



TONGJIXUE XILIE JIAOCAI

统计学系列教材

国家重点学科·国家精品课程

统计学 (第二版)

Statistics

向蓉美 王青华 主编



西南财经大学出版社



013049698
统计学系列教材

C8-43
144-2

统计学 (第二版)

Statistics

向蓉美 王青华 主编



C8-43

144-2



北航

C1656798



西南财经大学出版社

图书在版编目(CIP)数据

统计学/向蓉美,王青华主编. —2版. —成都:西南财经大学出版社,
2013. 2

ISBN 978 - 7 - 5504 - 0998 - 9

I. ①统… II. ①向…②王… III. ①统计学 IV. ①C8

中国版本图书馆 CIP 数据核字(2013)第 029524 号

统计学(第二版)

向蓉美 王青华 主编

责任编辑:李 雪

封面设计:杨红鹰

责任印制:封俊川

出版发行	西南财经大学出版社(四川省成都市光华村街 55 号)
网 址	http://www. bookcj. com
电子邮件	bookcj@foxmail. com
邮政编码	610074
电 话	028 - 87353785 87352368
照 排	四川胜翔数码印务设计有限公司
印 刷	四川森林印务有限责任公司
成品尺寸	185mm × 260mm
印 张	15
字 数	300 千字
版 次	2013 年 1 月第 2 版
印 次	2013 年 1 月第 1 次印刷
印 数	1—4000 册
书 号	ISBN 978 - 7 - 5504 - 0998 - 9
定 价	29.80 元

1. 版权所有,翻印必究。
2. 如有印刷、装订等差错,可向本社营销部调换。
3. 本书封底无本社数码防伪标志,不得销售。

修订版前言

统计学的作用越来越为人们所认识,对统计知识的学习需求越来越多。《统计学》第一版自 2011 年 1 月出版以来,已三次印刷。这两年来,我国坚持以科学发展为主题,以加快转变经济发展方式为主线,国民经济保持平稳较快发展,各项社会事业不断进步。为了及时反映这些发展和变化,使教材的理论和方法应用更加密切地联系实际,我们对教材做了修订,主要是全面更新了统计数据,在知识的讲解、图表的使用和举例中,都尽可能利用最新的实际社会经济数据。对第一版中个别表述不够准确和印刷有误的文字和公式符号做了修正。

作者

2013 年 1 月

前 言

我们生活在信息时代,数据无处不在、无时不有。哪里有数据,哪里就有统计。

统计学是一门研究数据的方法论学科。宇宙间万事万物,林林总总,各种事物可依特定的性质予以归类,形成各种群体——现象总体。无论是属于自然的、实验的,还是社会的、经济的,凡是可以数据表现的总体,都可以作为统计的研究对象。统计方法和统计思想渗透到了社会、经济、自然、科技、生活的每一个角落,统计学不仅在社会经济领域得到发展,而且一些过去与数量毫无联系的学科,如政治学、法学、历史学、艺术学、考古学等都在对应用统计方法技术进行研究和实践。正因为如此,统计学被教育部经济学和管理学教学指导委员会指定为经济类和管理类各专业必修的核心课程之一。本教材是关于统计学基本理论和方法的教材。适用于经济类和管理类专业学生,也适用于广大经济管理人员作为学习统计学的参考书。

统计学的内容十分丰富。在多年的统计学教学实践中,我们发现很多学生因这门课程中计算公式多、对知识点之间的联系感到难于掌握,对统计学的学习产生畏惧情绪。因此我们重思想、重应用,按照先描述统计、后推断统计的思路编写,把统计知识的学习与EXCEL软件运用有机结合,注重教会学生思考如何收集所需数据、应该采用什么方法去整理和分析数据以及怎样利用计算机软件输出的结果得出结论、进行解释,注重培养学生应用统计知识去分析和解决实际问题的能力。每章后面配有本章小结以及思考与练习题。为了帮助大家更好地学习和把握这门课程,我们还编写了与本教材配套的《统计学学习指导及应用实践》,该书包括“学习目的要求”、“基本知识梳理”、“重点难点点拨”、“范例解析”和“练习与实践”等五个方面的内容。

