



广东省精品资源共享课程教材

MANAGEMENT STATISTICS

管理统计学

主 编 张卫国

副主编 刘小龙 徐维军



华南理工大学出版社
SOUTH CHINA UNIVERSITY OF TECHNOLOGY PRESS

广东省精品资源共享课程教材

MANAGEMENT
STATISTICS
管理统计学

主 编 张卫国

副主编 刘小龙 徐维军



华南理工大学出版社
SOUTH CHINA UNIVERSITY OF TECHNOLOGY PRESS

·广州·

内容简介

本书是“华南理工大学精品课程建设”和“广东省精品资源共享课程建设”的研究成果，是在参阅了大量的国内外优秀统计学教科书的基础上，紧扣当前实际发生的管理和经济问题特点，结合近年来我们所承担的管理统计学课程的教学体会编写而成的。

本书的编写遵循以经济管理活动中的实际问题为先导，通过现实需求揭示蕴含的统计学思想、统计学理论与方法，循序渐进地详细展开各章教学内容，使得统计学理论和方法的学习与解决管理实践问题紧密结合。同时，本书精选数据较新、覆盖较全的案例进行章节贯穿和综合应用，以便读者在理解的同时能够学以致用。全书共分为九章，按照数据收集、整理、分析和解释的逻辑顺序，分别介绍统计基本概念、统计数据描述、抽样与抽样分布、参数估计、假设检验、方差分析与试验设计、相关与回归分析、时间序列、统计指数等基本内容。其目的是利用统计学理论和方法对管理实践活动进行分析，以探索管理问题和统计数据的内在数量规律性。

本书既具有系统的统计学理论知识，又具有超强的实践指导训练，能够很好地满足高校本科生以及相关专业人员在案例教学、数据分析方面的需要，可作为高等学校经济学类、管理类专业本科生统计学课程的教材，也可作为人文及其他学科的教材或辅导书，还可以作为从事统计工作的人员和研究人员的参考书。

图书在版编目(CIP)数据

管理统计学/张卫国主编. —广州: 华南理工大学出版社, 2014. 2
ISBN 978 - 7 - 5623 - 4155 - 0

I. ①管… II. ①张… III. ①经济统计学 - 高等学校 - 教材 IV. ①F222

中国版本图书馆 CIP 数据核字(2014)第 022789 号

管理统计学

张卫国 主编

出 版 人: 韩中伟

出版发行: 华南理工大学出版社

(广州五山华南理工大学 17 号楼, 邮编 510640)

<http://www.scutpress.com.cn> E-mail: scute13@scut.edu.cn

营销部电话: 020 - 87113487 87111048 (传真)

责任编辑: 张 颖

技术编辑: 杨小丽

印 刷 者: 佛山市浩文彩色印刷有限公司

开 本: 787mm × 1092mm 1/16 印张: 21.25 字数: 548 千

版 次: 2014 年 2 月第 1 版 2014 年 2 月第 1 次印刷

印 数: 1 ~ 1000 册

定 价: 38.00 元

版权所有 盗版必究 印装差错 负责调换

前言

统计学是工商管理、管理科学与工程等学科及相关专业的一门重要的理论与方法基础课。它作为收集、整理、归纳和分析反映客观数据的方法论科学,其目的是探索数据的内在数量规律性,以达到对客观事物的科学认识。

统计学是一门关于数据的学科,统计学的理论与方法越来越被广泛应用于多个学科领域,如管理、经济、社会、工程、物理、化学、生物等,用于分析和解决有关随机性数据的问题,这就对统计方法的应用及教材的编写提出了更高的要求。基于统计学课程的重要性和广泛应用,我们在广泛参阅近几年来国内外出版的优秀著作和教材的基础上,编写了该教材。它既可以作为高等学校经济类、管理类专业本科生统计学课程的教材,也适用于人文及其他学科,还可供从事统计工作的人员和研究人员学习参考。在此,特别感谢袁卫、庞皓、曾五一、贾俊平、马军海、张建同、陈珍珍等老师所编著的统计学相关教材,为我们的教材编写提供了素材。

本教材深入浅出、简明易懂,尽量避免统计方法的数学推导和证明,更加侧重于统计思想的介绍和统计方法的应用,并通过 Excel 软件辅助复杂的统计计算。通过本课程的学习,将使读者掌握统计数据内在的数量规律性,获得对统计数据进行分析的基础和前提,掌握统计分析的基本方法,结合计算机软件和应用统计方法提高分析与解决实际问题的能力。为了帮助读者掌握书中的概念,在每章的开头和结尾给出了实际案例和分析,开头的案例导出了本章节要讨论和学习的问题,结尾部分则是对本章综合知识的复习和检验。为了加深对本章的理解,本书还给出了本章小结和相关的思考题。标题中带“*”的为选讲内容。

