

普通高等学校社会学专业主干课系列教材

社会统计学

Social Statistics

主编 张彦



高等教育出版社

普通高等学校社会学专业主干课系列教材

社会统计学

Social Statistics

主编 张彦

高等教育出版社

内容简介

本书作为高校社会学专业主干课程教材之一,主要介绍社会统计学的基本概念、发展历史、应用准则、工作程序等,对单变项简化、集中趋势和离散趋势的度量、概率分布、参数估计、非参数检验、相关与回归分析、方差分析等统计方法作了详细而系统的说明。本书十分注意把知识体系的完整性同教学上的由浅入深、循序渐进的连贯性统一起来;特别强调社会统计学同另一主干课程——社会调查研究方法的联系,并注重误差分析。与此同时,书中的符号编排、概念阐述、图表配备、例题选择等前后呼应,融为一体,既便于教又便于学。本书“通”、“专”相宜,适于高校社会学专业本科教学使用,亦适于为政治学、管理学、人口学等许多学科及社会工作采用。

图书在版编目(CIP)数据

社会统计学/张彦主编.—北京:高等教育出版社,
2005.6

ISBN 7-04-017193-7

I.社… II.张… III.社会统计-高等学校-教材 IV.C91-03

中国版本图书馆 CIP 数据核字(2005)第 029743 号

策划编辑 干咏昕 责任编辑 蒋青 封面设计 于文燕 责任绘图 杜晓丹
版式设计 王莹 责任校对 杨凤玲 责任印制 杨明

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100011
总 机 010-58581000
经 销 北京蓝色畅想图书发行有限公司
印 刷 北京北苑印刷有限责任公司

图书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网址订购 <http://www.landaco.com>
<http://www.landaco.com.cn>

开 本 787×960 1/16
印 张 23
字 数 430 000

版 次 2005 年 6 月第 1 版
印 次 2005 年 6 月第 1 次印刷
定 价 26.40 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 17193-00

前 言

社会统计学是教育部规定的社会学专业的十门主干课程之一。在我国,社会统计学的地位和作用是随着改革开放的深化而不断得到加强的,过去在统计工作上反映的“经济情况多、社会情况少”的局面也随之得到很大的改观。进入21世纪,当前中国社会主义建设事业正在按照“统筹城乡发展、统筹区域发展、统筹经济社会发展、统筹人与自然和谐发展、统筹国内发展和对外开放”的科学发展观全面推进。社会统计学作为从数量方面描述社会状况、社会发展和社会问题的一门学科,其作用和地位也更显突出。本书编写的指导思想是,从实现建设全面小康社会的战略目标和与国际统计实践接轨的需要出发,系统阐述社会统计学的基本原理和方法。内容包括:社会统计资料搜集和整理、综合指标测量法、概率论基础和统计推断、动态分析法和指数分析法等等。全书根据变量的四个层次(定类、定序、定距、定比),以社会统计的实践过程为经,以社会统计学的原理和方法为纬,组成一个较为完整的篇章结构和知识体系。

本书的特点是:

1. 力求反映我国改革开放以来,社会各界关于社会统计学的性质、地位、特点、内容等一系列基本问题的共识,注意吸收本学科新的比较成熟的成果,发挥社会统计学本身的特长,反映科学发展观、全面小康目标的提出及实施的社会背景和有关理论概念,具有时代感。

2. 将数理统计融合于社会研究之中,特别强调社会统计分析和调查研究的联系。本书将时间数列的指标分析和趋势分析及综合指数法纳入教材,使篇章结构和知识体系更加完整,并使之更加切合社会研究的实际需要。

3. 以学以致用为目的,把知识体系的完整性同教学上的由浅入深、循序渐进的连贯性统一起来,在深入浅出上下工夫,帮助文科学生克服对数学的畏难心理,从而很好地提高他们学习社会统计学的兴趣和主动性。本书的所有知识点前后呼应,符号、概念、图表、例题等经精心编排和仔细推敲,既便于教更便于学。

本书由张彦教授主编,参编作者有邱海雄教授、郝大海教授、刘欣博士、甄志宏博士、王越平博士,他们分别参与了相关分析、回归分析、方差分析等章节的撰写。干咏昕同志、蒋青同志对本书的出版工作也付出了大量心血。