2007年西南财经大学的统计学学科被评为国家级重点学科,统计学课程被评为国家精品课程。不断提高统计学教材的质量,是我们建设重点学科和精品课程最重要的内容之一。本教材由西南财经大学统计学院和西南石油大学几位长期从事统计学教学的教授和副教授共同编写而成。作者和具体分工如下:

向蓉美,西南财经大学,编写第1章和第4章;苏远琳,西南财经大学,编写第2章;蒲

涛,西南石油大学,编写第3章;雷敏,西南财经大学,编写第5章;夏怡凡,西南财经大学,编写第6章;马丹,西南财经大学,编写第7章;王青华,西南财经大学,编写第8章和第9章;最后由向蓉美和王青华对全书进行了总纂。

随着社会经济的不断发展和统计理论的不断完善,我们将适时根据其发展变化情况修订本教材,也恳请读者将使用本教材的建议和意见及时反馈给我们,对此我们表示衷心的感谢。

作者

2010年10月于西南财经大学

目录

第 1 章 总论	(1)
1.1 为什么要学习统计学	(2)
1.2 什么是统计	(3)
1.3 统计学中的基本概念	(8)
1.4 用 EXCEL 进行统计分析	(14)
第 2 章 统计数据的收集、整理与显示	(19)
2.1 统计数据的收集——统计调查	(20)
2.2 统计数据的整理	(23)
2.3 统计数据的显示	(31)
第 3 章 统计指标	(39)
3.1 总量指标和相对指标	(40)
3.2 平均指标——统计数据集中趋势的描述	(45)
3.3 变异指标——统计数据离中趋势的描述	(54)
3.4 数据分布形态的测定	(58)
3.5 利用 EXCEL 计算数据分布特征指标	(62)
3.6 箱形图	(63)
第 4 章 时间序列分析	(69)
4.1 时间序列概述	(70)
4.2 时间序列的水平分析	(72)
4.3 时间序列的速度分析	(77)
4.4 时间序列构成因素分析	(82)

第5章 统计指数与综合评价	(103)
5.1 统计指数的概念、作用和分类	(104)
5.2 总指数的计算	(106)
5.3 指数体系与因素分析	(115)
5.4 综合评价	(120)
第6章 统计量与抽样分布	(127)
6.1 总体与样本的统计分布	(128)
6.2 统计量	(130)
6.3 抽样分布及抽样分布定理	(132)
第7章 参数估计	(145)
7.1 点估计	(146)
7.2 区间估计	(153)
7.3 抽样估计的进一步讨论	(158)
第8章 假设检验与方差分析	(167)
8.1 假设检验的一般问题	(168)
8.2 一个总体参数的假设检验	(174)
8.3 两个总体参数的假设检验	(178)
8.4 多个总体均值的检验——单因素方差分析	(184)
第9章 相关与回归分析	(193)
9.1 相关与回归分析的概述	(194)
9.2 一元线性相关分析	(198)
9.3 一元线性回归分析	(203)
附录	(218)

第1章

总论

国家教育部把统计学列为经济类、管理类大学本科学专业的专业(核心)基础课。为什么要学习统计学?什么是统计?如何进行统计?在这一章中,我们将学习一些统计学的基本问题。

1.1 为什么要学习统计学

§ 1.1.1 感悟统计

统计数字是量化的历史。

——温家宝为《新中国 60 年》所作的序言《闪光的数据 浓缩的丰碑》

你们籍助于发展成熟的理论和统计分析来创造经济政策和计划的合理基础的贡献,涉及重大科学突破。……我很荣幸地向你们转达瑞典皇家科学院的祝贺,并且请你,丁伯根教授,从国王陛下手中接受 1969 年度阿尔弗雷德·诺贝尔经济学奖金。