本书的显著特点是:1. 突出以实际案例贯穿章节内容的特色,强调统计学基本思想、原理和方法在管理和经济实际领域中的应用,力图使初学者通过系统学习本课程,掌握解决实际应用问题的统计方法和技巧。2. 突出解决现实经济管理问题的特色,使用来源于现实生活中的统计数据讲解,体现所学统计理论和方法的应用价值,提高学生学习的兴趣,注重培养学生形成合理有效的思维方式。3. 突出应用 Excel 等信息技术软件工具进行统计分析的特色,结合翔实生动、通俗易懂的案例分析和数据求解,培养学生的动手和实践操作能力。

前言

本书末附录给出了概率论相关知识的补充,有助于读者回忆概率论的内容,更好地掌握统计学思想和方法。另外,本书还为广大读者制作了与本书配套的试题库、教学课件和课程网站资源,有助于读者更好地掌握管理统计学。

本教材由张卫国教授担任主编,刘小龙博士和徐维军教授担任副主编,全书由张卫国、刘小龙和徐维军统稿修订。参与本书编写的人员如下:徐维军、张卫国、刘小龙、李牧南、周文慧、肖炜麟、苏越良、刘勇和徐禾芳。在教材的数次修改过程中,研究生罗伟强、刘幼珠、周平平、刘桂芳等同学全程参与。在本书付印之际,谨向所有帮助和支持本书编写和出版的同志表示衷心的感谢。

由于编者水平有限,书中难免有疏漏或错误之处,恳请同行和读者多提宝贵意见,以便我们进一步修改和完善。

编 者
2013 年 12 月

目录

第一章 绪论	1	第二章 统计数据的描述	18
1.1 统计及应用领域	2	2.1 统计数据的类型	19
1.1.1 统计学的定义	2	2.1.1 按计量尺度划分	19
1.1.2 统计学的应用领域	3	2.1.2 按收集方法划分	20
1.2 统计学的分科	4	2.1.3 按时间状况划分	21
1.2.1 描述统计和推断统计	5	2.2 统计数据的来源	21
1.2.2 理论统计和应用统计	6	2.2.1 统计数据的间接来源	21
1.3 统计学的基本概念	7	2.2.2 统计数据的直接来源	23
1.3.1 总体和样本	7	2.3 统计数据的质量	25
1.3.2 参数和统计量	8	2.3.1 统计数据的质量要求	25
1.3.3 指标	9	2.3.2 统计数据的误差	25
1.3.4 变量	10	2.3.3 统计误差的控制	26
1.4 几种典型的统计软件	10	2.4 统计数据的整理	27
1.4.1 Excel	10	2.4.1 统计数据整理概述	27
1.4.2 SPSS	11	2.4.2 统计分组	27
1.4.3 SAS	12	2.4.3 分配数列	30
1.5 统计学简史	13	2.4.4 洛伦茨曲线和基尼系数	33
1.5.1 古典统计学(17世纪中叶至18世纪末)	13	2.5 分布集中趋势的测度	35
1.5.2 近代统计学(19世纪初至20世纪初)	14	2.5.1 众数	36
1.5.3 现代统计学(20世纪初到20世纪中叶)	15	2.5.2 中位数	38
1.5.4 当代统计学(20世纪中叶到现在)	15	2.5.3 算术平均数	40
本章小结	16	2.5.4 调和平均数	43
思考题	17	2.5.5 几何平均数	44
		2.5.6 众数、中位数和算术平均数的比较	45
		2.5.7 四分位数	46
		2.5.8 切尾均值	47

