)
2.3 数据的净化	(49)
2.3.1 利用“探索(Explore)”清理极端值	(49)
2.3.2 利用“交叉表(Crosstabs)”检查互斥数据	(53)
2.3.3 重复个案的排查	(54)
2.3.4 答卷录入质量的检查	(57)
2.4 数据文件的整理	(58)
2.4.1 缺失值的处理	(58)
2.4.2 逆向题目的重新计分	(64)
2.4.3 选取数据子集	(66)
2.4.4 数据文件的拆分	(70)
2.4.5 数据文件行与列的转置	(72)
2.5 在数据文件中生成新变量	(73)
2.5.1 定类变量的计数	(74)
2.5.2 定序变量的综合指标	(77)
2.5.3 定量变量转化为定性变量	(77)
2.6 对个案加权	(82)
2.6.1 何时需要对个案加权	(82)
2.6.2 利用“加权个案(Weight Cases)”进行加权	(83)
2.6.3 对个案加权应注意的问题	(84)
附表	(84)
第3章 调查数据的分布特征	(86)
3.1 一个单选题的统计表与统计图——单变量的频数分析	(86)
3.1.1 频数分布表	(86)
3.1.2 常用的统计图	(89)
3.2 一个单选题的数据分布特征——单变量的特征量数	(92)
3.2.1 数据的集中趋势	(92)
3.2.2 数据的离中趋势	(98)
3.2.3 偏度与峰度	(102)
3.2.4 参数估计	(104)
3.2.5 相对量数	(108)
3.3 利用 SPSS 对一个单选题的统计分析	(109)
3.3.1 利用“频率(Frequencies)”作统计分析	(109)
3.3.2 利用“描述(Descriptives)”作数据特征分析	(114)
3.3.3 利用“探索(Explore)”作数据特征分析	(115)
3.3.4 利用“探索(Explore)”求总体比例的置信区间	(117)
3.4 多个单选题交叉分组下的频数分析——多变量的交互分析	(118)
3.4.1 交叉表	(118)
3.4.2 常用统计图	(120)

3.5 利用 SPSS 对多个单选题作交互分析	(123)
3.5.1 利用“交叉表(Crosstabs)”对多变量频数作交互分析	(123)
3.5.2 利用“探索(Explore)”计算分组数据的特征量数	(126)
3.5.3 利用“均值(Means)”计算分组数据的特征量数	(128)
3.6 利用 SPSS 做多项选择题的频数分析——多响应变量分析	(129)
3.6.1 多响应变量分析的提出	(129)
3.6.2 SPSS 中多响应变量分析的功能	(129)
3.6.3 利用“多重响应(Multiple Response)”做多项选择题的频数分析	(130)
3.7 利用“比率(Ratio)”做比率分析	(133)
3.7.1 “比率(Ratio)”的结构与功能	(133)
3.7.2 操作步骤	(135)
3.7.3 输出结果及其解释	(135)
附表	(136)
第 4 章 统计图的制作与编辑	(138)
4.1 复式条形图的绘制	(138)
4.1.1 “条形图(Bar Charts)”的功能与结构	(138)
4.1.2 “个案组摘要”模式下的条形图	(139)
4.1.3 “各个变量的摘要”模式下的条形图	(144)
4.1.4 “个案值”模式下的条形图	(146)
4.2 线图	(147)
4.2.1 “线图(Line Charts)”的功能与结构	(147)
4.2.2 “个案组摘要”模式下的线图	(148)
4.2.3 “各个变量的摘要”模式下的线图	(149)
4.2.4 “个案值”模式下的线图	(150)
4.3 人口金字塔图	(151)
4.3.1 “人口金字塔(Population Pyramid)”的功能与结构	(151)
4.3.2 绘制金字塔图的操作步骤	(152)
4.3.3 绘制金字塔图的几点说明	(152)
4.4 统计图的编辑	(153)
4.4.1 图形编辑窗口概述	(153)
4.4.2 对条形图的编辑	(156)
4.4.3 对其他图形的编辑	(163)
4.5 作图与读图	(166)
4.5.1 掌握制作统计图的基本原则	(167)
4.5.2 学会审图,谨防统计图中的“陷阱”	(168)
4.5.3 学会读图,抓住重点深入思考	(169)
附表	(170)
第 5 章 正态总体均值的差异检验——不同群体差异的比较之一	(171)
5.1 假设检验概述	(171)