——爱立克·伦德伯教授在第一届诺贝尔经济学奖颁奖大会上的讲话^①

在终极的分析中,一切知识都是历史;在抽象的意义下,一切科学都是数学;在理性的基础上,所有的判断都是统计学。

——C. R. Rao《统计与真理》^②

像今天有能力的公民能读会写一样,将来会有一天要求有能力的公民必须会计算,而且能够利用平均值、最大值和最小值。可以预期,这样的时代已经不远了。

——H. G. Wells^③

众所周知,《红楼梦》一书共 120 回,自从胡适作《红楼梦考证》以来,一般都认为前 80 回为曹雪芹所写,后 40 回为高鹗所续,然而长期以来这种看法一直都饱受争议。从 1985 年开始,复旦大学的李贤平教授带领他的学生从统计角度作了考证。一般认为,同一情节大家描述的都差不多,但由于个人写作特点和习惯的不同,所用的虚词是不一样的。他们创造性的想法是将 120 回看成是 120 个样本,然后确定与情节无关的 47 个虚词(之、其、或、亦、呀、吗、咧、罢、可、便、就……)出现的次数(频率),作为《红楼梦》各个回目的标志,利用统计方法果然能将 120 回分成两类,即前 80 回为一类,后 40 回为一类,很形象地证实了《红楼梦》不是出自同一人的手笔。之后又进一步分析前 80 回是否为曹雪芹所写。这时又找了曹雪芹的其他著作,做了类似计算,结果证实了用词手法完全相同,断定前 80 回为曹雪芹一人手笔,而后 40 回不是高鹗一个人所写。这个论证在红学界轰动很大,使红学界大为赞叹。

——红楼梦作者考证^④

① 王宏昌. 诺贝尔经济学奖金获得者讲演集[M]. 北京:中国社会科学出版社,1997.

② C. R. Rao 是当代国际最著名的统计学家之一。C. R. Rao. 统计与真理[M]. 北京:科学出版社,2004.

③ H. G. Wells, 英国作家、历史学家。这段话引自《统计与真理》。

④ <http://www.qintai.net/forum.php?mod=viewthread&tid=6254>

统计的力量是如此的强大、无处不在,哪里有数据,哪里就有统计,统计是现代人谋生的手段,因为我们正处于信息化的时代,数字化信息随处可见。

§ 1.1.2 统计学是一门应用范围很广的科学

无论我们是否学过统计学、懂得统计学,我们生活中的每一天都会遇到大量统计问题,新闻和大众媒体每天都在表现统计数字。例如每年的《政府工作报告》总是要列举大量的水平、比例、结构、速度等数据,说明国民经济的发展状况;统计部门每月、每年都要公布居民消费价格指数(CPI),反映一定时期内城乡居民所购买的生活消费品和服务项目价格变动趋势和程度。确实,“好”、“比较好”、“很好”、“非常非常好”、“比以往任何时候都好”等等这样的比较级或最高级的词语太苍白、太空洞了,而2011年我国国内生产总值达到473 104亿元,按可比价比上年增长9.3%、比1978年增长21.5倍,这样的统计数据实实在在地反映了中国改革开放以来取得的巨大成就。我们生活在信息时代,信息的主体之一就是统计数据。

宇宙间万事万物,林林总总,各种事物可依特定的性质予以归类,形成各种群体——现象总体。无论是属于自然的、实验的,还是社会的、经济的,凡是可以数据表现的总体,都可以作为统计的研究对象。统计方法和统计思想渗透到了社会、经济、自然、科技、生活的每一个角落。统计学不仅在社会经济领域得到发展,而且一些过去似乎与数量毫无联系的学科,如政治学、法学、历史学、艺术学、考古学等,都在对应用统计方法技术进行研究和实践。

1.2 什么是统计

§ 1.2.1 统计的涵义

“统计”作为社会经济生活中经常使用的名词,一般具有三种涵义:统计工作、统计资料和统计科学。

1. 统计工作

统计工作即统计实践活动,是人们为了说明所研究对象的某种数量特征和数量规律性,而对该现象的数据进行收集、整理与分析的活动。例如,为了获得粮食产量而进行的抽样调查活动,为了获得全国人口的数量和构成等而进行的人口普查活动等。

统计实践活动有很久远的历史,早在公元前两千多年,统治者为了征兵、征税、管理奴隶的需要,就有了人口、土地、财产的统计活动。《通典》记载了我们历史上最早的统计数据:“禹平水土,为九州”;“九州之地凡24 388 024顷,人口13 553 923人”。到中世纪,西