目录

2.6 分布离散程度的测度	47	3.4.2 分层抽样	85
2.6.1 极差	48	3.4.3 簇群抽样	86
2.6.2 内距	48	3.4.4 等距抽样	87
2.6.3 平均差	49	本章小结	87
2.6.4 方差与标准差	50	思考题	88
2.6.5 变异系数	52	案例分析	88
2.7 分布偏态与峰度的测度	52	第四章 参数估计	90
2.7.1 偏态	52	4.1 点估计	91
2.7.2 峰度	54	4.1.1 点估计介绍	91
2.8 统计表与统计图	55	4.1.2 点估计量的求法	91
2.8.1 统计表	55	4.2 评价估计量的标准	96
2.8.2 统计图	59	4.2.1 无偏性	96
2.8.3 雷达图	68	4.2.2 有效性	98
本章小结	68	4.2.3 一致性(相合性)	99
思考题	69	4.3 区间估计	101
案例分析	69	4.3.1 置信区间	101
第三章 抽样与抽样分布	72	4.3.2 单个总体参数的区间估计	103
3.1 抽样的概念和特点	72	4.3.3 两个总体参数的区间估计	109
3.2 抽样分布	73	4.3.4 样本容量的确定	117
3.2.1 基本概念	73	4.4 Excel 辅助参数估计	119
3.2.2 样本均值 $\bar{X}$ 的抽样分布形式	75	4.4.1 总体方差未知的情况	119
3.2.3 单个样本统计量的抽样分布	77	4.4.2 总体方差已知的情况	120
3.2.4 两个样本统计量的抽样分布	80	4.4.3 总体比例(成数)区间估计	121
3.2.5 常用的几个概率分布的关系	81	本章小结	122
3.3 中心极限定理的应用	83	思考题	122
3.4 常用的抽样方法	84	案例分析: 淘宝销售	122
3.4.1 简单随机抽样	85		

目录

第五章 假设检验	126	6.2 单因素方差分析	167
5.1 假设检验的基本问题	127	6.2.1 数据结构	168
5.1.1 假设的陈述	127	6.2.2 分析步骤	168
5.1.2 两类错误与显著性水平	129	6.2.3 应用实例分析	171
5.1.3 检验统计量与拒绝域	131	6.2.4 关系强度的测量	172
5.1.4 利用 P 值进行决策	131	6.2.5 用 Excel 进行单因素方差分析	172
5.1.5 统计显著与实际显著性	133	6.2.6 方差分析中的多重比较	174
5.1.6 假设检验结论的表述	134	6.3 双因素方差分析	176
5.2 单个总体的参数检验	135	6.3.1 无交互作用的双因素方差分析	176
5.2.1 总体均值的检验	135	6.3.2 有交互作用的双因素方差分析	181
5.2.2 置信区间和假设检验	140	6.4 试验设计*	186
5.2.3 总体比例的检验	141	6.4.1 完全随机化设计	186
5.2.4 总体方差的检验	143	6.4.2 随机化区间设计	187
5.3 两个总体参数的检验	144	6.4.3 因子设计	189
5.3.1 两个总体均值之差的检验	144	6.4.4 正交设计	190
5.3.2 两个总体比例之差的检验	154	本章小结	195
5.3.3 两个总体方差比的检验	156	思考题	196
本章小结	159	案例分析	196
思考题	160	第七章 相关与回归分析	198
案例分析: 女企业家与男企业家	160	7.1 相关分析	199
第六章 方差分析与试验设计	162	7.1.1 相关分析的概念及原理	199
6.1 方差分析引论	163	7.1.2 相关关系的测定	200
6.1.1 方差分析的基本概念	163	7.1.3 相关系数的检验	204
6.1.2 方差分析的基本假定	163	7.1.4 相关分析中应注意的问题	204
6.1.3 方差分析前提假定检验及破坏	164		
6.1.4 方差分析的基本思想和原理	167		

目录

7.2 一元线性回归	205	8.3 时间序列趋势变动分析	238
7.2.1 回归分析	205	8.3.1 测定长期趋势的移动平均法 ...	238
7.2.2 一元线性回归模型	207	8.3.2 测定长期趋势的线性趋势模型法	242
7.2.3 一元线性回归的参数估计	208	8.3.3 测定长期趋势的非线性趋势模型法	244
7.2.4 一元回归方程的检验	210	8.4 季节变动分析	249
7.2.5 一元回归方程的预测区间	214	8.4.1 季节变动分析的意义	249
7.3 多元线性回归分析	217	8.4.2 季节变动分析的方法	249
7.3.1 多元线性回归模型	217	8.4.3 季节变动的调整	252
7.3.2 多元线性回归模型的参数估计	219	8.5 循环变动分析	253
7.3.3 对多元线性回归方程的评价 ...	219	8.5.1 循环变动及其测定目的	253
7.4 非线性回归模型	223	8.5.2 循环变动的测定方法	254
7.4.1 可线性化的常用曲线类型	223	本章小结	258
7.4.2 非线性判定系数	225	思考题	259
7.4.3 虚拟自变量回归	225	案例分析：中国社会消费品零售总额分析 ...	259
本章小结	226	第九章 统计指数	262
思考题	227	9.1 统计指数的概念与分类	263
案例分析	227	9.1.1 统计指数的概念	263
第八章 时间序列分析	229	9.1.2 统计指数的性质	263
8.1 时间序列的对比分析	230	9.1.3 统计指数的种类	264
8.1.1 时间序列的概念	230	9.1.4 统计指数的作用	265
8.1.2 时间序列的速度分析	231	9.1.5 总指数编制方式	265
8.2 时间序列及其构成因素	235	9.2 综合指数	266
8.2.1 时间序列的趋势因素	235	9.2.1 数量指标综合指数	266
8.2.2 时间序列构成因素的组合适型	237	9.2.2 质量指标综合指数	268

