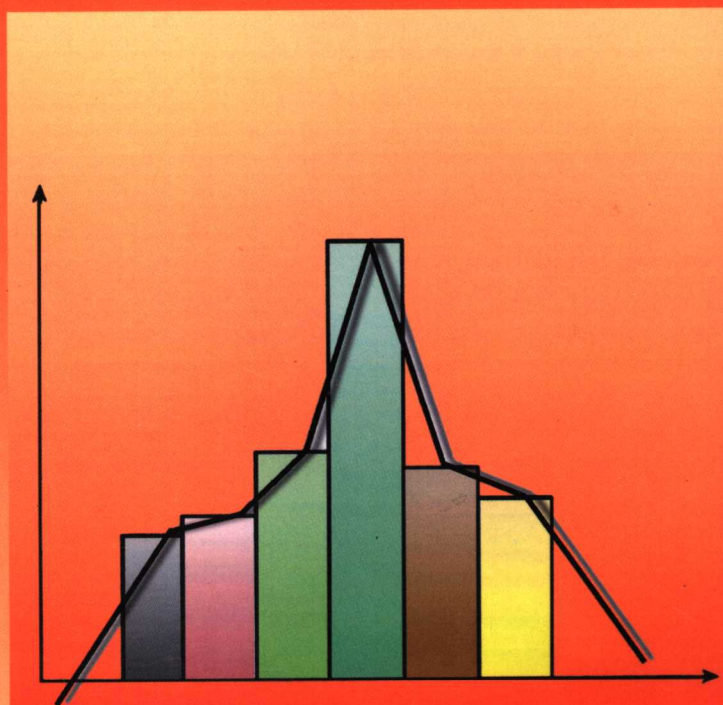


高等学校适用教材

# 应用统计学

(第二版)

陶靖轩 刘春雨 鲁统宇 叶祥北 编著



中国计量出版社

高等学校适用教材

---

# 应用统计学

(第二版)

陶靖轩 刘春雨 编著  
鲁统宇 叶祥北

中国计量出版社

## 图书在版编目(CIP)数据

应用统计学/陶靖轩等编著. —2 版. —北京:中国计量出版社, 2007. 6

高等学校适用教材

ISBN 978-7-5026-2658-7

I. 应… II. 陶… III. 应用统计学—高等学校—教材 IV. C8

中国版本图书馆 CIP 数据核字 (2007) 第 071822 号

## 内 容 提 要

本书共有 13 部分内容: 总论、统计工作基本过程、综合指标、时间序列、统计指数、相关分析与回归分析、方差分析与正交试验设计、抽样调查与推断、统计预测、统计决策、统计方法在微机上的实现、概率论基础、参数估计与假设检验。书后还附有 7 个常用的统计用表和习题参考答案。

本书可作为高等院校管理类专业的教材, 也可供统计人员和有关管理人员自学和参考。

中国计量出版社出版

北京和平里西街甲 2 号

邮政编码 100013

电话(010)64275360

E-mail jilfxb@263.net.cn

北京市密东印刷有限公司印刷

新华书店北京发行所发行

版权所有 不得翻印

\*

787 mm×1092 mm 16 开本 印张 21.25 字数 515 千字

2007 年 7 月第 2 版 2007 年 7 月第 4 次印刷

\*

印数 9 001—12 000 定价: 33.00 元

## 再版前言

《应用统计学》第二版与大家见面了。它与第一版有什么区别呢?

首先,作者总结了三年来各方面读者和专家提出的宝贵意见,从基本概念的阐述、基本算法的推导和基本数据资料的更新上做了比较仔细的修订。

在“总体和样本”这个统计学基本概念的表述中,我们认识到总体必定存在未知的需要推断的东西,比如分布或参数。因此,才在必要时抽取样本。否则,一个总体的一切都清楚了,就不需要再抽取样本。而样本是可观测的,利用观测样本的信息推断总体的统计特性是统计学的主要内容之一。所以,新版将摒弃“全及总体”,“样本总体”这些容易使人产生混淆的说法和概念。

在第四章时间序列里增加平均增长量的累计计算方法,这种算法可以比较完整地利用数列提供的信息。在时间序列的性质识别上,对计算自相关系数的传统近似算法进行了说明,因为这种近似算法是无计算机时代的一种便捷算法,但往往会导致错误的判断。所以我们在保留原方法的同时,特意说明不提倡用近似计算的原因(关于时间序列分析的注记——《数学的实践与认识》)。