5.1.1	假设检验的思路	(172)
5.1.2	假设检验的一般步骤	(174)
5.1.3	关于假设检验的几点说明	(175)
5.1.4	利用 SPSS 进行假设检验的步骤	(177)
5.2	统计检验的前期工作——对数据分布特征的检验	(178)
5.2.1	利用“探索：图(Explore: Plots)”考察数据特征	(178)
5.2.2	利用“单样本 K-S 检验(1-sample K-S)”检验考察数据分布	(185)
5.3	单个正态总体均值的检验——单个群体与其总体均值差异的比较	(187)
5.3.1	单样本 T 检验概述	(187)
5.3.2	“单样本 T 检验(One-Samples T Test)”的操作步骤	(188)
5.3.3	输出结果及其解释	(189)
5.4	两个独立正态总体差异的检验——两个群体差异的比较之一	(189)
5.4.1	使用两个独立样本 t 检验的条件及思路	(190)
5.4.2	利用“独立样本 T 检验(Independent-Samples T Test)”进行 t 检验	(190)
5.5	两个配对正态总体差异的显著性检验——两个群体差异的比较之二	(193)
5.5.1	使用配对样本 t 检验的前提条件与思路	(193)
5.5.2	利用“配对样本 T 检验(Paired-Samples T Test)”进行 t 检验	(194)
5.6	单因素方差分析——多个群体差异的比较	(196)
5.6.1	单因素方差分析概述	(196)
5.6.2	利用“单因素 ANOVA(One-Way ANOVA)”进行检验	(200)
附表		(209)
第 6 章 非正态总体的差异检验——不同群体差异比较之二		(210)
6.1	两个独立样本的非参数检验	(210)
6.1.1	非参数检验概述	(210)
6.1.2	SPSS 提供的四种检验方法	(210)
6.1.3	利用“两个独立样本(2 Independent-Samples)”进行差异检验	(213)
6.2	两个相关样本差异的非参数检验	(217)
6.2.1	SPSS 提供的四种检验方法之比较	(218)
6.2.2	利用“两个相关样本(2 Related-Samples)”进行差异检验	(219)
6.3	多个独立样本的非参数检验	(221)
6.3.1	使用多个独立样本的非参数检验的前提条件	(221)
6.3.2	SPSS 提供的三种检验方法	(221)
6.3.3	利用“K 个独立样本(K Independent Samples)”进行检验	(224)
6.4	多个相关样本的非参数检验	(225)
6.4.1	使用多个相关样本的非参数检验的前提条件	(226)
6.4.2	三种非参数检验方法的思路	(226)
6.4.3	利用“K 个相关样本(K Related Samples)”进行检验	(228)
6.5	对比例的一致性检验	(230)
6.5.1	单个总体比例的检验	(231)

6.5.2 多个群体比例差异的比较	(236)
附表	(244)
第7章 事物间的相关关系	(247)
7.1 相关关系概述	(247)
7.1.1 函数关系与相关关系	(247)
7.1.2 散点图	(248)
7.1.3 相关系数	(251)
7.2 两个定性变量的相关分析	(253)
7.2.1 “分析(Analyze)”中有关相关分析的菜单	(253)
7.2.2 利用“交叉表(Crosstabs)”进行 χ^2 独立性检验	(255)
7.2.3 两个定类变量间的相关系数	(257)
7.2.4 两个定序变量间的相关系数	(260)
7.3 定量变量的相关分析	(267)
7.3.1 两个定量变量的相关分析	(267)
7.3.2 定类变量与定量变量的相关分析	(274)
7.4 两个事物之间关系的进一步分析	(278)
7.4.1 详析分析的提出	(278)
7.4.2 利用 SPSS 做详析分析	(281)
7.5 单变量多因素方差分析	(287)
7.5.1 多因素方差分析概述	(287)
7.5.2 “单变量(Univariate)”的功能与结构	(289)
7.5.3 利用“单变量(Univariate)”进行单变量多因素方差分析	(294)
7.5.4 应用方差分析过程中的几点说明	(299)
附表	(303)
第8章 线性回归与曲线回归——事物间的非确定性因果关系之一	(305)
8.1 一元线性回归分析	(305)
8.1.1 回归分析概述	(305)
8.1.2 一元线性回归方程的建立	(306)
8.2 多元线性回归分析	(316)
8.2.1 一元与多元线性回归模型的比较	(316)
8.2.2 多重共线性的诊断	(318)
8.2.3 奇异值与影响点的诊断与处理	(320)
8.2.4 应用线性回归方程过程中的若干问题	(322)
8.3 利用“线性回归(Linear Regression)”进行线性回归分析	(323)
8.3.1 “线性(Linear)”的结构与功能	(323)
8.3.2 利用“线性(Linear)”进行线性回归分析	(328)
8.4 曲线估计	(338)
8.4.1 非线性关系的线性化	(338)
8.4.2 “曲线估计(Curve Estimation)”的功能与结构	(339)