目录

9.2.3 其他形式的综合指数	270	F.2.1 基本概念	296
9.2.4 综合指数的特点	272	F.2.2 离散概率分布	298
9.3 平均指数	272	F.2.3 连续概率分布	299
9.3.1 加权算术平均指数	272	F.3 概率的基本定义与性质	299
9.3.2 加权调和平均指数	274	F.3.1 事件的运算关系与概率的定义	299
9.4 指数体系因素分析	275	F.3.2 三种计算事件概率的方法	301
9.4.1 指数体系	275	F.3.3 概率的基本性质	305
9.4.2 因素分析	276	F.3.4 概率的加法原理	306
9.5 几种常见的经济指数	284	F.4 条件概率	307
9.5.1 居民消费价格指数	284	F.4.1 条件概率的定义与性质	307
9.5.2 工业生产指数	286	F.4.2 独立事件	309
9.5.3 股票价格指数	287	F.4.3 概率的乘法原理	311
本章小结	291	F.5 贝叶斯定理	312
思考题	292	F.5.1 基本概念	312
案例分析：“国民幸福指数”和“国民 生产总值”同样重要	292	F.5.2 贝叶斯公式	314
附录1 概率论基础	295	附录2 常用统计表及正交表	316
F.1 概率论简介	295	参考文献	330
F.2 事件与概率	296		

第一章 绪论

电影消费群体的研究

电影产业在中国的发展历程主要有三个阶段。第一个阶段是在 20 世纪初,那时电影刚刚进入中国,电影市场是以纯粹商品性为特征,是一种简单的文化消费,电影消费者主要来源于中产阶层。第二个阶段是新中国成立之后,当时的电影市场是在垄断机制下建立的,电影成为宣传国家主流意识的一种手段,工人、农民、干部、知识分子以及一般劳动者等社会主要阶层是电影消费的对象。第三个阶段是 20 世纪 80 年代之后,中国处于市场经济转型过程中,电影被定义为精神文化产品,特别是 20 世纪 90 年代电影体制改革的深化,使得电影消费成为一种特殊的文化消费。这种将电影商品属性过度挖掘使得票价提高,消费者流失,因此电影消费者主要是中高收入者,其构成主要来自于城市中的企业白领、大学生、党政机关干部、知识分子、文艺工作者、一般从业人员以及其他较高收入者。

如今,电影在全球已经进入了多媒体竞争时代,我国电影也逐渐淡化其意识形态,回归娱乐本性,成为文化产业的重要支柱。中国电影市场票房连续六年保持两位数增幅的快速增长态势,成为全球最具发展潜力的电影市场。据统计,我国电影市场票房 2007 年为 33.27 亿元人民币,2008 年 43.41 亿元,2009 年 62.06 亿元,2010 年突破百亿大关,达到 101.72 亿元,2011 年更达 131.15 亿元。这意味着 2011 年中国电影市场的范围仅次于美国和日本,排名第三。上映 18 天的国产电影《泰囧》以 9.4 亿元的票房成为最卖座的电影,观影人次近 3000 万,超过《阿凡达》《变形金刚 3》。这一连串数字彰显着我国电影业发展的潜力和希望。

“我们拥有世界上最广泛的受众人群,这是中国电影产业发展最雄厚的基础。”国家新闻出版广电总局电影局局长童刚说。对电影消费群体进行分析有利于电影投资者制作符合大众消费者口味的电影,有利于行政部门制定相关政策引导我国电影业朝着健康方向发展。对电影消费群体的分析须有数据支持,如何获得想要的的数据,并对所获得的数据进行加工处理,提炼出所需要的信息,如电影观众的收入分布和年龄分布、电影观众主要选择何种观看方式、电影观众收入与票房是否存在联系等,而这些数据分析与统计学息息相关。