欧各国都有了人口、军队、领地、财产等的统计活动。

2. 统计资料

统计资料即统计数据,是通过统计工作所获得的能够说明现象总体某种特征的数据,是统计实践活动的成果。例如2009年我国国内生产总值340 507亿元,按可比价计算比2008年增长9.1%等,这些数据就是广大统计部门和统计工作者辛勤劳动得到的,能够说明我国经济发展水平的统计资料。

统计数据最集中、最系统地反映在各种年鉴中,如《中国统计年鉴》、《国际统计年鉴》、《中国经济年鉴》、《中国金融年鉴》、《中国物价及城镇居民家庭收支调查统计年鉴》等。《中国统计年鉴》涵盖了国民经济所有主要数据,包括国民经济核算、各行业生产、就业、人民生活、对内对外贸易、社会活动、环境等方面的数据。随着互联网技术的发展和普及,网络提供的数据有成为统计数据主要来源的趋势。统计数据还可以从经济分析报告、专著、期刊里获得,比如《中国经济数据分析》、《经济形势分析与预测(经济蓝皮书)》,一些期刊里刊登的最新月度数据等。

3. 统计学

统计学是一门研究收集数据、表现数据、分析数据、解释数据,从而认识现象数量规律的方法论科学。统计学源于统计实践活动,是对统计实践活动的理论概括和总结,又用于指导统计实践活动。

收集数据需要对客观现象作周密细致的调查;表现数据需要对调查得到的数据加以整理,使之成为反映现象总体的条理化、系统化的数据;分析数据需要用科学的方法从数据中得出反映现象本质数量规律性的结果;解释数据需要用有关知识对数量规律性作出说明。所有这些就构成了统计学的研究内容。

统计学是一门为定量分析提供方法的方法论学科,因而是一门应用性很强的学科,它几乎与所有的学科领域都有着或多或少的联系,凡是有数据的地方,就有统计学的用武之地。统计学为其他学科提供研究数量规律性的方法,但是统计学绝不是万能的,各个学科数量规律的解释还需要由各学科的理论来完成。比如,大量观察发现2011年我国人口的性别比,从0~4岁组的119.15:100,到95岁以上组的33.82:100,随着年龄的增大呈下降趋势。形成这样的比例和趋势的原因,不是统计学能够解释的,要用遗传学、医学甚至社会学的理论来解释。又如大量观察表明,吸烟者患肺癌的比例大于不吸烟者患肺癌的比例。吸烟是否会导致患肺癌?为什么会导致患肺癌?这些是医学研究的问题。

尽管统计学不能解决各门学科的所有问题,但是,各门学科离不开统计学,统计学的理论和方法在各门学科的研究中会发挥越来越重要的作用。

作为一门科学,统计学初创于17世纪中叶至18世纪初,当时主要的学派有国势学派和政治算术学派。

国势学派的创始人是德国人赫尔曼·康令。康令1660年以后,开始定期地、系统地

用对比的方法讲授国家之间比较方面的知识,他不仅讲述事实,而且试图探讨事实的因果关系。他把这个课程叫做“欧洲最近国势学”,于是“国势学”由此产生。当时,康令的学说在学术界产生了很大的影响,德国大学中的许多教授都称赞并追随他的学说。其中戈特弗里德·阿亨瓦尔发扬了康令的学术思想,把这门课程定名为“统计学”,开始有了“统计学”这个名称。但是国势学派只是对各国情况作一般性的比较记述,如“某国人口众多”、“土地辽阔”之类,而没有进行数量研究和描述。所以国势学派是有统计学之名,而无统计学之实。

