



中央民族大学国家“十一五”“211工程”建设项目

应用社会统计 分析方法

◎ 焦开山 / 编著

中央民族大学出版社
China Minzu University Press



中央民族大学国家“十一五”“211工程”建设项目

Applied Social Statistical Methods

应用社会统计分析方法

◎ 焦开山 / 编著

中央民族大学出版社
China Minzu University Press

图书在版编目 (CIP) 数据

应用社会统计分析方法/焦开山编著. —北京: 中央民族大学出版社, 2012. 12

ISBN 978 - 7 - 5660 - 0324 - 9

I. ①应… II. ①焦… III. ①社会统计—统计分析—高等学校—教材 IV. ①C91 - 03

中国版本图书馆 CIP 数据核字 (2012) 第 274528 号

应用社会统计分析方法

编 著 焦开山

责任编辑 李苏幸

封面设计 布拉格

出 版 者 中央民族大学出版社

北京市海淀区中关村南大街 27 号 邮编: 100081

电话: 68472815 (发行部) 传真: 68932751 (发行部)

68932218 (总编室) 68932447 (办公室)

发 行 者 全国各地新华书店

印 刷 厂 北京宏伟双华有限公司

开 本 880 × 1230 (毫米) 1/32 印张: 9.375

字 数 240 千字

版 次 2012 年 12 月第 1 版 2012 年 12 月第 1 次印刷

书 号 ISBN 978 - 7 - 5660 - 0324 - 9

定 价 28.00 元

版权所有 翻印必究

民族学社会学教材与研究丛书

总 序

民族学与社会学学院的前身是建立于 1952 年的中央民族学院研究部。在 20 世纪五六十年代，研究部曾汇聚了中国大部分民族学与社会学的顶尖人才，如中国民族学与社会学的开拓者潘光旦、吴文藻、杨成志、吴泽霖、费孝通、林耀华和李有义等人以及他们的学生陈永龄、宋蜀华、施联朱、王辅仁、吴恒和王晓义等著名学者。

20 世纪 80 年代初，研究部更名为民族研究所，不久又建立了中国第一个民族学系，20 世纪 90 年代扩大为民族学研究院，2000 年更名为民族学与社会学学院。半个世纪以来，名称和建制的变化，并没有影响她致力于民族学教学与研究的宗旨，经过几代人的努力，从该院毕业的民族学专业的学士、硕士和博士已遍布全国各地，多为栋梁之材。同时出版了大量在国内影响巨大的专著和教材。如潘光旦、吴文藻、费孝通等人的文集，林耀华先生主编的《民族学通论》、宋蜀华先生的《民族研究文集》、陈永龄先

生的《中国民族学史》(英文版),还出版了全所历年研究成果的论集《民族研究论文集》(1981—1993年,共九册),这些出版物的共同特点是,以实地调查的材料为基础,以中国的56个民族为主要研究对象。几十年来,这已成为我院几代人的学术传统。

民族学(文化人类学)毕竟是一个自西方传来的学科,在中国发展历史较短,几十年来又多次受政治运动的干扰,所以与我国一些传统的老学科相比,中国的民族学无论在专业的理论、方法和研究成果方面,都是一个比较年轻、比较薄弱的学科。因此,今后本学科的重点是加强民族学专业的基础理论和方法的建设。为此,我们认为需要长期坚持两个方面的工作:

一、积极了解和借鉴国外学者有关的理论、方法和实践。这就要求我们既要翻译、介绍国外一些经典的名著,又要随时掌握国外研究的动向,将其最新的代表性作品翻译介绍给国内的读者和同行。

二、也是更重要的一个方面,继承我院50年来的传统,坚持实证性的研究方法,以中国的56个民族为主要研究对象,紧密联系实际,加强实地调查,以此为基础,进行理论总结,为建立独树一帜的、有中国特点的民族学理论而努力。

我们认为有必要使我们的学科建设和理论研究进一步系统化、规范化,并且在研究成果的基础上不断更新我们的教材。因此,我们于2000年成立了“民族学教材与研究丛书编委会”,目的是以民族学与社会学学院为基础,系统地编辑出版民族学专业的教材和以实证性研究为主的专著、调查报告和论文。

编委会将重点支持以下内容的教材和著作:

1. 民族学专业主干课和紧缺的必修课教材。
2. 以实地调查资料为基础的专题研究著作。
3. 国外民族学名著或前沿理论与方法的译著。
4. 有重要学术资料价值且规范的田野调查报告。