在数据资料的引用中,利用2006年出版的统计年鉴,更换了一些陈旧的资料,增加了一些新的例子和习题。

第五章指数分析中增加循环波动的测定以及不规则波动的测定,以便对时间序列的影响因素的分析更为全面。

对第七章方差分析与试验设计的部分习题和例题进行了更新,通过这些例题和习题的练习,学习与环境保护等国民经济发展有密切联系的分析方法,可以增强学员利用方差分析和试验设计解决实际问题的兴趣和能力。

第十一章的“统计方法在计算机上的实现”,考虑到目前SPSS的流行,将原先的EXCEL语言统统更换为SPSS,并编写了新的实验教学大纲和指导书。

其他各章内容也做了比较系统的修订和更新,同时增添了一些习题。

国内以及原教材介绍统计方法,大量引用经济学数据,这当然与统计学的起源与研究国家的国情、国势、国力有关系,同时也是因为经济数据与我们的密切联系,所以无可厚非。但统计学也应用于其他观测到的数据,如测量误差、质量特性值、实验结果数据等等。在解释数据内在变异和分布时,这些数据更加具备统计规律性。也就是说应用统计也非常适宜解决这些方面的问题。这在国外统计学书籍上比比皆是。所以本书准备在介绍原始数据时,更多称作“观测值”或“观测数据”,研究问题的对象既有经济数据,更包括其他观测数据。

需要说明的是，这次修订增添了新的力量。第一、二、五、八章仍由刘春雨老师负责；第三、四两章由叶祥北老师负责；第六、十一章由鲁统宇老师负责（其中第十一章为重新编写内容）；第七、九、十、及附录一、附录二的五章由陶靖轩老师负责。全书仍由陶靖轩老师担任主编和负责统稿工作。

当然，本书（第二版）存在的问题必然还会不少，我们衷心希望各位读者和专家继续提出宝贵建议和意见，以便再版时修改。

编著者

2007年6月

# 前 言

## (第一版)

应用统计学是高等学校管理类专业主要课程之一,为适应这门课程各类教学需要,曾经出版了多种版本的统计学教材。尤其在我国社会主义市场经济体制不断发展和完善的形势下,统计体制的改革势在必行,加之近年来统计学术界对一些基本理论问题的反思和重新认识,统计学教材也面临革故鼎新的任务,出现了不少新的版本。我国自1984年1月1日开始实施《统计法》,1996年5月中旬全国人大常委会对《统计法》重新进行了修订,其中重要的改动之一就是确定了“统计调查应以周期性普查为基础,以经常性抽样调查为主体,以必要的统计报表、重点调查、综合分析为补充,搜集、整理基本统计资料”,从而以法律形式确定了科学的统计调查方法。1995年8月于北京召开的国际统计学会第五十次会议,与会代表1800多人,交流中外学术论文1300多篇,许多优秀的数理统计成果在会议上进行了交流,极大地促进了“大统计”学科的组建。此后,我国统计界运作了几十年的政府统计部门与蓬勃发展的数理统计学术团体共同组建了统计科学联合会,并在1995年底全国统计专业技术中级资格甲种考试前,新编了统计基础理论和实务统计分析的考试用书,明显加大了数理统计内容。

客观地说,在此之前我国已有不少学者认识到数理统计理论及方法与社会统计的密切联系,在不少新版的统计学教材中加入了“概率论基础”、“参数估计与假设检验”等数理统计内容。但作为教材,由于受教学体系(比如大一、大二讲《概率统计》,已学习过这部分内容)和学时的限制,加之统计学教材引入数理统计内容时间不长,使得这些内容尚未完全与原有的统计学内容融为一体。针对这种情况,我们在本教材中以附录形式写入两章数理统计内容,以便学员随时参考和学习;教师则可以根据实际情况灵活掌握,决定是否讲授这些内容。

鉴于统计运算与大量数据处理的紧密联系,本书在最后一章加入了“统计方法在微机上的实现”一章内容,学员在学习统计知识的同时可学习各种统计软件 and 数据处理方法,以便更加卓有成效地进行统计分析。这里有一个问题应该说明:按照《不列颠百科全书》的定义,统计学是“收集和分析数据的艺术”。这“数据”是带有随机性的数据,即受统计规律的支配,这一特点使统计学与诸如数值分析之类的学科区分开来。因此,尽管数值计算技术(尤其计算机技术)的发展对统计学的发展起到了巨大的作用(比如验证计算结果及一些统计推理的微机实现),但也不应将统计学简单地理解为是纯粹的数值计算。