张 彦

2005年1月于上海

目 录

第一章 导论	1
第一节 什么是社会统计学	1
第二节 社会统计学的几个基本概念	9
第二章 社会统计资料的搜集	15
第一节 统计调查的方法及种类	15
第二节 统计调查的组织形式	16
第三节 概念的操作化与测量	20
第四节 统计误差	23
第三章 社会统计资料的整理	26
第一节 统计分组的原则与标准	26
第二节 统计表	27
第三节 变量数列的编制	32
第四节 统计图	39
第四章 集中趋势测量法	47
第一节 算术平均数	47
第二节 中位数	51
第三节 众数	55
第四节 几何平均数、调和平均数及其他	57
第五章 离中趋势测量法	62
第一节 全距与四分位差	62
第二节 平均差	64
第三节 标准差	66
第四节 相对离势	70
第六章 概率与概率分布	75
第一节 概率论	75
第二节 概率的数学性质	81
第三节 概率分布、期望值与变异数	89
第七章 假设检验	98
第一节 二项分布	98
第二节 统计检验的基本步骤	101
第三节 正态分布	107
第四节 中心极限定理	115

第五节 总体均值和成数的单样本检验	121
第八章 常用统计分布	128
第一节 超几何分布	128
第二节 泊松分布	130
第三节 χ^2 分布	134
第四节 F 分布	138
第九章 参数估计	142
第一节 点估计	142
第二节 区间估计	144
第三节 其他类型的置信区间	149
第四节 抽样平均误差	152
第五节 样本容量的确定	159
第十章 双样本假设检验及区间估计	162
第一节 两总体大样本假设检验	162
第二节 两总体小样本假设检验	165
第三节 配对样本的假设检验	170
第四节 双样本区间估计	174
第十一章 非参数检验	180
第一节 符号检验	181
第二节 配对符号秩检验	183
第三节 秩和检验	185
第四节 游程检验	188
第五节 累计频数检验	191
第十二章 相关与回归分析	196
第一节 变量之间的相互关系	196
第二节 定类变量的相关分析	198
第三节 定序变量的相关分析	211
第四节 定距变量的相关分析	224
第五节 回归分析	228
第六节 曲线相关与回归	234
第十三章 χ^2 检验与方差分析	242
第一节 拟合优度检验	242
第二节 无关联性检验	249
第三节 方差分析	255
第四节 回归方程与相关系数的检验	266
第十四章 动态分析与指数分析	274
第一节 时间数列及其指标分析	274
第二节 时间数列的趋势分析	287

第三节 指数分析法	295
附表 1 随机数表	313
附表 2 二项分布个别概率表	318
附表 3 二项分布累加概率表	322
附表 4 标准正态分布 (Z 值) 表	324
附表 5 t 分布表	326
附表 6 泊松分布表	328
附表 7 χ^2 分布表	330
附表 8 F 分布表	332
附表 9 配对符号秩检验表	343
附表 10 秩和检验表	345
附表 11 游程检验表 ($n_1 > n_2$)	349
附表 12 相关系数表	352
附表 13 累计法平均速度查对表	353
参考文献	358

第一章 导 论

统计是关于数字和数据合成的学问。马克思说，科学只有当它利用了数学的时候，它才达到了完善的程度。恩格斯也说，数学是我们所知道的最纯粹的数量规定，但是它充满了质的差异。毛泽东则反复告诫我们，如果不懂得注意事物的数量方面，不懂得注意基本的统计、主要的百分比，不懂得注意决定事物质量的数量界限，一切都是胸中无“数”，结果就不能不犯错误。社会现象同自然现象一样，既有质的规定性的一面，也有量的规定性的一面。社会现象的规模、水平、程度、比例等都是通过一定的数量特征表现出来的，例如社区规模、生活水平、都市化程度、人口迁移率等等。所以，任何一个立身于社会的有识青年，学习和掌握社会统计学，对于自己的成长和发展都是大有帮助的。特别是在急剧变革和转型的当代社会，我们面临的问题高度复杂，我们面临的情况千变万化，社会统计学不仅会增强我们认识社会的洞察力，透过现象看本质，而且可以增强我们对未来的适应能力，透过趋势作出正确选择和决策。