8.4.3	利用“曲线估计(Curve Estimation)”进行曲线估计	(341)
8.4.4	应用曲线估计过程中的若干问题	(343)
附表		(345)
第9章 Logistic 回归分析——事物间的非确定性因果关系之二		(346)
9.1	Logistic 回归分析概述	(346)
9.1.1	Logistic 回归分析的提出	(346)
9.1.2	Logistic 回归的基本思路	(347)
9.1.3	Logistic 回归方程中的虚拟变量	(347)
9.1.4	Logistic 回归方程中系数的直观解释	(348)
9.1.5	Logistic 回归方程的检验	(349)
9.2	二项 Logistic 回归	(350)
9.2.1	二项 Logistic 回归分析的适用范围与步骤	(350)
9.2.2	“二项 Logistic 回归分析(Binary Logistic)”的功能与结构	(351)
9.2.3	“二项 Logistic 回归分析(Binary Logistic)”的应用	(355)
9.3	多项 Logistic 回归分析	(359)
9.3.1	多项 Logistic 回归分析模型	(359)
9.3.2	“多项 Logistic 回归分析(Multinomial Logistic)”的功能与结构	(360)
9.3.3	“多项 Logistic 回归分析(Multinomial Logistic)”的应用	(363)
9.4	多项有序回归分析	(367)
9.4.1	多项有序回归分析的功能与结构	(367)
9.4.2	多项有序回归分析的应用	(369)
第10章 对调查对象的分类		(373)
10.1	距离与相似性度量	(373)
10.1.1	聚类分析概述	(373)
10.1.2	聚类分析中对“亲疏程度”的测量	(375)
10.1.3	进行“亲疏程度”度量时应注意的问题	(377)
10.2	系统聚类	(378)
10.2.1	使用系统聚类分析的条件与步骤	(378)
10.2.2	“系统聚类(Hierarchical Cluster)”的功能与结构	(380)
10.2.3	利用“系统聚类(Hierarchical Cluster)”进行分析聚类	(384)
10.3	K-均值聚类	(389)
10.3.1	使用 K-均值聚类的条件与步骤	(389)
10.3.2	“K-均值聚类(K-Means Cluster)”的结构与功能	(390)
10.3.3	利用“K-均值聚类(K-Means Cluster)”进行聚类分析	(393)
第11章 问卷的质量分析		(398)
11.1	问卷的项目分析	(398)
11.1.1	项目分析的基本方法	(398)
11.1.2	利用 SPSS 进行项目分析	(399)
11.2	问卷的信度分析	(400)

11.2.1 对信度的估计	(400)
11.2.2 “可靠性分析(Reliability Analysis)”的结构与功能	(403)
11.2.3 利用“可靠性分析(Reliability Analysis)”进行信度分析	(405)
11.3 问卷的效度分析	(409)
11.3.1 问卷的内容效度	(409)
11.3.2 效标关联效度	(411)
11.3.3 结构效度	(412)
11.4 主成分分析	(413)
11.4.1 主成分分析的基本思路	(413)
11.4.2 主成分分析的基本步骤	(416)
11.5 因子分析	(416)
11.5.1 因子分析概述	(417)
11.5.2 因子分析的基本思路	(417)
11.5.3 因子分析的基本步骤	(419)
11.5.4 “因子分析(Factor Analysis)”的功能与结构	(423)
11.5.5 利用“因子分析(Factor Analysis)”进行结构效度分析	(427)
11.5.6 利用因子得分进行分类与评价	(433)
附表	(437)
参考文献	(439)

