



21世纪高等院校经济·管理类规划教材

THE 21ST CENTURY
UNIVERSITIES
ECONOMY MANAGEMENT PLAN

统计学

Statistics

湘潭大学出版社

21世纪高等院校经济·管理类规划教材



THE 21ST CENTURY
UNIVERSITIES
ECONOMY MANAGEMENT PLAN

统计学

梁来存 罗彤◎副主编

湘潭大学出版社

图书在版编目(CIP)数据

统计学 / 郑贵华, 颜泳红主编. — 湘潭: 湘潭大学出版社, 2010.8

ISBN 978-7-81128-193-4

I. ①统… II. ①郑…②颜… III. ①统计学 IV. ①C8

中国版本图书馆 CIP 数据核字 (2010) 第 097366 号

统 计 学

郑贵华 颜泳红 主编

责任编辑: 黎 毅

封面设计: 黄 敏

出版发行: 湘潭大学出版社

社 址: 湖南省湘潭市 湘潭大学出版大楼

电话(传真): 0731-58298966 邮编: 411105

网 址: <http://xtup.xtu.edu.cn>

印 刷: 湖南天闻新华印务邵阳有限公司

经 销: 湖南省新华书店

开 本: 787×1092 1/16

印 张: 19

字 数: 463 千字

版 次: 2010 年 7 月第 1 版 2011 年 1 月第 2 次印刷

书 号: ISBN 978-7-81128-193-4

定 价: 30.80 元

(版权所有 严禁翻印)

21 世纪高等院校经济·管理类规划教材编委会

主 任 田银华
副 主 任 刘长庚 柳思维
编 委 (按姓氏笔画):
 刘友金 汤腊梅 匡 钰 肖东生 肖小勇
 邹安全 杨新荣 欧绍华 周发明 易伟义
 高玉泉 谢 恒 彭晓春 楚尔鸣

总 序

中国改革的目标是建立社会主义市场经济体系，这一历史使命的完成需要千千万万掌握市场经济运行规则的企业家、经济学家以及管理专业人才。中国需要人才，培养人才靠教育，高水平的教育迫切需要有一套与国际接轨的一流教材，尤其是大学本科生的教科书。应现实之需求，我们组织编写了这一套“21 世纪高等院校经济·管理类规划教材”。

本套教科书的适用对象是高等院校经济、管理专业的本科生和研究生，以及具有同等文化程度的自学者。

本套教科书拟推出 18 本，包括：政治经济学、微观经济学、宏观经济学、会计学、统计学、国际贸易学、国际金融学、货币银行学、管理学原理、财务管理、市场营销学、国际经济学、计量经济学、人力资源管理、经济法、物流管理学、现代企业管理、市场信息调查与分析等。基本涵盖了经济、管理类各专业的核心课程。

本套教科书有如下特点：第一，涵盖了现代经济学和管理学的主要领域和许多前沿专题，力图准确、全面、系统地阐述每个学科的基本内容，努力做到现代经济学、管理学教学研究的主流框架与我国经济改革具体实践的有机结合。第二，最大限度地方便读者。各教科书均在前言或第一章中告诉读者该书的层次。教科书的语言简明通俗，结构科学严谨，适合于教学与自学。第三，尽可能地为教员提供方便。这些教科书均配备有相应的教案，以节约教员备课的时间。

本套教科书的作者们来自 10 多所本科院校。他们不仅受过现代经济学、管理学的系统训练，而且他们都从事教学科研工作多年，其中不乏国内外的一流专家教授。他们既熟练掌握现代经济学、管理学理论，又对中国和发达国家经济的实际运行情况有深刻的了解，从而使这套教科书能结合中国国情来系统地、准确地介绍经济、管理理论与案例。

我们诚挚地将此套教科书献给读者，望不吝赐教，以便再版时修正。

田银华

2010 年 3 月

前 言

统计学是一门收集、整理、显示和分析数据的科学。无论是属于自然的、实验的，还是社会的、经济的，凡是可以用数据表现的现象，都可以作为统计学的研究对象。作为数据分析的一种有效工具，统计方法已广泛应用于生产、生活和科学研究的各个领域，成为各学科领域研究者和实际工作者的必备知识。世界发达国家历来都比较重视统计学和统计学教育。美国的高等院校几乎都开设《统计方法》选修课。在国内，统计学也越来越受到重视。1998年，教育部将统计学从原来的二级学科调整为理学类一级学科。2002年教育部高教司又确立了包括统计学在内的九门课程为工商管理类专业的核心课程。目前，统计学已成为全国经济学类和管理学类各专业均必须开设的核心课程。