本章将介绍统计学的一些基本知识,包括统计学的含义、统计应用领域、统计学的分科、统计中常用的基本概念、几种典型的统计软件和统计学简史等。

1.1 统计及应用领域

1.1.1 统计学的定义

每当提到统计,人们总是联想起计数、数字和数据等。在人们的生活、工作和学习中,数字和数据无处不在。小至一个人、一个家庭,大至一个国家都有需要处理的各种数据。一个月的收入与支出、一年的利润等都是我们关心的数据。世界各国都有相应的政府统计机构专门收集和处理关于社会各方面的数据。统计学是一门关于数据的学科,由于数据无处不在,因此统计学领域宽广、分支多,统计学家们给统计学下的定义也众多。维基百科全书上说:“统计学是在统计实践的基础上,自17世纪中叶产生并反映客观现象总体数量的学科。”它是研究如何收集、整理、归纳和分析反映客观现象数量的数据,以便给出正确认识的方法论科学,被广泛应用于各学科中,甚至被用于工商业及政府的情报决策上。统计学另一个权威定义来自大英百科全书:统计学是关于收集和分析数据的科学和艺术。简单对统计学给出一个普遍赞同的定义是困难的,也是没有必要的。更多的统计理论学习和统计应用的实践会让读者对统计学有个基本清晰准确的认识。

统计学目前已成为单独的一级学科,是一门研究随机现象,以推断为特征的方法论科学,“由部分推及全体”的思想贯穿于统计学的始终。具体地说,它是关于收集、整理、分析和解释数据,并从数据中得出有用结论的一门广泛应用的学科,具有科学性兼少部分艺术性的科学。

统计学的定义告诉我们如下事实,首先,统计学是关于数据的学科。这个数据不是数学中抽象的数和形的数据,而是实际发生的实测、观测、试验等数据。因此,统计学是以实用性为出发点和目标。对于数据收集通常有两种方法,一种是人为地控制某些相关因素来对我们感兴趣的研究对象做试验,另一种是直接观察研究对象。通常得到的原始数据是纷繁复杂的,甚至夹杂着一些混乱,因此必须对数据进行整理,然后深入分析,挖掘数据中所含有的信息和找出对应的原因及机制,并最终得到一些结论为将来的实践和预测提供指导,这也是数据模拟的目的。其次,统计学是以归纳推理为研究方法的学科,归纳推理是由个别到一般的研究方法。统计问题中所要处理的数据,往往只是研究对象中的部分个体,但我们却要用统计方法来对所有的研究对象做出判断和得出一些有用结论,这就是归纳推理的方法。当然,这些归纳推理方法必须建立在概率论等理论的基础上,数据本身有很大的随机性,而概率论等则是处理随机性的数学理论。再次,统计学是一门具有广泛用途的学科。统计学是关于数据的学科,数据无处不在,它充满在我们的工作、生活、学习和其他任何人类实践与生产活动之中。统计学具有极其广泛的研究领域,理、工、农、医、艺、文、体几乎没有不用统计学的学科。最后,统计学不完全是—门科学,它带有少部分艺术特征。科学有它的准确性,它的推理和推导是严格正确的。由于数据在收集过程中带有各种各样的随机性,后续数据的整理、分析、解释和总结都需要有科学系统的理论和方法来处理这种随机性,这种科学系统的理论和方法就是概率论及其相关的数学理论方法。在大多数情况下,统计学是严格的和精确的,但由于数据的复杂性以及统计方法的简单实用性要求,统计学在处理数据的复杂性以及构造简单实用统计方法时经常要求它自身

的直观合理性以及简洁性，并且有时还带有统计学家对直观合理性与简洁性的个人理解和选择，所以说统计学是带有部分艺术特征的关于数据的科学。

用统计来认识事物的一般步骤是：研究设计→抽样调查→统计推断→结论。其中，研究设计就是对实际问题制订调查和实验研究的计划，抽样调查是收集资料、获取数据的过程，统计推断是处理数据、分析数据、解释数据的过程，其统计的研究过程如图 1-1 所示。显然统计的主要功能是推断，用部分资料来推断总体，所以其推断的方法是一种不完全归纳法。