第 1 章 抽样调查与 SPSS 概述

SPSS 作为社会统计分析的强有力工具,在社会调查中主要应用于采用定量研究范式的抽样调查,因此本书将以抽样调查过程中对问卷的统计分析工作流为主线展开。

1895 年,挪威统计学家凯尔(A. N. Kiaer)在国际统计学会(ISI)第五届大会上提出了“用代表性样本方法来代替全面调查”的建议,这一年被认为是抽样调查历史的开端。一百多年以来,随着抽样理论、统计分析技术、计算机技术和统计软件的发展,抽样调查作为社会调查的一种重要方法及获取统计资料的重要手段,日益受到政府各部门、企业、学术界与社会公众的重视,在行政管理、学术研究、民意调查和市场调查等领域,无论从应用的广度还是深度都有了极大的发展。

SPSS 是进行抽样调查研究的重要工具。最初,SPSS 是社会科学统计软件包英文名称(Statistical Package for the Social Sciences)首字母的缩写,现在全名为 SPSS Statistics,以区别于 SPSS 公司的其他产品。从开发至今已有 30 多年的历史,是国际上最通用的三大统计软件之一。

为使读者能够顺利地将 SPSS 应用于抽样调查的统计分析,本章将对调查问卷、概率抽样以及 SPSS 的必备知识做出简要的介绍。更为详尽的内容,可参见本书作者编著的《社会调查方法与实践》。

1.1 抽样调查概述

1.1.1 抽样调查的概念与特点

抽样调查(Sampling Survey)是从全体研究对象(称为总体)中,按一定方式选择或抽取一部分对象作为样本,调查工作仅在样本中进行,是一种非全面调查。从研究范式上看属于定量研究,抽样调查有一套完备的操作技术,包括抽样方法、资料收集方法和统计分析方法等。

抽样是一种选择调查对象的程序和方法。抽样方法有两种:概率抽样与非概率抽样。概率抽样也称随机抽样,是指按照随机原则,从总体中抽取一定数目的个体作为样本,总体中的每一个个体被选入样本的可能性是一样的,可以通过样本的信息来对总体进行描述。非概率抽样也称为方便抽样,是从方便的角度出发或根据研究者主观的判断来抽取样本,每个个体进入样本的可能性有多大是未知的,不能通过样本的信息来对总体进行描述,只能对样本进行描述。具体的抽样方法如图 1-1 所示。

问卷是抽样调查的主要工具,问卷的结构化是抽样调查所用的工具或手段区别于其他社会调查方法的重要特征。在数据资料的收集上,也有两种方式:自填式问卷调查和结构式访谈。自填式问卷调查是由调查对象自己在印制好的问卷上按着要求填答问卷中的题项;结构式访谈也称标准化访谈或者问卷访谈,是指访谈员根据事先准备好的问卷进行的访谈,访谈员必须严格按着规定的方式和顺序向被访者提出问卷中的问题,被访者只能根据问卷中的备选项选择回答,然后由访谈员填写。如果访谈员与被访者面对面地坐在一起进行访谈,称为直接