一个本科生，无论其主修方向是什么，如果能够掌握一些统计数据分析方法，无疑会受益匪浅。本书是根据“技能型、应用型”人才培养目标以及“21世纪高等院校经济·管理类规划教材”的要求，为广大非统计专业（尤其是经济管理类）的本科学生编写的教材。它强调统计学最基本的内容及应用，使读者不但能够系统、准确理解统计学的基本理论、方法和技能，而且能学会如何通过计算机利用 Excel 进行数据处理与分析。

本书具有如下特点：

1. 通俗易懂。在编写过程中，着重直观解释统计学的基本概念，尽量避免引入只有专业统计人员才需要了解的数学推导及定理证明。

2. 科学实用。统计学是一门独立的方法论科学，存在通用的数据搜集、整理和分析的方法。本书在贯彻“大统计”学科建设的基础上，吸收了统计学前沿的最新成果，构建了一个严谨、科学的理论框架，并力求理论联系实际，尽可能用大量最新的实际数据为基础，通过案例分析，展示统计世界丰富多彩的本来面目，让读者觉得统计学并非纸上谈兵，而能学以致用。

3. 统计方法与计算机应用有机结合。每一章都结合实例详细讲解了 Excel2003 软件在统计中的应用，强调通过对 Excel 结果的说明来分析数据，而不只是如何计算，以提高读者分析解释数据的能力，从而提高读者学习统计学的兴趣和应用统计方法分析解决实际问题的能力。

4. 为了便于学生掌握和理解本书的内容，每章开篇有学习目标、结尾有小结，并附有适量配套的思考与练习题。练习题型标准化、规范化，方便学生围绕每章的“学习目标”及教学内容进行课后练习，巩固所学知识，使读者更加易学易练。

本教材以数据的搜集、整理、显示和分析的基本理论和基本方法为主线，以数据分析为核心，以透过数据发现规律为目的，精心组织和编排内容体系。全书共分十章，内容涵盖统计数据的搜集、整理、显示和分析的一般原理和方法，主要包括绪论、统计数据的搜

集、统计数据的整理与显示、统计数据分布特征的度量、时间序列分析、统计指数分析、抽样与抽样分布、参数估计、假设检验、相关与回归分析等内容。

本书由湖南工业大学、湘潭大学、南华大学、湖南工程学院、湖南城市学院、湖南文理学院和邵阳学院等7个学校合作编写。

由郑贵华教授、颜泳红教授担任主编，由梁来存副教授、罗放华教授、罗金华副教授、胡琳副教授、张彤副教授担任副主编。具体编写分工如下：第1章（郑贵华）；第2章（罗金华）；第3章（刘智勇，周雪元）；第4章（单勤琴）；第5章（颜泳红）；第6章（罗放华，朱伟华）；第7-8章（梁来存），第9章（胡琳），第10章（谢卓）。全书有关 Excel 在各章的应用以及章后练习题，除了第7、8、10章外，其余均由颜泳红完成。全书最后由郑贵华、颜泳红总纂定稿。

本书的编写参考了大量文献资料与国内外同行的最新研究成果，在此表示衷心感谢！本书只列出了部分参考文献，若有遗漏，万望见谅！

由于编者水平有限，书中难免有不尽人意之处，恳请同行与读者提出宝贵意见，以便我们进一步修改与完善。

编者

2010年5月

目 录

第一章 绪 论

一、统计与统计学	1
二、统计学的几个基本概念	7
三、统计计算工具	12
本章小结	17
思考与练习	17

第二章 统计数据的搜集

一、统计数据的来源	20
二、统计调查	22
三、统计数据的质量	29
四、Excel 在数据搜集中的应用	31
本章小结	38
思考与练习	38

第三章 统计数据的整理与显示

一、统计数据的整理	41
二、统计数据的显示	52
三、Excel 在数据整理与显示中的应用	60
本章小结	77
思考与练习	78

第四章 统计数据分布特征的度量

一、集中趋势的度量	83
二、离散程度的度量	96
三、偏态和峰态的度量	103
四、Excel 在数据分布特征度量上的应用	105
本章小结	111
思考与练习	111