第一节 什么是社会统计学

“统计”一词，英语为 statistics，用作复数名词时，意思是统计资料，用作单数名词时，指的是统计学。一般来说，统计这个词包括三个含义：统计工作、统计资料和统计学。这三者之间存在着密切的联系，统计资料是统计工作的成果，统计学来源于统计工作。

1. 社会统计的产生与发展

社会统计的实践已有几千年的历史，是最早产生的统计。古埃及在建造金字塔时，为征集建筑费用对全国人口和财产进行过简单的调查和统计。古希腊在公元前 400 年就进行过人口普查。古罗马则在公元前 400 年就建立了出生、死亡登记制度。从奴隶社会到封建社会这段漫长的时期，社会统计主要局限于对事物进行原始的调查登记和简单的计算汇总，这是奴隶主或封建王朝为实现征税或服兵役和服劳役的需要而进行的。

华夏文化是世界上有数的几大文明之一，有关统计的方法、统计思想及统计工作，与中国的文化渊源紧密相连，都可以追溯到远古时代。根据《尚

书·禹贡》记载，公元前2 000多年大禹治水时，依据山川土质、人口物产等分全国为九州，从而有了九州人口和土地数字，称为九州表。东周及其以后的各个朝代，关于土地、人口的数字不绝史册，如明初编制的黄册和鱼鳞册。黄册乃全国户口名册，记载各户所有丁粮；鱼鳞册系全国土地图籍，绘其地形，记明地位，具有现代统计图表的性质。进入封建时代，我国还有一些杰出的思想家提出过不少科学的统计思想和先进的统计方法。例如，汉代司马迁在他编的《史记》中多次用到统计表，晋代在分组上所采用的两端开口组，宋代对中位数的应用等等，这一切都远远走在当时的西欧诸国之前。

尽管如此，由于中国始终没有进入资本主义社会，商品经济始终没有形成一种社会经济形态，中国统计工作犹如一条流入沙漠的河流，慢慢地消失、衰亡了。清人孟森在其译作《统计通论》中写道：“吾国人方将就学于日本，以开统计之路。夫学问为天下之公器，在先觉者固荣，后起者亦不得为辱。吾统计之传者，已四千余年于兹。吾知孰国学术，有退无进，即以统计论，方今稍耳其名。观西人之勤于此者三百年，日本之急起而直追者亦四十年。以如是一日不可少之学。而吾士大夫犹梦梦，可哀也。”

17世纪后，由于资本主义生产关系和商品经济形态的形成，在西方各国，简单的人口、土地、资源统计已不能满足社会发展的需要，因而产生了以工业、农业、贸易、交通等方面统计为主的社会经济统计。在不断丰富的统计实践经验的基础上，比较系统的统计理论知识逐步形成，这便产生了统计学。目前在统计分析中经常使用的一些基本方法和术语都始于这一个时期，比如最小方法、正态分布曲线与误差计算等等。

但在近代统计学的初创阶段，由于统计学者所处的历史环境不同，对统计实践的理解不同，从而总结出来的经验和概括出来的理论也有所差别，这就产生了不同的统计学派。统计学的发展，很大程度上正是在不同学派的争论中得以实现的。

国势学派 这一学派产生于德国，其创始人为康令（1606—1681）和阿享瓦尔（1719—1772）。阿享瓦尔在1749年所著《近代欧洲各国国家学纲要》一书绪言中，把国势学定名为“Statistik”（统计）这个词。他根据拉丁文 Status（指“现状”、“状态”）、意大利文 Stato（指“国家”）和 Statista（指“熟悉国情者”）及德文 Statisti（指“国情”、“国势”）等字根，创造出“Statistik”这个新词，其原意是指“国家显著事项的比较和记述”。从此该学派一直以统计学命名。但由于其特征是只用文字记述，不用数字计量，所以历史上人们又将这一学派称为“有名无实”的学派。国势学派又可称为记述学派和历史学派。该学派有一句名言：“统计是静止的历史，历史是流动的统计。”

“统计”这个词创造出来后，被各国相继采用，并译成各国文字：法国译

为 Statistique, 意大利译为 Statistica, 英国译为 Statistics. 日本最初译为“政表”、“政算”、“国势”、“形势”等, 直到 1880 年在太政官中设立了统计院, 才确定以“统计”二字为正名, 1903 年(清光绪廿九年)钮永建、林卓南等翻译了横山雅南所著的《统计讲义录》一书, 从此“统计”一词从日本传到我国。