访谈或面访；双方事先约定好时间，通过电话、网络等通信手段，按既定的目的所进行的访谈，则称为间接访谈。

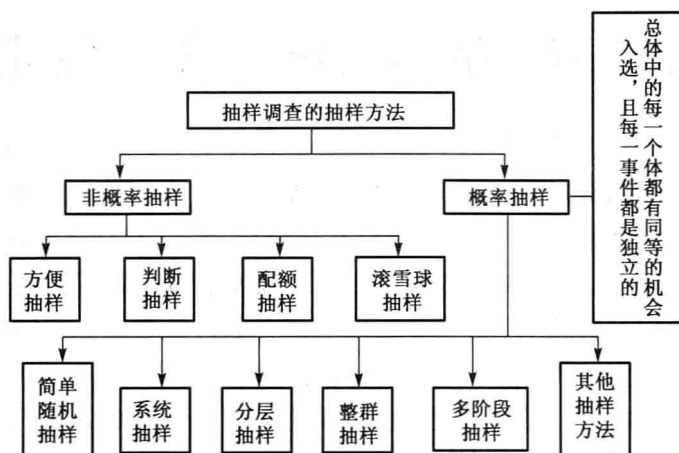


图 1-1 抽样方法分类

抽样调查的标准化工具就是问卷，因此抽样调查也称为问卷调查。问卷收回后，使用统计软件对数据进行统计分析，本书就是介绍如何利用统计软件包 SPSS 对问卷进行数据分析。

抽样调查的最大特点是采用定量的或数字的描述方式来研究总体，利用样本所产生的数据进行统计分析，并且不断地将各种新的统计分析方法引入到我们的研究中，使对问题的分析更加深入，这是其他调查方法所不具备的，鉴于此，抽样调查也称为“统计调查”。对于通过概率抽样方法得到的样本，不仅可以通过统计分析描述其样本的特征，而且可以对总体特征进行推断。综上所述，抽样调查与其他非全面调查比较，具有三个主要特点：一是随机原则；二是以结构化问卷为工具；三是从数量上推算总体。

1.1.2 抽样调查的过程

抽样调查是一种标准化程度较高的研究方法，调查过程有比较固定的程序，通常将调查过程分为五个阶段：选题阶段、准备阶段、调查阶段、分析阶段和总结阶段，具体抽样调查流程如图 1-2 所示。

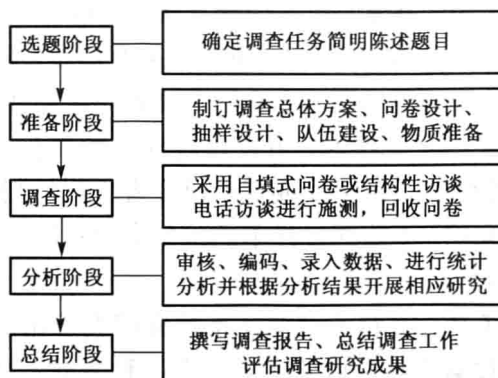


图 1-2 抽样调查过程流程

1.1.3 对抽样调查的评价

(1) 抽样调查与其他调查方法相比，有以下的优势：

第一,采用自填式问卷或结构式访谈进行调查可以直接从调查对象那里获取第一手资料,比某些间接地、利用文献等得到的第二手资料更准确、更真实、更符合研究者的需要。

第二,当研究对象是一个大型的群体时,抽样调查比其他调查方法更能节省时间、经费和人力,具有更高的效率。如采用团体调查方式进行问卷调查,在很短的时间内能够同时调查很多人;采用邮寄调查、网络调查,可以突破时空的限制,所有工作只需较少的研究人员来完成。

第三,结构化的问卷、概率抽样的方法,使统计分析更为深入,其他方法不可比拟。

(2)抽样调查也有一定的局限性,具体表现在以下几点:

第一,问卷调查缺少弹性和深度。如问卷是在实施调查之前设计的,在实施调查时,一旦发现问题或要增减题目,很难做出相应的改变。问卷中大量的问题是要描述人们的态度、行为和特征,对“为什么”只能给出有限的几个选择,很难做到对原本复杂的问题进行比较深入的探讨。

第二,容易受到人为因素的影响,效度较低。调查对象回答问卷的态度、对问卷的理解程度等都直接影响回收问卷的质量,所以收回的问卷有时可能并没有反映出调查对象潜在的价值观,调查对象的行为也并不一定总是与表述保持一致,通过问卷是否真正调查出我们所要调查的内容就特别值得关注。