第五章 时间序列分析

一、时间序列概述	116
二、时间序列的描述性分析	118
三、时间序列的分解分析	128
四、Excel 在时间序列分析中的应用	139

本章小结·····	148
思考与练习·····	149
第六章 统计指数分析	
一、统计指数概述·····	153
二、总指数·····	156
三、指数体系与因素分析·····	165
四、几种常见的经济指数·····	169
五、Excel 在统计指数分析中的应用·····	175
本章小结·····	180
思考与练习·····	181
第七章 抽样与抽样分布	
一、抽样概述·····	186
二、抽样分布·····	193
三、Excel 在抽样与抽样分布中的应用·····	205
本章小结·····	209
思考与练习·····	210
第八章 参数估计	
一、总体参数估计的一般问题·····	213
二、总体参数的区间估计·····	218
三、样本容量的确定·····	224
四、Excel 在参数估计中应用·····	227
本章小结·····	230
思考与练习·····	231
第九章 假设检验	
一、假设检验的一般问题·····	235
二、总体参数的假设检验·····	240
三、Excel 在假设检验中的应用·····	246
本章小结·····	249
思考与练习·····	250
第十章 相关与回归分析	
一、简单线性相关分析·····	253
二、简单线性回归分析·····	262
三、Excel 在简单线性相关与回归分析中的应用·····	273
本章小结·····	280
思考与练习·····	281
附 表 ·····	287
参考文献 ·····	294

第一章 绪 论

学习目标：

1. 了解什么是统计。
2. 明确统计学的研究对象与研究方法。
3. 理解统计学的性质与分科。
4. 熟悉统计学中的基本概念。
5. 对 Excel 处理数据的功能有个初步了解。

一、统计与统计学

（一）什么是统计

在现实生活中，“统计”一词经常出现。然而不同背景的人对统计的理解是有差别的。有的人一提起统计，就必然要说到数据，认为没有数据，统计就成了无源之水，无本之木；有的人一提起统计就想到“计数”，认为没有“计数”这一行为过程，就无法取得数据资料；还有人一提起统计，就想到要对数据及其内在的含义进行研究。

其实，统计一词源于中世纪拉丁文的“status”，意思是各种现象的状态和状况。最早使用“统计”一词的是 18 世纪的德国人——阿亨瓦尔，他把拉丁文“status”引申为德语“statistika”（国势学、统计学），意思是主要用文字来记述国家应注意的事项的学问，后来传入英国被译为“statistics”。英语中的“statistics”以复数形式出现时，其意思为统计活动、统计数据；以单数形式出现时，其意思为统计学。也许正因如此，统计一般包括统计工作、统计资料和统计学三层含义。

1. 统计工作

统计工作即统计实践活动，是人们利用各种科学的统计方法，搜集、整理、分析统计数据和提供统计资料的工作总称。例如，国家和各级省市的统计部门搜集反映其所属地区

的工业、农业、商业、教育以及交通运输业等国民经济部门运行情况的各项数据资料，并将这些资料汇总、加工整理与分析等，这一系列的活动就是统计工作。统计作为一种社会实践活动由来已久，可以追溯到远古时期的“结绳记事”。随着社会、经济的发展，统计的应用领域不断扩大，经验积累得越来越多，客观上为统计学的产生和发展奠定了基础。

2. 统计资料

统计资料亦称为统计信息，是指通过统计工作而得到的各项数据资料（简称统计数据）以及利用这些数据资料所编制的统计图表、进行的统计分析及统计预测等资料的总称。如统计局公布的一年一度的统计公报，省、地、部编制的统计年鉴以及我们经常从电视、报刊、杂志等媒体上听到或看到的数据资料等。随着信息技术的发展与网络的普及，统计资料的公布不再仅仅是纸质资料，大量的电子版的数据可以方便地从官方统计网站上获得，而且大部分都是免费的。如中华人民共和国国家统计局网站（<http://www.stats.gov.cn/>）就以“进度统计数据”、“年度统计数据”、“普查数据”、“专题数据”和“部门数据”五个栏目及时公布许多重要的统计数据。

3. 统计学

统计学是统计工作的经验总结和理论概括，它根据自己的研究对象，系统地阐述了统计的理论和方法。统计学的定义有 180 多个版本，但其基本观点是：统计学是一门搜集、整理、显示和分析统计数据的方法论科学。如何搜集统计数据？怎样整理和显示统计数据？采用什么样的方法分析统计数据以探索数据内在的规律性？解决这些问题的方法就构成了本学科的内容体系。