政治算术学派的创始人是英国人威廉·配第。配第首先提出了用数量科学地研究社会经济现象的方法——政治算术。在他的名著《政治算术》序言中,配第明确指出:“我进行这项工作所使用的方法,在目前还不是常见的。因为和只使用比较级或最高级的词语以及单纯作思维的论证相反,我却采用了这样的方法(作为我很久以来就想建立的政治算术的一个范例),即用数字、重量和尺度的词汇来表达我自己想说的内容,只进行诉诸人们感官的论证和考察在性质上有可见的根据的原因。”他还说:“用数字、重量和尺度(它们构成我下面立论的基础)来表示的展望和论旨,都是真实的,即使不真实,也不会有明显的错误。……因为,能够证明为确实的东西,也就是确实的。”他在研究社会经济现象的规律时,还应用推算法、分组法,编制了原始数据的图表,计算了一系列的总量指标、相对指标和平均指标。但是配第始终没有用“统计学”三个字,所以政治算术学派是有统计学之实,而无统计学之名。

近代统计学产生于18世纪末至19世纪末,当时主要的学派是数理统计学派。数理统计学派的代表人物是比利时著名的统计学家、数学家、物理学家、天文学家和人类学家兰贝特·阿道夫·雅克·凯特勒。他融会贯通各家各派的统计思想,博采众长,把统计学发展中的三个主要源泉,即德国的国势学,英国的政治算术和意大利、法国的古典概率论加以协调、统一、改造和融合成为具有近代意义的统计学,把统计学推向了新的发展阶段。他将统计方法用于研究人类,促进了人口统计学的发展;他提出“平均人”理论,用平均值作为实际值的一种代表值;他创建“国际统计会议”组织,促进了国际统计交流与合作。可以说凯特勒是古典统计学的终结者,近代统计学的先驱者,在统计学发展史上起着承先启后、继往开来的作用。

现代统计学一般认为从19世纪末20世纪初开始,其标志是推断统计学的问世。英国生物学家、统计学家卡尔·皮尔森1894—1895年提出包括正态分布、矩形分布、J型分布、U型分布等13种曲线及其方程式。他的这一研究成果,打破了以往次数分布曲线的“唯正态”观念,推进了次数分布曲线理论的发展和应用,为大样本理论奠定了基础;他还提出了卡方检验和标准差概念,发展了相关回归理论等。英国统计学家威廉·西利·戈塞特创立了 t 分布,开创了小样本理论的先河。

统计学的发展史表明,随着社会的发展与实践的需要,统计学越来越多地依赖和吸收

数学方法,使统计方法不断丰富和完善,不断发展和演变,成为研究社会经济现象、自然技术现象数量方面的方法论科学。目前,统计学越来越多地向其他学科领域渗透,形成各种以统计学为基础的边缘学科,随着统计学应用日益广泛和深入,特别是借助电子计算机,统计学所发挥的功效必将日益增强。

§ 1.2.2 统计学的类型

1. 描述统计学

描述统计学是关于如何对现象的数据特征进行观测、整理、计量、表述的理论和方法论科学。其特点是用从一个总体或样本中收集到的数据,来对这个总体或样本进行描述或得出有关这个总体或样本的结论。如全班有 50 名同学或从全校 15 000 名同学中抽取 400 名同学,把这 50 名同学或 400 名同学的成绩用图、表或特征值如平均分数、分数的标准差和及格率等表示出来,从而得出针对该班 50 名同学或 400 名同学学习情况的结论。

2. 推断统计学

推断统计学是关于如何抽取样本并利用样本数据推断总体有关数据的理论和方法论科学。其特点是用从一个总体中随机抽取的样本数据,得出关于这个总体的结论。如从全校 15 000 名学生中,随机抽取 400 名学生进行学习情况调查。用这 400 名学生平均分数、分数的标准差和及格率等推算出全校 15 000 名学生的平均分数、分数的标准差和及格率等,从而得出全校学生学习情况的结论,这就是在进行推断统计。

只有对样本进行了描述,才能对总体进行推断,所以描述统计是推断统计的基础。本教材第 2 章至第 5 章属于描述统计学的基本理论和方法,第 6 章至第 9 章属于推断统计学的基本理论和方法。

§ 1.2.3 统计的特点

1. 数量性

统计的数量性特点指统计总是用数字作为语言来表述事实。统计运用科学的方法收集、表现、分析和解释数据,并用统计指标表明所研究现象的规模、水平、比率、依存度、发展变化趋势和规律等。