政治算术学派 该学派的创始人是英国人格朗特(1620—1674)和威廉·配第(1623—1687), 它以后者的著作《政治算术》而得名。1662 年格朗特发表了第一篇关于人口统计的论文《自然和政治观察》, 从宗教管理、商业、气候、疾病等方面, 对当时伦敦人口的出生率、死亡率和性别比例等方面进行了综合的统计分析, 这是政治算术学派的第一篇研究论文。1676 年《政治算术》一书问世, 威廉·配第运用有关人口、土地税收和国家收入等方面的数字资料, 对英国、法国、荷兰的经济实力进行比较, 首创了一种数字对比分析的方法。用配第自己的话来说, “即用数字、重量、尺度来表达自己想说的的问题”。配第没有使用统计学这一名词, 但他使用的社会宏观数量对比和分析方法揭示了统计学所要研究的内容, 因此历史上人们又将这一学派称为“有实无名”学派。马克思对配第评价很高, 誉他为“政治经济学之父, 在某种程度上也可以说是统计学的创始人”。

政治算术学派与记述学派互相影响, 互相争论, 共存了近 200 年之久。直至 19 世纪中叶, 统计学被公认必须与数量分析方法和手段联系在一起, 争论才告平息。

数理统计学派和社会统计学派 19 世纪末 20 世纪初, 随着资本主义的发展, 统计理论没等上述两个学派争论结束, 就有了新发展。数理统计学派的创始人是比利时的凯特勒(1796—1874), 其最大的贡献就是将法国的古典概率论引入统计学, 用纯数学的方法对社会现象进行研究。由于把概率论引进统计学, 使社会随机现象数量方面的研究提高了准确性。凯特勒应用大量观察法研究犯罪、人口等现象, 曾准确地预测到 1830 年法国犯罪行为的种类和数量。1867 年, 一门兼有数学和统计学双重性质的学科被命名为“数理统计学”, 凯特勒也被人称为“现代统计学之父”。

凯特勒的另一个重要贡献, 是他把政治经济学、数学和当时政府统计工作的方法结合在一起, 建立了一个专门研究社会现象的统计学派。后来这个学派传到德国, 就出现了以克尼斯(1821—1898)、梅尔(1841—1925)和恩格尔(1821—1896)为代表的德国社会统计学派。但在近代统计学的发展过程中, 数理统计学派和社会统计学派的争论也是比较大的。

第一次世界大战前后, 随着社会统计学派的中心逐步向英、美等国转移, 社会统计学与社会学的关系日益明确。1900 年美国统计学家马约·史密斯发

表了《统计学和社会学》，他认为“统计学是使用特殊方法的社会学的一个部门，只有利用特殊方法才能得到最好的答案”。他解释说：“所谓特殊方法，即是用大量观察法解释社会生活问题。通过大量事例的数量观察，可以概括出有关人类社会的系统知识。”1920年，另一位美国统计学家史特威·恰平出版了《实地调查与社会研究》，还著有《社会学中的科学方法》，提出“统计学不仅仅是一种方法，也是一门科学”的论点。这表明，现代社会统计学不但应用数理统计学的理论和方法形成了自己的科学方法论，而且它同社会学的发展紧密地结合起来，与此同时一种狭义上的社会统计学逐步为人们所重视。

应该指出，二次大战后，社会统计学在广义和狭义两方面的实践意义都逐步得到了人们的公认。众所周知，20世纪下半叶，世界政治格局发生了重大变化，科学技术迅速进步，经济发展成为全人类面临的共同问题。但在注重经济增长的同时，在工业发达国家出现了一系列难以解决的社会问题，诸如环境污染、犯罪率和离婚率上升、失业和贫困加剧等等。在发展中国家，经济增长反而带来了政治的不稳定、社会动荡和国内的贫富差距。这一切使人们逐渐认识到，经济增长并不一定意味着社会发展，经济高涨不等于社会进步，经济效益好不等于社会效益佳。这就要求从社会整体发展的观点出发，应用社会调查和各种统计方法，收集大量的、更全面的事实来描述、分析、研究社会发展状况和发展趋势，从而监测社会发展和采取相应措施，达到统计为社会服务的目的。