以上三者既有独立的相对性，更有密切的联系性。第一，统计工作是基础。没有统计工作，就不可能取得统计资料，也不可能获得丰富的统计实践经验，统计科学也无从谈起。第二，统计学与统计工作是理论与实践的关系。统计学是随着统计工作的产生和发展而产生的，是统计工作实践经验的总结和概括，是人们正确从事统计实践活动所必不可少的理论知识，它指导人们如何正确地搜集、整理客观事物的数据资料，如何用数字去描述和分析客观事物的现状和变化过程，预测客观事物发展的前景。由此可见，学好统计学是搞好统计工作的前提。第三，统计工作与统计资料是劳动与成果的关系。统计工作只有依据科学的统计理论和方法，才能取得符合客观实际的统计资料，满足党政领导制定方针政策、企事业单位领导做出决策所需。如果没有科学的统计方法论做指导，在认识社会、管理国家的过程中就不可能取得有价值的统计资料，这样的统计工作非但没有任何意义，而且可能给国家和社会带来不可设想的后果。曾几何时，急功近利的短期行为盛行，弄虚作假的形式主义泛滥，虚报浮夸现象成风，统计数据严重失实，造成党和国家决策失误，最终导致整个国民经济重大比例严重失调，经济濒于崩溃边缘。

在现实生活中，如果只提“统计”二字，而不附加任何词汇，那么“统计”可以是指统计工作，也可以是指统计资料，还可以是指统计学。诸如“干统计”就是指从事统计工作，“学统计”就是学习统计学，“据统计”就是根据统计资料等。但绝大多数情况下，我们所说的统计是指统计工作。

（二）统计学的产生和发展

中国最早的统计工作，起于公元前 21 世纪的夏朝。据《尚书》记载，公元前两千多

年前,统治者为了治理国家,进行过人口和土地统计,出现了“九州”地理区划、“九山九水”治理方案和“上中下三等九级”的贡赋标准。随着社会的进步,统计实践的内容日益丰富,已经渗透到自然、社会、经济和科学技术等各个方面,所涉范围非常广泛,几乎无所不包。

到了17世纪,随着资本主义的发展,统计也适应社会经济的发展而得到广泛的发展,加上数学及其他科学的影响,统计学作为一门社会科学应运而生,并形成了不同的流派。

1. 国势学派

国势学派产生于17世纪的德国,主要代表人物是德国的康令(H. Conring, 1606—1681)和阿亨瓦尔(G. Achenwall, 1719—1772),他们在大学中开设了国势学课程,介绍如何记录国家发展的重要事件。后人把从事这方面研究的德国学者称为国势学派,或记述学派。当时记载关于国家组织、人口、军队、领土、居民职业以及资源财产等事件主要用的是文字叙述的形式,基本上没有量的描述与分析。

国势学派对统计学的最大贡献就是它提出了一个世界公认的名词“统计学(Statistics)”。有人评价该学派是“有统计之名,无统计之实”的学派。

2. 政治算术学派

政治算术学派产生于17世纪的英国,主要代表人物是威廉·配第(William Petty, 1623—1687)和约翰·格朗特(John Graunt, 1626—1687)。配第在其著作《政治算术》中,对当时的英国、荷兰、法国之间的财富进行数量上的计算和比较,开启用数量方法研究社会经济现象之先河。在这个意义上,马克思称配第是“统计学的创始人”。该批学者也因此获得“政治算术学派”之名。格朗特通过对伦敦市50多年的人口出生和死亡资料的计算,写出了第一本关于人口统计的著作《对死亡率公报的自然观察和政治观察》,统计的含义也因此从记述转变为专指用数量来说明国家的重要事件。