但是统计不研究抽象的数量,它是在质的规定性下研究数量。例如,要说明一国的经济的状况,需要统计国内生产总值数据,但是只有明确了国内生产总值的概念和范围,才能得到国内生产总值数据;只有规定了工资总额的内涵和外延,才能得到工资总额数据。

2. 总体性

统计的总体性特点指统计不是研究个别事物的数据,而是研究大量个别事物构成的现象整体的数据,只有这样才能达到认识现象的数量规律的目的。例如,劳动力资源统计,不是要了解个别劳动力的情况,而是要反映一个国家或地区的劳动力资源总数及其构

成、就业总数及其分布等。

但是,统计是从认识个别事物入手来认识现象整体的数量特征的。例如要了解劳动力资源总数及其构成、就业总数及其分布等,必须从每一个劳动力开始,对其性别、年龄、文化程度、职业等进行调查登记,然后经过分类汇总计算,才能了解劳动力资源的整体情况。

§ 1.2.4 统计研究的基本方法

1. 大量观察法

所谓大量观察法指为了对现象整体的数量规律有所了解,必须对所研究现象的全部或足够多的个体进行调查的方法。只有通过大量观察法才能消除偶然的、次要的因素的影响,以反映主要的、共同起作用的因素所呈现的规律性,达到对现象总体数量规律的认识。例如,就个别家庭来说,可能有的男性人口多些,有的女性人口多些,似乎没有什么规律,但是如果观察上千上万个家庭,就会发现人口的性别比例大约为1:1;又如掷硬币和掷骰子,每掷一次出现哪一面或哪个点子是不确定的,但是当我们掷很多次硬币或骰子时,就发现掷硬币出现正面和反面的可能性几乎各为50%,掷骰子出现1~6点的可能性几乎各为1/6。

大量观察法实际上不是指一种具体方法,而是一种统计思想方法,强调观察的总体单位要充分多,只有这样才能将现象的个别偶然差异充分抵消,从而准确地揭示出所研究现象的数量特征和规律性。否则,就可能以偏概全,得到片面的或错误的结论。

2. 统计分组法

所谓统计分组法指根据统计研究的需要,按一定的标志把总体分成若干组别的方法。通过统计分组,突出组与组之间的差异、抽象组内各单位差异,以便划分现象的类型、反映总体的内部结构和现象之间的相互关系。统计分组法贯穿于统计研究和统计工作的全过程。

3. 综合分析法

所谓综合分析法指运用各种经过科学分类汇总的综合指标和各种分析方法,如时间数列分析法、指数分析法、相关回归分析法等,反映现象总体在一定时间、地点、条件下的规模、水平、对比关系、集中趋势、差异程度、依存关系、发展趋势和变化规律等。

4. 归纳推断法

所谓归纳推断法指由个别事物的事实,概括为现象总体的一般特征的推理方法。归纳推断可以使我们从具体的事实中得到一般的知识,扩大知识领域,加深认识程度。社会经济现象是复杂的,常常会出现这样的情况:我们所观察的只是部分单位或有限单位事实,而我们需要分析的却是现象总体的全部单位的事实,这就需要我们z从部分单位的事实归纳推断出现象总体的数量特征。例如调查万分之一的城市居民户的收入水平,推断出