20世纪60年代以来，西方发达资本主义国家先后都制定了社会发展计划。法国较早地强调了计划的社会方面，在1962—1965年的战后第四个计划中，法国政府把原来的计划名称“经济现代化与投资计划”改为“经济与社会发展计划”。该计划规定，要关心经济增长成果的分配和区域发展。在计划中，和很高的总增长率（每年5.5%）并列的还有社会投资（住房、教育和卫生）每年增长10%的计划目标。其他发达资本主义国家在制定的计划中，也越来越重视计划的社会方面。

垄断资本主义国家制定与实施社会发展计划的活动直接推动了社会统计的发展。20世纪60年代首先在美国掀起了一个颇有声势的“社会指标运动”。1966年美国的雷蒙·布埃尔提出用社会指标的方法解决社会分析和社会规划的基本理论，出版了《社会指标》一书，其后产生了深远的影响。应该指出，“社会指标运动”是社会思潮的一种具体表现，它使社会学和社会统计学在宏观研究方面进了一大步。

从20世纪70年代起，一些区域性或世界性的组织开始颁发统一使用的社会统计指标体系。1976年经互会为各成员国颁发了《社会统计基本指标体系》。同年，联合国经济合作与发展组织（OECD）编制了《社会生活质量的

计量》的社会统计指标体系。

建国以来，我国已经建立了以社会主义计划经济为基础的统计组织机构和统计制度方法，在社会经济统计的实际工作和理论研究方面，积累了不少适合我国国情的宝贵经验，取得了很大成就。但也应该看到，我国过去的统计工作，存在着反映物质生产领域情况多、反映非物质生产领域情况少的弊病。1982年，我国把国民经济发展计划改为国民经济与社会发展计划。为了适应这一状况，我国的统计工作发生了一系列变化。

首先，是统计组织机构与统计制度的变化。1982年，国家统计局成立了社会统计司，并开始建立与完善社会统计制度。过去，统计工作、统计理论与方法侧重于经济统计，是为了适应恢复生产与发展社会主义计划经济的需要。现在，经济统计与社会统计同时受到重视，是为了更好地完成我国经济与社会协调发展的战略任务。

其次，是统计指标与制度化报表的变化。1983年7月全国统计工作会议审查通过《社会统计指标体系》（草案）。从此，反映社会统计指标体系的各种月、季度、年度报表，进入了各行各业，成为各级政府了解情况、检查工作与制定规划所不可缺少的工具。1983年，由国家统计局社会统计司编写的《中国社会统计资料》首次公开出版。

2. 社会统计学的对象与特点

在社会学研究中，社会统计学具有极为重要的地位和作用。这是因为，作为社会学研究对象的社会现象有其质和量两方面。社会学研究社会现象，不能不了解它在数量上的表现，这就决定了存在于社会学和社会统计学之间的内部联系。

社会统计学是运用统计的一般原理，对社会各种静态结构与动态趋势进行定量描述或推断的一种专门方法与技术。人们既用它来分析已经发生和正在发生的现象，也用它来估计预测未来可能发生的现象。我国于20世纪80年代初恢复社会学这门学科以来，各个社会学系（专业）为建构社会统计学理论与方法、形成具有中国特色的社会统计学教材，做了不少有益的工作。目前，社会统计学是社会学专业的主干课程，其他一些社会科学学科，如政治学、历史学、法学等，也纷纷把这门课程列入教学计划。

任何一门独立的学科，都有区别于其他学科的独特研究对象。社会统计学的研究对象，概括而言是指社会现象的数量方面。但值得注意的是，由于客观需要和具体任务的不同，社会统计学具有广义与狭义之分，它们的对象范围是不同的。

广义的社会统计学，在我国实际上就是社会经济统计学。20世纪80年代

初,我国把国民经济发展计划改为国民经济与社会发展计划,与此同时,中国社会学在中断30多年之后恢复并重建,由此产生了“社会统计学从社会经济统计学独立出来”的需要。从社会经济统计学中独立出来的社会统计学便是狭义的社会统计学,它是本书讨论的重点。

社会统计学从社会经济统计学中独立出来,造成了社会统计学与经济统计学相并列的格局。那么狭义的社会统计学同经济统计学是怎样一种关系呢?