政治算术学派对统计学的最大贡献就是它以数量分析为特征,研究客观现象的数量方面。有人评价该学派是“有统计之实,无统计之名”的学派。

3. 数理统计学派

数理统计学派产生于19世纪中叶,最初把古典概率论引进统计学的是法国数学家、统计学家拉普拉斯(P. S. Laplace, 1794—1827),他阐明了统计学的大数法则,进行了大样本推断的尝试。而最终完成统计学和概率论结合的则是比利时统计学家、数学家凯特勒(Adolphe Quetelet, 1796—1874),1841年他出任比利时中央统计委员会会长,1851年积极筹备国际统计学会组织,并任第一届国际统计会议主席。他在著作《社会物理学》中利用大数法则论证了社会生活中的随机偶然现象贯穿着必然的规律性。在此基础上,英国学者高尔顿(F. Galton, 1822—1921)提出了生物统计学,皮尔逊(K. Pearson, 1857—1936)将生物统计一般化而发展为描述统计,埃奇渥斯(F. Y. Edgeworth, 1845—1926)、鲍莱(A. L. Bowley, 1869—1957)则侧重于描述统计在经济领域中的应用和方法研究,费舍尔(R. A. Fisher, 1880—1962)创立了推断统计学。20世纪50年代,又出现了贝叶斯统计学,将统计推断运用于决策问题。

数理统计学派对统计学的最大贡献就是发现了大量现象的统计规律性和开创性地应用了许多统计方法,为统计学的发展奠定了基础。

20 世纪 60 年代以后,随着统计学理论的发展和完善,统计学的方法论已广泛应用于人类活动各个领域,统计学在开拓新知识领域方面的应用,已经远远超过了 20 世纪内的任何技术或科学发明,统计学已逐步成为一个成熟的学科,并为各个领域提供了方法论上的支撑。正因如此,我国教育部社政司原司长顾海良在一次重要会议上说:“语言学、心理学和统计学是 21 世纪中国最有发展前途的三大学科。”

随着计算机技术和网络技术的不断完善和各种新技术的不断创新,统计学的发展有如下趋势:第一,统计学从面对小批量的数据转变为面对海量数据,因此使用计算机统计分析软件对数据进行处理成为必然;第二,统计学已渗透到各个领域,可以应用于各行各业的数据分析,使得统计学成为一门“万能”的方法论学科。

(三) 统计学的研究对象和研究方法

1. 统计学的研究对象

统计学是长期统计工作实践经验的科学总结和理论概括,是搜集、整理和分析统计数据的方法论科学,其目的是探索数据内在的数量规律性。所以,统计学的研究对象是客观事物总体的数量方面。其基本特征如下。

(1) 数量性。数量性是统计学研究对象的基本特征,是统计学区别于其他学科的本质特征。不是计量的研究,不是统计研究。统计学总是在和数据打交道,我们也因此称统计学为“数据的科学”。其数量表现包括:第一,数量多少,如广东省的人口多少、耕地多少等等,以了解现象的规模或水平;第二,各种现象之间的数量关系,如湖南人口的城乡比例、义务教育的普及率、GDP 增长率等,以了解现象之间的相互关系;第三,质与量互变的界限,如江苏省居民生活收入的贫困线、居民家庭恩格尔系数等,以判断事物是否发生了质的变化。统计不仅要研究这些数量的现状,而且要研究其过去和未来的变化。

需要注意的是,统计研究的数量都是客观现象在一定时间、地点、条件下的具体数量表现,而不是抽象的数量。

(2) 总体性。总体性是指统计是从整体上反映和分析事物的数量特征,而不是着眼于个别事物。因为事物的本质和发展规律只有从整体上观察,才能作出正确的判断。例如,要了解湖南省的居民状况,其目的在于了解湖南省居民的总体特征,即湖南省居民的男女比例、年龄构成、文化程度、职业构成、平均收入等等,而不是为了掌握个别湖南居民的具体情况。

2. 统计学的研究方法

(1) 大量观察法。大量观察法是指在客观事物的研究中,从总体出发对其全部单位或足够多数的单位进行观察和分析研究的方法。它是由统计学的研究对象及研究目的决定的。统计学的研究对象是客观现象总体的数量方面。由于构成总体的个体有其特殊性、偶然性。如果只研究其中的几个或少数个体,其结果就不足以说明总体的基本特征。只有对总体中足够的个体进行综合研究,才能消除个体的偶然性和特殊性,反映总体现象的必然性和普遍性。

(2) 统计描述法。统计描述法是对已经搜集到的统计数据进行初步加工、分组、编制统计表、绘制统计图,计算出各种能反映总体数量特征的综合指标并加以分析研究的统

计方法。统计描述法主要通过统计分组法和综合指标法等得到现象总体的数量特征,反映客观事物的内在数量规律性,以达到进行统计分析和研究的目的。统计描述是统计研究的基础,它为统计推断、统计咨询和统计决策提供必要的统计数据资料。