第三,问卷的回收率有时难以保证,根据抽样设计方案抽到的调查对象有可能出现拒答的问题。

1.2 调查问卷的一般问题

1.2.1 问卷的结构

问卷一般由问卷标题、封面信、指导语、问题与选项、编码、结束语和其他资料七个部分组成。

问卷的标题向调查对象概括地说明了调查的主题,起到画龙点睛的作用。标题的结构一般是:调查对象+调查内容+“调查问卷”。例如,“天津市居民住房状况调查问卷”、“中国公众科学素养调查问卷”、“青年发展状况调查问卷”等。

问卷标题之后是给调查对象的一封短信,通常称为封面信或封面语。在封面信中应说明调查的目的和价值;调查对象的范围;调查对象做出回答的重要性;说明确保调查的秘密性以及回答本身对调查对象没有负面影响,要说明“对每个问题如何选项,没有对错之分”,以解除调查对象的心理压力,激励调查对象以认真、积极的态度对待调查。在信的开始或最后应说明调查者的身份,调查单位的名称,甚至将联系人的姓名、联系方式(电话号码或通信地址或电子信箱)也写上。信要写得简明、亲切,在字数上不要超过两三百字,在信的结尾处一定要真诚地对调查对象表示感谢。

指导语的功能有如冰箱、电视之类家电的使用说明书,是向调查对象或施测人员说明填答问卷的方法、要求和注意事项,有时还要说明回答所需要的时间,用以指导调查对象如何正确回答问卷。在具体编写和安排上,如果问卷形式比较简单,而且调查对象文化层次比较高,指导语可以在封面信中做出说明。如果指导语篇幅比较长,就需要将封面信与指导语分开,在封面信的后面紧接着写指导语。另外,有时需要通过指导语对问题或概念含义做出解释。

问题和选项是问卷的主体,一般包括调查所要询问的问题、回答问题的选项和方式,以及对回答方式的说明等。问卷中的问题从不同的视角可以做不同的分类,按问题提问的方式可分为开放式问题和封闭式问题。开放式问题(Open-ended question)是指调查者对所提出的问题没有规定答案的选择范围,调查对象可以按照自己的意愿自由地回答。例如,“您对目前商品房价格有何看法?”封闭式问题(Close-ended question)也称为选择题,是调查者将问题的内容和可选择的答案做了精心设计后编制的题目,调查对象只能根据自己的实际情况,从所给出的若干个可能的答案中进行选择,无法进行自由的发挥。

编码是将文字答案转变为数字代号的过程,以便于计算机做数据处理,主要包括两项工作:设计每一个题目所占的栏位以及规定每一个问题的每个选项所对应的数字,当问卷回收后,便可以将每份问卷的所有答案都转化为按规定应标记的数字。多数问卷不放编码,但也有问卷将编码放在问卷中,为计算机录入方便,编码放在问卷每页的最右边,并用一条竖线将它与问题及答案分开。如中国科学技术协会编制的《科技工作者状况调查问卷》每题均设有编码:

101 您的年龄

1. ☐ 29 岁及以下 2. ☐ 30~39 岁 3. ☐ 40~49 岁
4. ☐ 50~59 岁 5. ☐ 60 岁以上

101
☐

对于年龄的 5 个选项,1、2、3、4、5 分别表示 5 个年龄段,是选项的编码,101 是问题的编码。实施调查时,要求调查对象将自己的选择项所对应的编码填写在方框中。

对于自填式问卷往往有一个结束语。可以是简短的几句话,对调查对象表示真诚感谢,也可以在其后附一个开放式的问题,如“我们的调查到此就结束了,对您的配合与支持表示衷心的感谢!”、“欢迎您就本次调查及其有关问题提出建议和进一步的看法”。

其他资料主要是指采用面对面访问方式收集资料时对问卷的施测过程的记录,是对问卷进行审核和分析的重要依据,内容包括:面访完成的情况,包括访问所用时间、访问员姓名及访问员对回答的评价、审核员姓名及对审核的意见,未完成的原因等。

1.2.2 问卷的类型

从不同的视角可以将问卷做不同的分类,按使用问卷的方法可以分为自填式问卷(由调查对象自己填写)与访问式问卷(由访问员根据调查对象的回答进行填写);按问卷的功能可以分为主体问卷和过滤问卷(为筛选出符合条件的调查对象而设计的问卷);按问卷出题的方式又可以分为无结构型(开放式问卷)、结构型(封闭式问卷)和半结构型问卷。