城市全部居民户的收入水平;调查 1‰ 的农田的收获量,推断出上万亩农田的收获量等等。

1.3 统计学中的基本概念

§ 1.3.1 总体、个体与样本

总体是在一定的研究目的下,所要研究事物的全体,它是由客观存在的、具有某种共同性质的众多个别事物构成的整体。^①

构成总体的个别事物是个体或总体单位。个体是所要研究具体问题的承担者。在统计调查中,常常称总体为调查对象,称个体为调查单位。

样本是从总体中抽取的一部分个体的集合,构成样本的个体的数目称为样本容量。从总体中随机抽取一部分个体作为样本,目的是要根据样本提供的有关信息去推断总体的特征。

比如,了解某校学生的学习情况,学习情况具体体现在学生身上,所以全校所有的学生是总体,每一个学生是个体,从全校所有的学生中随机抽取 400 名学生就构成了一个样本,通过 400 名学生的学习情况如平均成绩、及格率等,可以推断全校学生的学习情况;若要研究某市的工业生产情况,工业生产情况具体体现在工业企业身上,该市每一个工业企业是个体,所有的工业企业是总体,从中抽取的若干个工业企业构成一个样本,通过样本工业企业的产值、利润、上缴税金、劳动生产率等,可以推算全市工业产值、利润、上缴税金、劳动生产率等;若要研究某市的工业生产设备情况,工业生产设备情况具体体现在设备上,所以每一台工业生产设备是个体,该市所有的工业生产设备是总体,从中抽取的部分工业生产设备是样本,通过这些设备的净值、生产能力等,可以推算全市所有工业生产设备的净值、生产能力等。

在这些例子中,“学习情况”、“工业生产情况”、“工业生产设备情况”是研究目的,某个学校的学籍、进行工业生产、用于工业生产的设备分别是这些学生、工业企业、工业生产设备的“共同性质”,若干名学生、若干个工业企业、很多很多的工业生产设备分别是“众多个别事物”。

总体具有以下特点:

总体具有同质性。这是指构成总体的总体单位在某一方面性质是相同的,只有性质相同的人、单位、物等才能集合在一起,研究其数量表现和数量联系才有意义。因此,同质

^① 本章的“总体”指实物总体,在推断统计学中,称随机变量为总体(见第 6 章)。

性是构成总体的基础。

总体具有大量性。这是指构成总体的总体单位必须足够多。总体单位是总体数量特征最原始的承担者,总体的数量特征很多时候是无法直接观测到的,只能通过对总体单位的数量特征进行观测得到。而总体单位的数量特征可能各不相同,没有规律可循,只有对大量总体单位的数量特征进行综合,才能体现总体的数量特征。因此,大量性是构成总体的条件。

总体具有差异性。这是指构成总体的总体单位在某一方面性质是相同的,而在其他方面都是不尽相同的。例如上面的例子,如果每一个学生的学习情况都一样,每一个工业企业的生产情况都一样,每一台设备的状况都一样,我们就无需总体了,只要了解一个学生、一个工业企业、一台工业生产设备,就知道了所有的学生的学习情况、所有的工业企业的生产情况、所有的工业生产设备的状况。因此,差异性构成总体的前提。

按构成总体的总体单位是否可以计量,总体可以分为有限总体和无限总体。构成一个总体的总体单位无论有多少,只要其数量是有限的,就是有限总体。如全国人口普查,总体单位多达十几亿人,但它是有限的,是有限总体。构成一个总体的总体单位若是不可数的,即为无限总体。如果没有时间界线,可以把连续生产线的产品产量视为无限总体。社会经济现象绝大多数是有限总体,而在推断统计中,往往把总体视为无限总体。

§ 1.3.2 标志、指标与变量

1. 标志

标志是说明总体单位特征的概念,所以也称为单位标志或单位标识。在统计调查中称为调查项目或登记项目。总体单位具有很多特征,人口调查中,说明每一个人特征的标志有:性别、籍贯、文化程度、婚姻状况等;在工业生产调查中,说明每一个工业企业特征的标志有:所有制性质、职工人数、固定资产数量、产量、利润额等。

标志的具体表现称为标志值。如人口调查中,性别这个标志具体可以表现为男或者女,年龄这个标志具体可以表现为2岁、19岁、85岁等。这里“男”、“女”、“2岁”、“19岁”、“85岁”等是标志值。

根据标志的具体表现不同,标志可以分为品质标志和数量标志。若一个标志的具体表现只能用文字表示,不能用数字表示,则这个标志为品质标志,品质标志表明总体单位的属性。如性别表现为男或女,籍贯表现为北京、上海、成都等,婚姻状况表现为未婚或已婚,所有制性质表现为国有经济、集体经济、股份制经济等。所以,性别、籍贯、婚姻状况、所有制性质等是品质标志。就所研究的问题,有的品质标志只有“是”或者“非”两种表现,则称其为是非标志。如产品合格与否,每一件产品要么合格,要么不合格;家庭是否有电脑,每一个家庭要么有电脑,要么没有电脑等。产品质量、家庭是否有电脑就是是非标志。