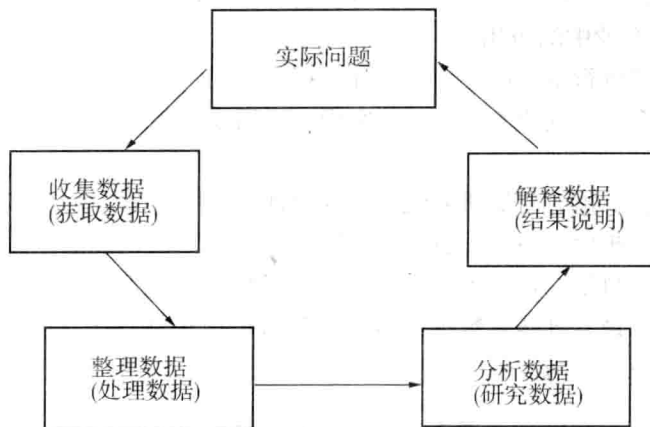


图 1-1 统计的研究过程

1.1.2 统计学的应用领域

统计学是应用性很强的学科，已被广泛应用于各个专业领域（如物理、化学、工程、生物、经济、社会等）中有关随机性数据问题的分析。可以说，几乎所有的研究领域都要用到统计方法，如产品生产工艺流程、医学研究、企业经营管理、人口普查等，下面将给出统计学在生活中的一些应用。

1. 统计方法在工业生产中的应用

一是应用正交设计、回归分析、方差分析、多元统计分析等统计方法解决试制新产品和改进老产品、使用代用原材料和寻找适当的配方、判定影响产品质量的重要因素、决定最优的生产条件等问题；二是应用统计质量管理方法，通过质量控制图、抽样检验、可靠性统计分析以解决工业生产中工序控制、成批产品的抽样验收、大批生产元件可靠性检验等问题。

2. 统计方法在医学、药学中的应用

在医学研究中，统计方法常被作为重要的研究工具，用来发现和验证导致某种疾病的种种因素。如应用统计方法证实肺癌与吸烟的关系。在药学研究中，通过临床试验，应用正交设计、交叉设计、回归分析、方差分析、列联表分析等统计方法，来确定一种药物对治疗某种疾病是否有效，用处多大，以及比较几种药物或治疗方案的效果。

3. 统计方法在市场研究中的应用

市场研究是为某一特定的市场营销问题的决策而开发和提供其所需信息的一种系统

的、有目的的活动或过程。这里所说的信息，不仅是市场调查所得的数据资料，还包括市场研究人员对资料进行分析所得的结果(如结论、建议等)。市场研究的范围包括：产品研究、销售研究、市场与销售潜量的估计、价格研究、购买行为研究、竞争分析、广告及促销研究、销售成本和利润分析、营销环境研究。

4. 统计方法在社会、经济领域中的应用

在西方发达国家，统计方法在这些领域中的应用要比其在自然科学和技术领域中的应用更早且更广泛。如社会学中的抽样调查、列联表分析，人口学中的人口发展动态模型、随机过程统计，经济学中的经济计量模型等社会现象的定量化研究。

5. 统计方法在农业中的应用

通过对田间试验进行适当的设计和统计分析，得到收获量最大的方案。另一应用是数量遗传学，如在培育高产优质农产品的研究中，应用复杂的回归分析和方差分析的方法来计算遗传力。

在应用性研究中，由于对所研究现象的规律性认识不充分，人们不得不依靠对实验和观测数据进行统计分析(统计推断与统计预测)，以提出解决问题的办法。一般而言，无论是自然科学还是技术科学，都离不开实验观察，都有处理数据的问题，因此也就有统计方法的用武之地。可以说，凡是有数据的地方，就可以应用数理统计方法。表 1-1 给出了统计学一些相关的应用领域。

表 1-1 统计学的应用领域

animal science(动物学)	agriculture(农业)	geology(地质学)
genetics(遗传学)	archaeology(考古学)	auditing(审计学)
epidemiology(流行病学)	econometrics(经济计量学)	ecology(生态学)
dentistry(牙医学)	actuarial work(精算)	geography(地理学)
anthropology(人类学)	finance(金融)	industry(工业)
ophthalmology(眼科学)	education(教育学)	linguistics(语言学)
pharmaceutics(制药学)	engineering(工程)	sociology(社会学)
psychology(心理学)	physics(物理学)	taxonomy(分类学)
demography(人口统计学)	literature(文学)	crystallography(晶体学)
meteorology(气象学)	hydrology(水文学)	marketing(市场营销学)
psychophysics(心理物理学)	political science(政治学)	survey sampling(调查抽样)
military science(军事科学)	religious studies(宗教研究)	management science(管理科学)
medical diagnosis(医学诊断)	quality control(质量控制)	human genetics(人类遗传学)
gambling(赌博)	historical research(历史研究)	manpower planning(劳动力计划)