1. 开放式问卷

如果一份问卷是由开放式问题组成的,称问卷是开放式问卷或无结构型问卷,其特点是所提问的问题没有在组织结构上加以严密的设计安排,只是围绕研究的目的提出一些问题。开放式问卷的最大优点是灵活性大,适应性强,能够发挥调查对象的主动性。当对一个问题的全部选项不十分清晰时,利用开放式问题,能够收集到各种不同的答案,甚至是意想不到的回答,为设计封闭式问题奠定了基础。但在整理问卷时比较困难,工作量比较大,对资料难以量化,无法做深入的统计分析。

2. 封闭式问卷

由封闭式问题组成的问卷称为封闭式问卷,即结构型问卷,是根据研究目的和主题精

心设计的具有结构的问卷,不仅包括有一定数目的问题,而且要按一定的提问方式和顺序进行,调查中不能随意增加或减少问题,也不能变动顺序和字句。大量的调查研究(特别是大型的调查)都是采用封闭式问卷方式进行的,因为只要研究者围绕研究假设精心设计问卷,调查对象认真回答(由于回答简便,较开放式问卷容易做到这一点),调查后对数据做必要的整理和转换,那么就可以进行各种深层次的统计分析,发现各种现象背后所具有的规律性。

封闭式问卷的缺点,一是在设计问题的选项时要比开放式问题花费更多的时间与精力;二是限制了调查对象的回答,问题难于深入;三是封闭式问题可能由于选项不全或太多、选项排列的次序,特别是仅仅要求调查对象在某个答案上画圈或勾,对于笔误或有意打错的,随意乱打记号的往往难以发现,所有这些问题都会在一定程度上影响了调查结果的准确性和真实性。

1.2.3 编制问卷的过程

1. 编制问卷的基本过程

形成一份问卷主要是通过以下几个步骤:准备工作、探索性研究、编制问卷初稿、根据专家调查和试测的结果修改问卷,形成最后的正式问卷。

问卷的准备工作包括:明确调查目的与主题、查阅文献资料、明确问卷调查收集资料的方式(采用自填式还是访谈式)、分析调查对象特征和初步界定概念的操作定义。概念的操作定义就是用一系列可以观察、可以测量的事物、现象和方法,对抽象概念做出界定和说明。例如,“发散思维”概念的操作定义为:在限定的时间内,学生列举出砖的各种用途,所得到的测试分数。

探索性研究主要包括:进行开放式问卷调查、召开座谈会、个别访谈及实地考察。通过这些工作,帮助我们弄清楚对某个指标可以提出哪些问题,某个问题可能会有多少种回答,经过归纳便可以设计出这个问题的选择项;还能够对各种问题的提法、不同类型的回答者所使用的语言、对不同问题的关注程度等获得第一手资料,有利于操作定义的完善,并将问题编写得更加清晰、选项更加客观具体,为开放式问卷转化为封闭式问卷奠定基础。

编制问卷初稿可以直接在计算机上操作,首先将所有问题整理出来并录入计算机,不必考虑题目的顺序,但要注意问题的提法和答案的设计;然后将一个个问题分别嵌入相关的主题中,并对每一部分的题目进行排序,形成问卷的主体;最后写好封面信、指导语和编码等内容。

问卷初稿完成后,要在听取专家对问卷的意见和建议的基础上修改问卷,然后进行小范围的试测,从三个方面发现问卷中的问题,一是对回收问卷的填答情况进行考查;二是对题目进行项目分析,删去鉴别度低的题目;三是进行信度效度分析,以便对问卷进行再修改,形成最终问卷。

2. 对试测问卷填答情况的考察

(1)如果回收率在 60% 以下,说明问卷设计存在较大的问题,必须做较大的修改;

(2)如果在回收的问卷中废卷很多,例如有的答卷中很多问题没有回答、填答方法出现错误,或在全部回收问卷中对某个问题的回答比较集中在一个选项上,都说明题目设计有问题,或封面语、指导语写得不够好,要分析产生的原因,加以修改。