从理论上讲,社会统计学与经济统计学都是社会经济统计学的有机组成部分。但从实际上看,它们所司的领域、研究对象以及内容和方法,都各具特点。就研究对象而言,经济统计主要是从研究和反映生产力关系方面去认识社会;社会统计则主要是从研究和反映一定经济基础之上的上层建筑方面去认识社会。就研究内容而言,经济统计侧重于研究和反映物质资料的生产、分配、交换和消费过程,经常使用度量衡单位、货币单位和劳动时间单位等;社会统计也使用这些计量单位,但在社会研究的许多领域,经常需要对人们的态度、观念、行为进行度量,测量这些社会现象目前还没有一个精确而统一的尺度,只能以近似估算或词语表达等方式来代替。就调查方法而言,在经济统计中,由于其研究对象所具有的特征,全面调查、经常性调查较常用,特别是统计表这种调查组织形式,可以定期取得系统的资料;在社会统计中,由于其研究对象(如社会行为等)所具有的特征,抽样调查更为常用。即便是经济统计和社会统计都采用抽样调查方法,前者常常重视抽样推断的方法,以达到从部分到全体的认识,后者则常常是先依据社会研究的目的建立假设,然后再应用抽样法来检定假设是否成立。

3. 社会统计学的方法

一门学科的方法,取决于该门学科的研究对象。所谓社会统计方法,简单地指搜集、整理与分析资料的研究技术或手段。同众多定量研究方法一样,社会统计方法的根本特征也是数量分析。但社会统计学所研究的对象具有大量性、变异性的特点,这就决定了社会统计研究必须采用大量观察法。所谓大量观察,即是就总体中足够多的单位进行调查和综合分析,用以反映社会总体的数量特征。大量观察法是统计调查阶段的重要方法。

大量观察法之所以成为统计上特有的方法,是与大数定律的作用分不开的。大数定律是随机现象出现的基本规律,它的一般意义是:观察过程中每次取得的结果可能不同(因为具有偶然性),但大量重复观察结果的平均值却几乎接近某个确定的数值。

社会统计以人为中心,它所考察的社会现象是由具有主观能动性的人们之间的社会活动所造成。因此我们在观察社会现象的数量特征时,对象总体各单

位的具体特征和数量表现出较大的差异性。大数定律说明，由大量的相互对立的随机现象所构成的因素总体，如果每个因素对总体的影响相对较小，则这些偶然性质的影响可以相互抵消，使大量社会现象的数量特征借助平均数形式，显示出必然的规律性。

例如，当我们观察个别家庭或少数家庭新生儿的性别时，男女的比例可能极为参差不齐——有的生男不生女，有的生女不生男，有的是女多男少，有的是女少男多。然而经过大量观察，我们发现男婴、女婴的出生数趋于接近。也就是说男婴或女婴的出生数各占新生儿总数之半。新生儿的性别比经过大量观察趋于一个特定值，这就是我们在第六章将要详细阐述的概率论的一个实例。大数定律和概率论实质上是讨论同一性质的问题，只是概率论侧重于各种抽样结果发生之概率的具体计算，而大数定律侧重于统计规律性的理论说明，这是我们一开始学习社会统计时就要加以认识 and 理解的。

通过社会统计调查所获得的大量原始材料，还必须运用统计学所特有的分组方法，进行加工整理，才能供综合分析之用。统计分组法是统计资料整理的基本方法，对于这一点，我们学习了第三章之后，将会有较深刻的理解。

统计的重点是统计分析，这是在统计资料收集和整理之后展开的。统计按其内容主要包括两个方面：描述统计和推断统计。实际上，描述统计和推断统计也就是统计分析的两种基本方法，它们之间有紧密联系也有重要差别。

所谓描述统计就是讨论范围仅以搜集的资料本身为限，而不予以扩大。早期统计都是描述统计。例如国势学派对欧洲各国的政治、经济、地理、人口等状况用文字及数字进行描述以说明国情；政治算术学派也基本以数字资料为基础，对多种社会现象进行比较分析。统计为执行其描述功能，已发展出综合指标法、离势分析法、动态分析法、指数分析法、相关分析法等一系列重要方法，这些都是我们要加以学习的。