5. 本院教师实证性研究的论文集。

我们要求教材的编写者，应具有多年讲授该课程的资历，并且发表过有关的研究论文。我们要求丛书中的教材和论著应参考并引用国内外最新的相关研究成果，能够与国际学术界对话。

我们希望经过若干年的努力，本套丛书能够为民族学与社会学学院 50 年学术传统的发扬光大，为中国民族学学科的建设和中国民族学在国际学术界中较高地位的确立做出贡献。

杨圣敏

前 言

社会统计分析方法是一种应用极其广泛的方法，也是社会科学专业学生的必修课。目前，关于社会统计分析方法的教科书很多，这些教科书要么侧重于统计分析方法的数理方面，要么侧重于统计软件的操作方面，但把两者结合起来的却很少。鉴于此，本书介绍了社会统计分析方法的基本原理、技术以及 SPSS 统计软件操作方法。能够进行统计分析的软件有很多，本书之所以介绍 SPSS 统计分析软件是因为本软件简单易学，非常适合初学统计的人。

社会统计分析方法是一个不断发展着的开放的体系，且内容庞大，因此，本书的侧重点放在了统计分析的基本概念、基本原理和基本方法上。本书主要介绍了统计分析方法的应用层面，而不是背后的数理层面，故本书被称为“应用”社会统计分析方法。考虑到中国社会科学专业（比如社会学、社会工作、政治学等）的学生大多数学基础薄弱，对学习统计学可能会有恐惧的心理，本书力求减少数学公式的应用，而更加看重统计学方法的实际意义。在我看来，统计分析方法不仅仅是一种数学运算或者统计软件的操作，而

且是一种思维方式，是一种看待问题并且解决问题的视角。此外，需要强调的是，统计分析方法是我们开展社会研究的一种手段而不是目的，因此统计分析方法是为研究目标服务的。学习和应用统计分析方法，应该紧紧围绕着自己的研究目标，而不能“为了统计而统计”。

本书的主要内容包括了基础统计学的一些知识，它们是通往高级统计学的基础，因此掌握好和应用好这些知识是非常必要的。首先，本书的第1章介绍了定量研究的基本逻辑以及统计分析方法在其中的地位。接下来，第2章介绍统计分析方法的一些基本概念，在本书的后面几章中会不时地出现这些概念。统计分析的基本概念构成了统计分析方法的基础，也反映了统计分析的基本逻辑，因此掌握这些基本概念是进一步学习统计分析方法的前提。第3章简要介绍了SPSS统计软件的基本操作，包括如何把数据录入到SPSS中以及SPSS的运行方式。由于本书的一项重要内容就是介绍如何用SPSS进行统计分析，因此第3章的内容是后续学习的一个基础，尤其是那些对SPSS还比较陌生的读者。第4章介绍了数据的清理和整理方面的内容，这些工作是后续统计分析的基础。第5章介绍了描述性统计分析的方法，包括单一变量和双变量的描述性统计分析。社会研究的主要目的之一就是对社会现象的状况、过程和特征进行客观的、准确的描述，因此描述性统计分析方法在社会研究中占据着非常重要的地位。第6章介绍了推断性统计的基本概念和原理，包括概率论、抽样分布等方面的内容，这是后续几章内容的基础。掌握这些概念和原理对于学习和理解推断统计学的内容至关重要。第7章和第8章分别介绍了参数估计和假设检验的基本方法，这是推断性统计分析的基本内容。第9章和第10章分别介绍了一元线性回归模型和多元线性回归模型，这是社会统计分析方法应用最广的一种方法，掌握其基本原理对于学习更高级的统计分析方法也是非常有帮助的。

本书的出版得到了中央民族大学“211 工程”的资助，对此表示感谢。此外，还要感谢本书的评议人包智明教授以及我的博士生导师郭志刚教授，他们为本书的出版提出了非常宝贵的建议。由于能力有限，本书难免有某些不足之处，敬请读者批评指正，以便今后进一步改进和提高。关于本书所用数据，读者可发邮件索取。本人电子邮箱地址：muc_statistic@163.com。