(3) 统计推断法。统计推断法是指以一定的置信水平,根据样本数据资料来判断总体数量特征的归纳推理方法。它是现代统计的基本方法。由于种种主客观方面的原因,通常进行观察的只有部分或有限单位,而需要判断的总体对象范围是大量的,甚至是无限的,这就产生了根据局部的样本数据资料对全部总体数量特征所作判断的问题。这种情况十分普遍,例如,要对一批商品的质量进行破坏性检验,我们只能根据部分商品质量结果来推断该批商品的质量;根据湖南省部分职工家庭(如1000户职工家庭)的平均收入,去推断该省全部职工家庭的平均收入水平。

(4) 统计模型法。统计模型法是根据一定的经济理论和假定条件,用数学方程去模拟现实经济现象相互关系的一种研究方法。它既是人们认识事物的手段,又是人们对事物认识结果的描述。利用这种方法可以对现象状态和变化过程进行数量上的评价、预测和控制。

上面介绍的只是统计学研究的基本方法,在统计学方法体系中,还有许多其他专门方法,如指数分析法、时间序列分析法等。但无论是基本方法还是其他专门方法,在实际研究中都应根据研究目的和任务的需要而结合运用。

(四) 统计学的性质与分科

1. 统计学的性质

统计学是一门独立的方法论学科。它既不属于数学,也不属于其他任何实质性科学。但现代统计学仍应坚持以概率论等数学理论为指导,密切结合有关应用领域的实质性科学进行研究。否则,在统计实践中我们不可能发展新的统计理论和统计方法,也不可能对认识事物本质有任何帮助,甚至有可能导致错误结论。

2. 统计学的分科

历经300多年的发展,统计学目前已经成为横跨社会科学和自然科学领域的多科性的科学。统计学的广泛应用,产生了多种不同的分支,形成了统计学的学科体系。从统计分析方法的研究和应用来分,统计学可以分为理论统计学和应用统计学;从统计方法的构成来分,统计学可以分为描述统计学和推断统计学。

(1) 理论统计学与应用统计学。理论统计学是研究一般的收集数据、整理数据和分析数据方法的理论和统计方法的数学理论。它以数学中的概率论为基础,从纯理论的角度,对统计方法加以推导论证,其中心内容是以归纳方法研究随机变量的一般规律。理论统计学的特点是计量不计质,它具有通用方法论的理学性质。

应用统计学是统计学方法应用于各种实质性科学所产生的一系列专门领域的统计学。如经济统计学、社会统计学、教育统计学、卫生统计学等。应用统计学包括一般统计方法的应用,更包括各自领域实质性科学理论的应用。从所研究的领域或专门问题出发,根据研究对象的性质采用适当的指标体系和统计方法,解决所需研究的问题。应用统计学不仅要进行定量分析,还需要进行定性分析。所以应用统计学通常具有边缘交叉和复合型学科

的性质。

理论统计学和应用统计学总是互相促进，共同提高的。理论统计学的研究为应用统计学提供方法论基础，应用统计学在对统计方法的实际应用中，又常常会对理论统计学提出新的问题，开拓理论统计学的研究领域。

(2) 描述统计学与推断统计学。描述统计学研究如何取得反映客观现象的数据，并通过图表形式对所收集的数据进行加工处理和显示，进而通过综合、概括与分析得出反映客观现象的数量规律性。内容包括统计数据的收集方法、数据的加工处理方法、数据的显示方法、数据分布特征的概括与分析方法等。