描述统计是统计分析方法的基础。包括推断统计在内，没有描述统计作为基础，想要运用好也是不可能的。描述统计所用数学较少，实用性又很强，因此在社会统计中使用的机会很多。这里需要指出的是，当研究者需要处理两个以上变量的相互关系时，描述统计往往特别有用。

所谓推断统计，主要是依据概率论，研究如何依据有限资料对总体性质作推断，从而使统计的功能大为扩充。统计推断需要的统计资料相对而言较少，所以在人力、物力及时间等方面比较经济，研究和应用的范围也拓宽了许多。但推断统计所用数学较多，它是在数理统计学派之后发展起来的，属于比较现代的统计分析方法。

推断统计是对抽样调查来讲的。描述统计固然对处理样本资料也有效，但样本能否代表总体，能在多大的程度上代表总体，只有通过统计推断才能得出

结论。所以抽样调查一定要有推断统计。那种认为样本理所当然能够代表总体的看法是没有根据的。在这里，最基本的一点就是要求样本必须是随机产生的。

推断统计有两个基本内容：①参数估计，即由样本的指标数值推断总体的相应的指标数值，它包括点估计和区间估计；②假设检验，即就社会研究中提出的某种假设应用抽样方法来加以统计检验。推断统计在社会统计方法中占有相当重要的地位。从总体而言，社会统计学较之经济统计学在内容上要深一些，难度也大一些。

4. 社会统计工作的程序

程序化是社会研究的一个重要特点。所谓程序是指研究工作的顺序和步骤，这个顺序和步骤由研究者制定，看上去好像是主观的产物，而实际上它是人们根据客观法则和人的认识规律而作出的一种逻辑安排。统计工作的程序，须视具体情况具体对待，一般可分为以下五步：

(1) 制定计划

“凡事预则立，不预则废”，统计工作也是如此。一项具体的统计工作确定后，在统计调查展开之前，我们必须事先就任务的性质、目的、范围、人员、经费、表格、进度、方法等作通盘考虑，做好计划，然后按计划行事，方能取得事半功倍的效果。

(2) 统计调查

统计调查就是按照统计研究所预定的目的和任务，运用各种调查的组织形式，有组织、有计划地向调查对象搜集原始资料的工作过程。如果把整个统计工作看作是建造一座楼房，制定计划就是给这座大楼绘制蓝图，统计调查就是给这座大楼提供砖块。没有砖块，再好的蓝图也无法实施。

(3) 统计整理

统计整理就是根据统计研究的目的和任务，对调查所得的资料进行加工整理，使之条理化、系统化。统计调查取得的资料，只是说明总体单位特征的零乱的原始资料，属于反映事物表象、某一侧面或外部联系的感性材料。只有通过对这些原始资料的加工整理，呈现出总体内部的分布特征，才能为统计分析提炼各种有用信息打下基础。因此，统计整理是统计调查和统计分析的一个中间环节（详见第三章）。

(4) 统计分析

统计分析是对研究对象进行定量分析，它是一项专门技术，也是本书的重点。在社会统计中，统计被分成描述统计和推断统计两大部分（详见第四章之后）。

(5) 统计报告

大量统计资料经系统整理和深入分析之后, 就应将所得结果用文字报告的形式表达出来, 以供有关方面参考和利用. 报告中应将研究的目的、运用的方法、工作的经过、结果的检验及未来的推估等择要写明.

应该指出, 统计工作的程序, 从宏观上讲离不开统计立法, 而且统计活动的首要前提是统计立法. 按国际惯例, 统计立法应明确、具体地规定出统计活动所涉及的各方面的法律责任, 以保证统计资料的准确性和客观性. 例如对提供虚假数据的被调查对象, 统计法应具体规定出惩罚条款, 违法必究. 我国统计法 1983 年 12 月 8 日通过, 1996 年 5 月 15 日经第八届人民代表大会常务委员会第 19 次会议修改后重新颁布, 第一章“总则”第三条规定: “国家机关、社会团体、企业事业组织和个体工商户等统计调查对象, 必须依照本法和国家规定, 如实提供统计资料, 不得虚报、瞒报、拒报、迟报, 不得伪造、篡改.”

第二节 社会统计学的几个基本概念

每一门学科都有它自己的专门术语和概念, 统计学也是如此. 本节我们叙述的概念有总体与单位、标志与变量、指标与指标体系, 它们是各门统计学都少不了的基本概念.