此外,目前在管理学科大门类里还细分为许多专业门类或方向(如质量管理、财务管理、行政管理等),不同门类或方向对统计学知识有不同的特殊要求,如质量管理方向要求学员掌握方差分析、试验设计、抽样检验、过程控制、回归分析、可靠性基础与计量管理等

知识。为适应这一需要,我们在本书加入了“方差分析与正交试验设计”一章。不同类型的专业统计课程可根据需要选择讲授。

我国已于2001年11月加入了世界贸易组织(WTO),中国经济将进一步与世界经济融合。世界要进一步了解中国,中国也要进一步了解世界。作为宏观决策依据的统计,也要适应中国加入WTO这一新形势的需要,改革现行统计,特别是统计制度和方法要进一步与国际接轨,加强统计数据的国际可比性,更好地为宏观和微观主体服务。同时,要依法规范政府的统计行为,其主要的工作是:(1)改进和完善我国国民经济核算体系;(2)进一步推进中国官方统计与国际接轨;(3)加快建立、完善国际服务贸易统计;(4)改进有关统计指标体系。为此,我们在教材的编写中,尽量注意吸收反映我国加入WTO以后社会主义市场经济体制的社会背景的理论概念和统计方法。

本书所指“应用统计学”不同于数理统计。在西方,数理统计仅是统计学的数学理论部分。本书也不同于前苏联体制下的统计学。在前苏联,把统计学定义为一门研究大量社会现象的社会科学。本书的宗旨是编写研究现象总体数量关系的方法论科学著作。这种现象既包括自然现象,也包括社会现象。这就将它从实质性科学中分离出来,即它不是反映特定对象本身及研究现象本身的发展规律,而是采用适当的统计方法研究现象总体数量之间关系的方法论学科。本书的特点是:(1)适当引入统计推断内容,具有一定理论性;(2)密切结合我国社会主义市场经济体制建设的实际;(3)注意吸收国内外基本理论和基本方法的最新成就;(4)文字简练、语言通俗,具有较强的可读性。它既可以作为学习统计学的基础理论读物,又可以作为统计专业的基础课教材。

全书共分13部分,内容有总论、统计工作基本过程、综合指标、时间序列、统计指数、相关分析与回归分析、方差分析与正交试验设计、抽样调查与推断、统计预测、统计决策、统计方法在微机上的实现、概率论基础、参数估计与假设检验。其中第一、二、四、五、十一章由刘春雨同志编写;第三、六、七章及附录1由王敏华同志编写;第八、九、十章及附录2由陶靖轩同志编写;习题及答案由各章作者拟出。全书由陶靖轩同志负责统稿。

本书是我们在多年的统计学教学实践过程中,不断总结过去的教学经验和教训写作而成。各位作者除了自己的讲义和笔记外,还参考了不少中外统计学基础著作及学术研究成果,在一些统计学术会议上与同行专家的交流也使我们深受启发。在此,对所有给予我们帮助的专家和同仁,表示衷心的感谢!

限于编者的水平,书中讹误之处一定不少,盼望得到读者宝贵的批评和指教,以便再版时修正。

编著者

2003年1月

# 目 录

|                         |         |
|-------------------------|---------|
| 第一章 总 论 .....           | ( 1 )   |
| 第一节 统计学及其产生与发展 .....    | ( 1 )   |
| 第二节 统计学的特点与方法 .....     | ( 4 )   |
| 第三节 统计学基本概念 .....       | ( 6 )   |
| 习题一 .....               | ( 11 )  |
| 第二章 统计工作基本过程 .....      | ( 12 )  |
| 第一节 统计设计 .....          | ( 12 )  |
| 第二节 统计调查 .....          | ( 14 )  |
| 第三节 统计整理 .....          | ( 20 )  |
| 第四节 统计资料的表现形式 .....     | ( 29 )  |
| 第五节 统计分析 .....          | ( 34 )  |
| 习题二 .....               | ( 36 )  |
| 第三章 综合指标 .....          | ( 37 )  |
| 第一节 总量指标 (绝对数) .....    | ( 37 )  |
| 第二节 相对指标 (相对数) .....    | ( 39 )  |
| 第三节 平均指标 (平均数) .....    | ( 45 )  |
| 第四节 变异指标 .....          | ( 56 )  |
| 第五节 标准差的应用 .....        | ( 60 )  |
| 习题三 .....               | ( 62 )  |
| 第四章 时间序列 .....          | ( 65 )  |
| 第一节 时间序列概述 .....        | ( 65 )  |
| 第二节 动态比较分析 .....        | ( 68 )  |
| 第三节 动态平均分析 .....        | ( 71 )  |
| 第四节 长期趋势的测定 .....       | ( 79 )  |
| 第五节 季节波动的测定 .....       | ( 90 )  |
| 第六节 循环变动和不规则变动的测定 ..... | ( 95 )  |
| 习题四 .....               | ( 97 )  |
| 第五章 指数分析 .....          | ( 99 )  |
| 第一节 统计指数概述 .....        | ( 99 )  |
| 第二节 综合指数 .....          | ( 101 ) |
| 第三节 平均数指数 .....         | ( 107 ) |