推断统计学是研究如何根据样本数据去推断总体数量特征的方法，它是在对样本数据进行描述的基础上，对统计总体的未知数量特征作出以概率形式表述的推断。

描述统计学和推断统计学的划分，一方面反映了统计方法发展的前后两个阶段，同时也反映了应用统计方法探索客观事物数量规律性的不同过程。如图 1-1 所示。

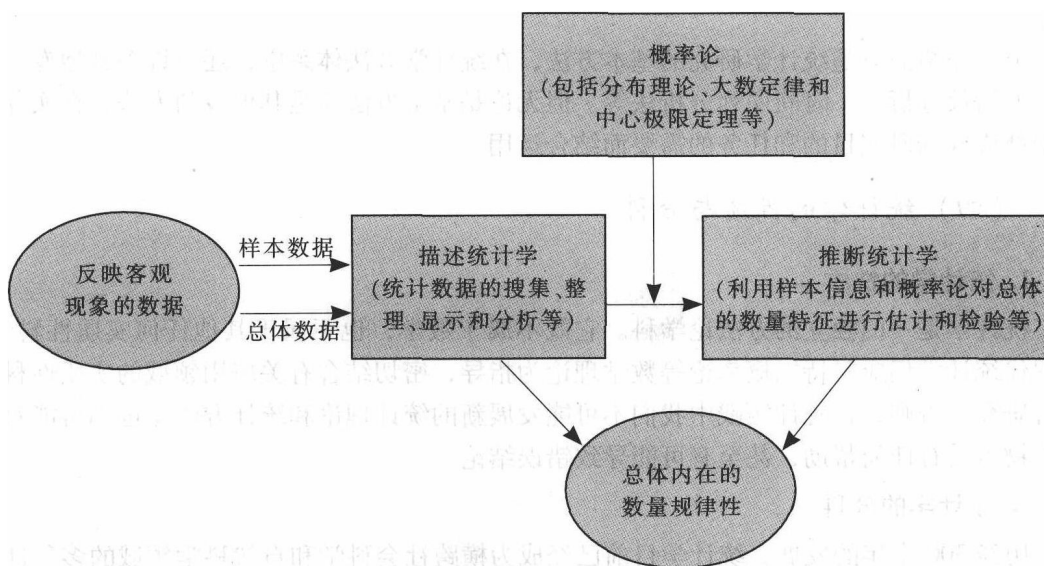


图 1-1 描述统计与推断统计的关系

不难看出，统计研究过程的起点是统计数据，终点是探索出客观现象总体内在的数量规律性。在这一过程中，如果搜集到的是总体数据（如普查数据），则经过描述统计之后就可以达到认识总体数量规律性的目的；如果所获得的只是研究总体的一部分数据（样本数据），要找到总体的数量规律性，则必须应用概率论的理论并根据样本信息对总体进行科学的推断。

显然，描述统计学是整个统计学的基础。由于在对现实问题的研究中，所获得的数据主要是样本数据，因此，推断统计学在现代统计学中的地位和作用越来越重要，已成为统计学的核心内容。当然，这并不等于说描述统计不重要，如果没有描述统计收集可靠的统计数据并提供有效的样本信息，即使再科学的统计推断方法也难以得出切合实际的结论。

本书内容以统计方法为线索展开。第二至第六章是描述统计的内容。第二章介绍统计数据的来源、统计调查方案、统计调查的组织形式以及搜集数据的具体方法；第三章介绍数据整理和显示的方法；第四章介绍数据集中趋势、离散程度及分布形态的测度方法；第五章介绍时间序列的分析方法；第六章介绍统计指数的编制方法，了解常用经济指数。从第七章至第十章都属于推断统计的内容。第七章介绍推断统计的基础理论抽样分布；第八章介绍总体参数的估计方法以及必要样本容量的确定；第九章介绍总体参数的假设检验；第十章介绍根据样本变量的相关关系与简单线性相关模型的建立去估计总体变量的相关关系与简单线性相关模型。

二、统计学的几个基本概念

（一）总体和样本

1. 总体 (Population)

统计总体，简称总体，就是统计所要研究的事物或现象的全体，即由客观存在的、具有某种共同特征的许多个别事物构成的整体。如“所有湖南人”、“所有株洲市工业企业”、“所有长沙市低保家庭”、“所有广州市的乡镇”等都是定义合适的总体。构成总体的个别事物叫个体或总体单位或元素。个体可以是一个人，也可以是一个企业、一个家庭，一个乡镇，还可以是一个时间单位。

在一次特定范围、目的的统计研究中，统计总体与总体单位是固定不变的，是包含与被包含的关系。但是随着统计研究任务、目的及范围的变化，统计总体和总体单位可以相互转化。

统计研究是以个体为基础的，总是从搜集个体数据开始，然后将每个个体的数据进行汇总综合，以反映总体的综合数量特征。这就要求总体必须同时具备以下三个特征：即同质性、大量性和变异性。同质性是指构成总体的个体在某一方面或多方面必须具有相同性质，如“湖南人”总体中的每一个个体必须是湖南省境内常住居民，“株洲工业企业”总体中的每一个体必须是在株洲登记注册、生产工业产品的企业；大量性是指构成总体的个体必须有许多个，如“湖南人”总体由六千多万个体构成，“株洲工业企业”总体由约2000个体组成；变异性是指构成总体的个体必须存在差异，如每一个湖南人在性别、性格、年龄、受教育程度等方面应不完全相同，每一个株洲工业企业在办厂位置、所有制性质、职工人数、生产规模、生产成本等方面不尽相同。有变异才有必要统计。