1.2 统计学的分科

统计学的内容十分丰富，随着人们对定量研究的日益重视，统计方法已被广泛应用于自然科学和社会科学的众多领域，统计学也已发展成为由若干分支组成的学科体系。从统

计方法的构成来看,统计学可分成描述统计学和推断统计学;从统计方法的研究和应用角度看,统计学可分为理论统计学和应用统计学。

1.2.1 描述统计和推断统计

描述统计学(descriptive statistics)是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规律性数量特征的一门学科。描述统计学的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。图1-2表明描述统计是对数据进行处理的第一阶段,即用直观的图表和概括性的数字表示数据的分布、形状等特征,并为进一步的统计推断提供依据。

推断统计学(inferential statistic)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,对总体进行估计、假设检验、预测或其他推断,进而对统计总体的未知数量特征做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。二者关系如图1-2所示。

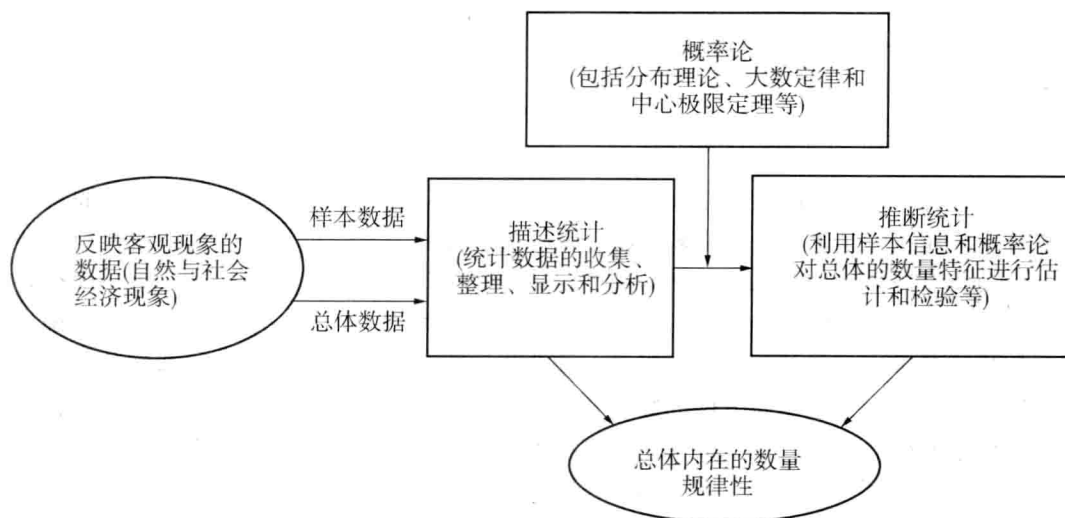


图1-2 统计学探索客观现象数量规律的过程

统计研究过程的起点是统计数据,终点是探索出客观现象内在的数量规律性。在这一过程中,如果收集到的是总体数据(如全国人口普查数据),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,根据样本信息对总体的数量特征进行科学的估计和检验等推断。

显然,描述统计和推断统计是统计方法的两个组成部分。描述统计是整个统计学的基础,它包括对客观现象的度量,科学经济地收集与整理数据,确定观测对象的数量特征,科学度量和描述其数量特征等;推断统计则是现代统计学的主要内容,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,选用恰当的统计量和推断方法,科学准确地探索总体内在的数量规律性至关重要。因此,推断统计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和统计研究工作的关键环节。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。因而,推断统计对描述统计又有很强的依赖性。从描述统计学发展到推断统计学,既反映了统计学发展的巨大成就,也是统计学发展成熟的重要标志。

应该认识到,尽管描述统计可以在获得总体数据时直接探索出总体数量的规律性,但这种情况对实际工作中有无限总体的自然现象来说是不太可能的。而尽管社会现象的总体多数为有限,但考虑到获得数据及推断总体的经济性、时效性和准确性,往往抽样调查方式比普查更普遍、有效。例如,全国人口调查,全国及各省、市、地区的普查可以摸清基本情况,获得丰富的统计数据;但普查涉及千家万户,所花费的时间、人力、财力和物力都极其巨大,因而只能间隔较长时间进行一次,而两次普查之间的年份以抽样调查方法获得连续的统计数据。此外,社会中大量的管理和研究工作不可能都进行全面普查,如某种商品的市场销售量调查、电影消费群体调查等,只能通过抽样调查的方式对总体数量规律性进行科学的推断。因而,设计合适的调查方法、科学获取观测对象样本数据、图表显示数据的特征和规律、分析和提取数据中的有用信息以最终推断总体特征就显得格外重要了。