|      |                      |       |
|------|----------------------|-------|
| 第四节  | 总平均水平指数 .....        | (110) |
| 第五节  | 指数体系与因素分析 .....      | (113) |
| 第六节  | 指数的应用 .....          | (119) |
| 习题五  | .....                | (124) |
| 第六章  | 相关分析与回归分析 .....      | (127) |
| 第一节  | 相关分析 .....           | (127) |
| 第二节  | 回归分析的基本概念 .....      | (131) |
| 第三节  | 一元线性回归模型 .....       | (133) |
| 第四节  | 多元线性回归模型 .....       | (140) |
| 第五节  | 非线性回归模型 .....        | (145) |
| 习题六  | .....                | (149) |
| 第七章  | 方差分析与正交试验设计 .....    | (151) |
| 第一节  | 方差分析 .....           | (151) |
| 第二节  | 正交试验设计 .....         | (162) |
| 习题七  | .....                | (174) |
| 第八章  | 抽样调查与推断 .....        | (177) |
| 第一节  | 抽样推断概述 .....         | (177) |
| 第二节  | 抽样调查的组织方式和抽样方法 ..... | (181) |
| 第三节  | 抽样误差 .....           | (185) |
| 第四节  | 抽样估计 .....           | (190) |
| 第五节  | 样本容量的确定 .....        | (196) |
| 习题八  | .....                | (198) |
| 第九章  | 统计预测 .....           | (200) |
| 第一节  | 统计预测概述 .....         | (200) |
| 第二节  | 简单模型预测 .....         | (205) |
| 第三节  | 长期趋势预测 .....         | (207) |
| 第四节  | 季节变动预测 .....         | (213) |
| 第五节  | 回归预测 .....           | (217) |
| 习题九  | .....                | (222) |
| 第十章  | 统计决策 .....           | (224) |
| 第一节  | 统计决策的意义和分类 .....     | (224) |
| 第二节  | 风险型决策方法 .....        | (226) |
| 第三节  | 不肯定型决策方法 .....       | (232) |
| 第四节  | 效用理论在决策分析中的应用 .....  | (236) |
| 习题十  | .....                | (240) |
| 第十一章 | 统计方法在微机上的实现 .....    | (242) |
| 第一节  | SPSS 软件概述 .....      | (242) |
| 第二节  | 数据文件的建立与数据的录入 .....  | (245) |

|                         |       |
|-------------------------|-------|
| 第三节 单变量的描述统计分析 .....    | (247) |
| 第四节 两个变量之间的线性相关分析 ..... | (253) |
| 第五节 线性回归分析 .....        | (256) |
| 第六节 统计图 .....           | (265) |
| 附录 1 概率论基础 .....        | (269) |
| 第一节 概 率 .....           | (269) |
| 第二节 概率分布 .....          | (275) |
| 第三节 常用离散型随机变量分布 .....   | (279) |
| 第四节 常用连续型随机变量分布 .....   | (282) |
| 附录 1 习题 .....           | (285) |
| 附录 2 参数估计与假设检验 .....    | (287) |
| 第一节 对未知参数的点估计 .....     | (287) |
| 第二节 假设检验的一般方法 .....     | (291) |
| 第三节 常用参数的统计假设检验 .....   | (293) |
| 第四节 参数的区间估计 .....       | (296) |
| 第五节 非参数检验 .....         | (299) |
| 附录 2 习题 .....           | (305) |
| 附录 3 常用统计用表 .....       | (307) |
| 习题参考答案 .....            | (325) |
| 参考文献 .....              | (330) |