1. 总体与单位

统计是大量观察的结果. 统计所研究的不是个别的数量关系而是总体的数量关系, 而没有大量个体的数量特征当然就谈不上总体的数量表现. 所以总体和总体单位是统计学的两个基本概念. 所谓总体, 就是作为统计研究对象的、由许多具有共性的单位构成的整体. 总体也有人称之为母体. 构成总体的每一个个体称为总体单位, 简称单位, 也称为个体.

总体作为统计研究的对象, 显然具有三个基本特征: 大量性、同质性和变异性. 总体必须由足够多的单位所组成, 个别或少数几个单位不足以构成总体. 大量性是统计运用大数定律的前提, 也是统计与其他数学分析方法的一个基本区别. 总体的同质性是构成总体的另一个必要条件. 同质性是相对研究目的而言的, 当研究目的确定之后, 同质性的界限也就确定了. 例如, 我们要研究青年的择业观, 男女青年都属于同质总体; 而如果我们要研究女青年的择偶标准, 男青年就属于“异质”而被排除在总体之外了. 变异性是构成总体的第三个条件, 如果个体之间没有差异, 也就没有必要进行统计研究了.

总体按其包括的单位的数目是有限还是无限, 分成有限总体和无限总体.

社会统计研究的总体一般都是有限总体，如世界人口，数目虽大但总是有限的，可以计数的。无限总体是指在一个合理的时间内不能将全部单位一一列举。如在商场来往不断的顾客人数，在生产线上大量生产的产品，都可以认为是无限总体。在社会调查时，对有限总体既可以进行全面调查，也可以采用各种非全面调查方法；对无限总体则不能做全面调查，只能通过调查其中部分单位来研究它。

总体在抽样调查及推断统计中，还引出了它与样本总体这个概念的联系与区别。推断统计有“部分推断总体”的特征。通过抽样得到的用以推断总体特征的那个“部分”，在统计学上称为样本或样本总体。样本中所含的单位数，在统计学上称为样本大小，也叫做样本容量。从总体抽选出样本的过程叫抽样，也有叫取样的。

在推断统计中，当要研究概率及概率分布时，我们还将接触到现实总体和想像总体这一对概念。对有些问题的研究，总体是由一起看得见、摸得着的个体所构成。对另一些问题研究则不然，它只存在于我们的想像之中。例如手中有十枚硬币，我们做掷硬币的试验，我们做10次、100次、1000次，看看有六枚硬币面朝上的这种结果实现的机会是多少。每一次试验，我们都会得到一个结果（但不一定是六枚硬币面朝上）。只要试验的次数有限，我们面对的都是现实总体。不过我们会发现，试验的次数越多，也就是总体包含的单位数越多，六枚硬币面朝上的频率越来越接近于0.205这个数值。通过想像，我们可以做无数次试验，每一次试验得到一个结果，我们再看看六枚硬币面朝上的可能性有多大，这就过渡到了想像总体。想像总体都是无限总体。在这个想像总体中，我们运用概率论算出，抛十枚硬币得到6枚硬币面朝上结果的可能性为0.205。这个数值不是通过观察得来的，被称为先验概率。先验概率不可能用经验方法准确求得，却可以估计。为了学习和应用推断统计，我们必须根据概率论先验地假定某种概率。然后才说，如果这个先验的概率是可能的，那么某些经验的结果是很可能的（或是不太可能的）。

总体和总体单位的概念是相对于一定的统计研究的目的而言的。随着研究目的的不同，同一事物可以作为总体，也可以当作总体单位。例如，当我们要了解某市一学校学生的健康状况时，该校所有学生便成了统计总体。而当我们要研究该市学校教育的发展状况时，这所学校就成了统计总体的一个单位了。

总体是统计的研究对象，但大量观察却必须从总体中的个别单位开始。而把观察得来的个别事实综合起来，结果所反映的却是整个总体的规律。这就是说，统计所能表示的总体的数量关系虽然代表着这个集合体中的每个个体，但又具有不同于每个个体的特征。用日本统计学家横山雅男的话来讲，“一勺之水无色可辨，而大海则蔚然而紺碧；少量之气无迹可寻，而太空则穹然而