焦开山

中央民族大学民族学与社会学学院社会学系

目 录

前 言	(1)
第1章 定量研究与统计分析方法	(1)
1.1 定量研究的方法论	(1)
1.2 统计分析方法及其作用	(8)
1.3 计算机在统计中的作用	(11)
1.4 本章小结	(13)
1.5 本章重要概念	(14)
1.6 思考与练习	(14)
第2章 统计分析的基本概念	(16)
2.1 总体与个体	(16)
2.2 抽样和样本	(18)
2.3 变量	(19)
2.4 数据	(25)
2.5 本章小结	(29)
2.6 本章重要概念	(30)
2.7 思考与练习	(30)
第3章 SPSS 基本操作	(32)
3.1 SPSS 的主界面	(33)
3.2 界定变量的属性	(35)
3.3 在 SPSS 中录入数据	(42)

3.4 SPSS 的运行方式	(44)
3.5 SPSS 的结果输出	(47)
3.6 本章小结	(48)
3.7 本章重要概念	(49)
3.8 思考与练习	(49)
第4章 数据的清理和整理	(50)
4.1 数据清理和整理的含义及其注意事项	(51)
4.2 数据资料清理的基本方式	(52)
4.3 数据资料整理的基本方式	(61)
4.4 本章小结	(67)
4.5 本章重要术语	(67)
4.6 思考与练习	(68)
第5章 描述性统计分析	(69)
5.1 描述统计分析的基本技术	(70)
5.2 单一变量的描述性统计	(80)
5.3 双变量的描述统计	(92)
5.4 用 SPSS 进行描述性统计分析	(102)
5.5 本章小结	(112)
5.6 本章重要概念	(114)
5.7 思考与练习	(114)
第6章 推断性统计分析的基础	(117)
6.1 概率引论	(118)
6.2 概率分布	(123)
6.3 正态分布	(131)
6.4 抽样分布	(139)
6.5 本章小结	(143)
6.6 本章重要概念	(144)
6.7 思考与练习	(144)

第 7 章 参数估计	(147)
7.1 估计概述	(147)
7.2 参数的点估计	(149)
7.3 区间估计	(151)
7.4 本章小结	(155)
7.5 本章重要概念	(156)
7.6 思考与练习	(156)
第 8 章 假设检验	(158)
8.1 假设检验的基本逻辑	(159)
8.2 关于总体平均值和比例的假设检验	(161)
8.3 关于两组比较的检验	(166)
8.4 单因素方差分析	(173)
8.5 对两个变量相关关系的检验	(177)
8.6 关于假设检验的一些说明	(182)
8.7 用 SPSS 进行假设检验	(185)
8.8 本章小结	(197)
8.9 本章重要概念	(198)
8.10 思考与练习	(198)
第 9 章 一元线性回归模型	(204)
9.1 一元线性回归模型的基本原理	(205)
9.2 回归模型的拟合优度和检验	(213)
9.3 实例分析	(217)
9.4 本章小结	(220)
9.5 本章重要概念	(221)
9.6 思考与练习	(221)
第 10 章 多元线性回归模型	(225)
10.1 多元线性回归模型的设置和估计	(226)
10.2 回归模型的拟合度测量	(233)

• 应用社会统计分析方法 •

10.3 回归模型的检验	(242)
10.4 虚拟变量的应用	(251)
10.5 回归诊断	(262)
10.6 本章小结	(275)
10.7 本章重要概念	(276)
10.8 思考与练习	(276)
参考文献	(283)

第1章 定量研究与统计分析方法

本章要点

- 定量研究的基本特征
- 定量研究的基本过程
- 统计分析在社会研究中的作用
- 统计分析软件

在社会科学研究中，定量研究已经成为一种主要的研究方式。在定量研究中，统计分析方法是最主要的一种工具，具有非常重要的地位。大量统计软件的出现，方便了对统计方法的使用。本章主要介绍定量研究的基本逻辑和基本特征、统计分析方法在社会研究中的地位以及统计软件在统计分析中的作用。

1.1 定量研究的方法论

1.1.1 数量与性质

数量和性质是事物的两种基本属性，数量是外在的表现，是可以通过经验观察到的，而性质则是事物的内在规定性，无法直接观察，需要通过逻辑思维能力获得。数量和性质不是对立的，两者是辩证统一的。数量是性质的反映，而性质要通过数量进行体现。请看下面的例子：

- 一个家庭每月食品支出占总支出的 60%，这说明了什么？
- 孩子每周会遭受父母 3 次以上的毒打，这说明了什么？

我们看到，60%、每周 3 次这些都是数量，但却反映了问题的性质。一个家庭中每月食品支出占总支出的 60%，说明这个家庭是贫困家庭^①；孩子每周会遭受父母 3 次以上的毒打，说明孩子正遭受父母的家庭暴力。我们看到，通过对社会现象的数量性的观察，可以获得对现象的性质判断。同样，如果我们对社会现象的内在性质进行说明，我们应该怎么办呢？比如：