1.2.2 理论统计和应用统计

理论统计学(theoretical statistics)是统计学的一个分支,主要研究如何将数学原理和计算机技术应用于统计学,发展出新的统计学方法和技术的数学理论。它是把研究对象一般化、抽象化,以概率论为基础,从纯理论的角度对统计方法加以推导论证,中心内容是统计推断问题,实质是以归纳方法研究随机变量的一般规律。例如,统计分布理论,统计估计和假设检验理论,相关与回归分析,方差分析,时间序列分析,随机过程理论等。由于现代统计学用到了几乎所有方面的数学知识,所以从事统计理论和方法研究的人员需要有扎实的数学基础。此外,由于概率论是统计推断的数学和理论基础,因而广义地讲统计学也是应该包括概率论在内的。理论统计学是统计方法的理论基础,没有理论统计学的发展,统计学也不可能发展成为像今天这样一个完善的学科体系。

应用统计学(applied statistics)主要是研究如何将统计学的方法和原理应用于自然、经济、工程等实际领域,解决实际问题,它是统计学和其他学科之间形成的交叉学科,也是理论统计学发展的源泉。统计学是一门收集和分析数据的科学。由于在自然科学及社会科学研究领域中,都需要通过数据分析来解决实际问题,因而统计方法的应用几乎扩展到了所有的经济和社会科学研究领域。例如,统计方法在物理研究中的应用形成了统计物理;在医学中的应用形成了医疗卫生统计学;在管理领域的应用形成了管理统计学;在经济领域的应用形成了经济统计学;在社会学研究和社管理中的应用形成了社会统计学,等等。以上这些应用统计学的不同分支所应用的基本统计方法都是一样的,即都是描述统计和推断统计的主要方法。但由于各应用领域都有其特殊性,统计方法在应用中又形成了一些不同的特点,如经济学中的统计指数、现代企业中的管理决策等。

1.3 统计学的基本概念

在开始学统计学时需要理解几个重要的概念,它们对课程的学习和理解至关重要,有必要单独加以介绍。这些概念包括总体和样本、参数和统计量、指标和变量等。

1.3.1 总体和样本

1.3.1.1 总体

总体(population)是根据一定的目的和要求所确定的研究对象的全体。它是由客观存在的、具有某种共同性质的许多元素构成的整体,将其中构成总体的每一个元素称为总体单位,简称个体。在这里,我们将总体单位具有的某种属性或数量单位统称为标志。例如,考察京东商城各种品牌冰箱的销售量,总体就是京东商城的全部冰箱,每一种冰箱品牌就是总体单位,而冰箱的销售量称为标志;研究某个学院教工的身高和体重,总体就是这个学院的全体教工,个体就是学院里的每个教工,教工的身高和体重则是标志;考察某企业员工的学历分布,总体就是该企业员工,个体就是企业的每个员工,标志是学历。

标志是统计调查的具体项目,也是统计调查所要采集的实际资料内容。对于标志,按照其性质的不同可以细分为品质标志和数量标志,标志的填写内容称其为标志表现。那些用文字描述的属性特征,一般称其为品质标志;那些用数字描述的数量特征,一般称其为数量标志。比如姓名和性别等标志,其标志值填写属于文字,这类标志是品质标志;而年龄和月工资的标志值填写属于数值,这类标志就是数量标志。

准确地界定总体对后续的统计研究和实践显然是非常重要的。总体范围的确定有时比较容易,如要了解一批出口汽车的质量时,则该批所有出口汽车就是总体,汽车的质量级别就是标志。但有些场合总体范围的确定则比较困难,比如,新推出的一种产品,要想知道消费者是否喜欢,首先必须弄清哪些人是消费者,也就是要界定该产品的消费者这一总体,但事实上这一总体范围的确定十分复杂。显然,正确地界定总体,需要具备问题所涉及领域的专业知识和工作经验。当总体的范围难以确定时,可根据研究的目的来定义合适的总体。

1. 统计总体的特点

在明确了以上一些基本概念之后,将它们联系起来观察,深入地认识总体,可以看出,统计总体具有同质性、大量性和差异性三个主要特点。

(1)同质性。同质性是指总体中的各个个体必须具有某种共同的属性或标志数值。如国有企业总体中每个企业的共同属性是国家所有。同质性是总体的根本特征,只有个体单位是同质的,统计才能通过对个体特征的观察研究,归纳和揭示出总体的综合特征和规律性。

(2)大量性。大量性是指总体中包括的总体单位有足够多的数量。总体是由许多个体在某一相同性质基础上结合起来的整体,个别或很少几个单位不能构成总体。总体的大量性可使个别单位某些偶然因素的影响——表现在数量上的偏高、偏低的差异相互抵消,从而显示出总体的本质和规律性。