构成总体的总体单位数目称为总体单位数，通常用“N”表示。如果构成总体的总体单位数目可数，我们称该总体为有限总体。大部分社会经济现象都属于有限总体，总体单位数目一般很大。如人口总体、企业总体、产品总体等都是有限总体。如果构成总体的总体单位数目不可数，我们称该总体为无限总体。在客观现象中，无限总体极少，但也存在，如昼夜连续生产的某产品便构成无限总体。

在现实生活中，大的总体研究起来比较困难，通常的做法是选择一个样本，并研究样本中的数据，然后用样本数据去推断总体数据。那么，什么是样本呢？

2. 样本 (Sample)

样本是指从统计总体中抽取出来作为代表这一总体的、由部分个体组成的集合体。构成样本的个体称为样本单位。构成样本的样本单位数目称样本容量。通常用“ n ”表示,相对于 N 而言, n 一般只是一个很小的数。习惯上,样本容量大等于30的样本称大样本,少于30的样本称小样本。统计研究的样本一般为大样本。比如想了解流水线上产品的合格率,我们可随机抽选100个产品进行检验,这100个产品就构成了一个大样本,样本容量是100。我们可根据对这100个产品检验的结果,计算出样本的合格率,并用它来代表总体的合格率。

总体与样本是整体与局部的关系。如果抽取样本的方式方法科学,样本的代表性就高,就可用样本去推断总体,达到对总体的全面认识。

(二) 标志和指标

1. 标志 (characteristic)

标志是说明个体的属性或特征的具体名称。标志在各个个体的具体表现称为标志值。正是由于各个个体就某些标志而言具有相同的标志值,才构成了统计总体;也正是由于各个个体就某些标志而言具有不相同的标志值,才有了统计的必要。例如,我们要研究北京市下岗职工的生活状况,显然,北京市的下岗职工就构成了我们要研究的总体,每个下岗职工是个体。反映个体的标志可以有很多:户籍、家庭住址、性别、民族、年龄、是否在岗、下岗年限、生活来源、家庭收入等等。对每个北京市的下岗职工而言,她们的“户籍”和“是否在岗”所对应的标志值都是相同的,这也是构成总体的条件——同质性;而其他标志的标志值就不会完全相同了,这就是我们要调查研究的内容。

(1) 标志按其性质可分为品质标志和数量标志。品质标志是说明个体属性特征的名称,其标志值只能是文字,但可以赋予数字代码。如北京市下岗职工的“户籍”、“家庭住址”、“性别”、“民族”等,其中“性别”标志的具体表现为“男、女”,在统计实践中,我们可以用“0”表示“男”,用“1”表示“女”。数量标志是说明个体数量特征的名称,其标志值通常为数值。如北京市下岗职工的“年龄”、“下岗年限”、“家庭收入”等,其中下岗职工的年龄可能是35岁、40岁、45岁、50岁等随机数值。

(2) 标志按其标志值是否可变分为不变标志和可变标志。不变标志是指所有个体具有相同标志值的标志。如北京市下岗职工的“户籍”和“是否在岗”。可变标志是指所有个体具有不完全相同标志值的标志,如北京市下岗职工的“家庭住址”、“性别”、“年龄”、“下岗年限”、“生活来源”、“家庭收入”等。

(3) 是非标志。只有两种标志值的品质标志称为是非标志。如“性别”、“是否合格”、“是否成功”、“是否购买”等。其标志值只能是“男、女”,“合格、不合格”,“成功、失败”,“有购买、无购买”。这种标志也叫交替标志或01标志。

2. 指标 (indicator)

指标是反映现象总体数量特征的概念及其具体数值。任何一个统计指标都是经过了从个别到一般、从具体到抽象的过程,它体现总体特征,具有综合性。一个完整的统计指标应该由指标名称、指标数值、时间、地点和计量单位等要素构成。如2008年末中国农民