- 中国是发展中国家。
- 你是一个健康的人。

要说明中国是一个发展中国家，就需要列举一系列的经济社会指标，比如人均国内生产总值（GDP）是多少、人均收入是多少等等，而要说明你是一个健康的人，就需要列举一系列的健康测量指标，比如血压多少、患病次数等。这些指标都是可以观察到的，我们就是根据这些指标而形成对事物性质的判断。当然了，你也可以通过其他方式形成对事物性质的判断，比如直觉的方式，即“我感觉你是个健康的人”；或者思辨的方式，比如“因为……所以我认为你是个健康的人”。但是，通过外在数量的特征判断事物的性质，是相对最为科学、可靠的方式。

① 食品支出总额占个人消费支出总额的比重被称为恩格尔系数（Engel's Coefficient）。联合国根据恩格尔系数的大小，对世界各国的生活水平有一个划分标准，即一个国家平均家庭恩格尔系数大于 60% 为贫穷、50% - 60% 为温饱、40% - 50% 为小康、30% - 40% 属于相对富裕、20% - 30% 为富裕、20% 以下为极其富裕。

当然了,有些事物是很难量化^①的,但是,这并不能成为批评通过数量特征去判断事物性质这种方法的理由。很难量化,只是表明了测量工具和测量技术的局限性而已。随着测量工具的改进,先前不能进行量化的现象也许就可以量化了。

1.1.2 什么是定量研究

社会研究是为了描述、解释以及预测人类行为以及相关社会现象而去收集资料、并对资料进行分析和解释的过程。在这一过程中,需要使用一些方法和技术,其中定量研究方法(quantitative research method)和定性研究方法(qualitative research method)是社会研究中应用最广、影响最大的两种方法。与侧重于对事物的意义、情境的描述和理解的定性研究不同,定量研究侧重于对事物的测量与计算,通过对量的考察去把握事物内在的性质。

概括起来,定量研究方法的特征主要有以下几个方面:

● 在哲学基础上,定量研究主要秉承实证主义的观点,即把社会现象看成是一个客观的事实,社会研究要以经验观察方法为主。

● 在研究目的上,定量研究旨在确定或者验证变量之间的关系、相互影响和因果联系。此外,定量研究通常是用来进行理论检验的。如果我们要对一个理论进行检验,会先提出一系列假设,然后通过问卷或者实验等方法去收集量化的数据资料,然后对量化的资料进行统计分析,来验证预想的理论或者模型。

● 在研究的手段上,定量研究是用数字而不是文字来描述研

^① 所谓量化就是用一些具体的经验指标来表示一个抽象概念的内涵。所谓“很难量化”并不是说不能量化,而是说在用哪些指标来表示一个概念内涵时,不同的人有不同的看法,没有形成相对一致的看法。

究对象，用统计学方法对收集的数据资料进行分析。因此，统计学是定量研究的基础和核心知识。

● 在研究程序上，强调标准化、系统化和操作化。在定量研究过程中，要尽量避免人的主观意识和看法对研究的干扰，以数据资料为依据，获得客观的研究结果。当然了，在研究问题的选择以及研究结果的解释及应用方面，定量研究无法做到也无须做到客观。

● 在研究结果上，定量研究侧重于把研究结论推论到更广泛的总体中。比如，在有一万名学生的学校中，随机抽取 500 名学生，根据对这 500 名学生的研究结果，可以获得对整个学校的结论。

1.1.3 定量研究的基本过程

作为一种系统的、科学的认识活动，定量研究遵循着一个基本的程序。之所以成为基本的程序，是因为就某个具体研究而已，可能并不会完全按照既定的程序展开。

一般而言，定量研究的基本程序主要包括以下几个阶段：

(1) 研究问题的选择阶段

选择研究问题是一项社会研究活动的起点，是整个研究的第一步。研究问题可以来自于以往的理论，也可以来自于自己的生活经历和兴趣。选择研究问题对于做好一项具体的社会研究具有十分重要的意义。在选择研究问题时，应该考虑的因素有：

● 研究题目所具有的意义和价值，即该题目是否“值得去做”。一个研究题目的意义和价值主要体现在两个方面：在理论上，该题目的研究是否会对相关的理论知识有所贡献；在实践上，该题目的研究是否有利于提出某些具体问题的解决之道。在一年一度的