# 第一章 总 论

## 第一节 统计学及其产生与发展

### 一、“统计”的含义

统计的英文单词是 Statistics，它包括三个方面的含义：统计工作（或统计活动）、统计资料和统计学。这已成为国际上通行的有关统计的习惯术语。

#### 1. 统计工作

统计作为一种社会实践活动有着悠久的历史，它是适应国家宏观管理和社会经济发展的需要而产生和发展起来的。统计工作是指国家行政机关（主要是统计机构）、企事业单位为满足社会、经济、政治、科技等方面管理需要或从事科学研究的需要，而对社会经济现象的数量方面进行搜集、整理和分析有关数据的一系列统计活动过程。统计活动的一般过程包括统计设计、统计调查、统计整理、统计分析等。

#### 2. 统计资料

统计资料是指统计工作过程中所产生的原始的或加工过的统计数据、统计报表、统计图表、统计分析报告，以及与之相联系的其他资料的总称。对统计资料通常有如下要求。

- ① 客观性。即统计资料必须能够反映现实而不受任何偏见的影响和任何势力的干扰。
- ② 准确性。即由客观性所要求，统计数据的偏差不能超出统计目的事先确定的允许误差范围。
- ③ 及时性。即统计资料为满足统计目的的需要，应及时搜集、及时加工、及时发表。
- ④ 连续性。统计资料应能在时间上提供动态对比的数据。
- ⑤ 系统性。统计资料在时间上应能以客观现象之间的内在联系为基础，各项数据相互之间也应保持这种内在联系，并能相互结合或对比运用。

#### 3. 统计学

统计学是指系统地阐述统计实践活动基本原理和研究方法的理论。它是一门研究如何搜集、整理和分析统计资料的理论和方法论科学。统计方法是社会、经济、科学技术等各研究领域和工作部门必要的基本的数量分析手段，是从大量数据资料中提取主要的有用信息的重要工具。

统计的三种含义密切相关：统计工作和统计资料是过程和结果的关系；统计资料是统计工作达到的成果或结晶，它来源于统计工作又服务于统计工作；统计工作和统计学是理论与

实践的关系；统计学是统计活动发展到一定阶段的产物，是统计实践经验的理论概括，并用于指导统计实践，为统计工作提供理论依据；统计工作的发展会不断完善和丰富统计理论，而统计理论的丰富又促进了统计工作的发展，使之更有效、更为科学。

以上对“统计”一词含义的理解，有助于对统计学的初步认识。但这还不够。要真正了解统计学这一门科学的性质、特点、功能，还要从其产生的历史来考察。

## 二、统计学的产生与发展

一切科学都来源于实践，统计这门科学的产生和发展也是如此，它是适应社会经济和科学技术的发展的需要而产生和发展起来的。

统计的产生有着很悠久的历史。最早的统计活动可以追溯到原始社会末期，当时的人们打猎捕鱼、结绳记事等活动，就孕育了统计的最初含义。在奴隶社会和封建社会里，统治阶层要巩固其对内统治和对外战争，加强对国家的宏观管理，就要进行征兵、课税，了解和掌握军队、农产品、世袭领地和国民财富等国情国力的资料。于是就产生了与人口、土地和国民财富等有关的统计调查。由于生产力发展缓慢，当时统计还只是简单意义上的登记数字和比较，还没有被当作一门学问去研究，尚处于萌芽阶段。

据历史记载，我国古代的一些政治家、军事家早就意识到统计的重要性。秦国的商鞅在《商君书》的《去强篇》中这样说：“强国当知十三数：境内仓、口（府）之数，壮男、壮女之数，老、弱之数，官士之数，以言说取食者之数，利民之数，马、牛、刍、藁之数。欲强国，不知国十三数，地虽利，民虽众，国愈弱至削。”2000多年前，春秋战国时代的统治阶级就已认识到统计是一种重要的认识工具和管理工具，它对战争的胜负、国家的安危有着重要意义。秦汉之时，已有地方田亩和户口记录；唐宋时则有计口授田田亩鱼鳞册等土地调查和计算；到明清时代，已经建立了经常的人口登记和保甲制度等。

同样在欧洲，古希腊和古罗马时代，就已开始了人口数字和居民财产的统计调查工作。封建主国家根据其需要也进行了有关人口、军队、世袭领地和财产等统计。但是那个时代的统计活动和统计资料，无论中国还是外国，都是一些原始的、简单的汇总和计算。

随着资本主义的兴起和社会生产的不断发展，社会分工越来越科学，社会生活日趋复杂。为了解国内外社会经济状况、市场状况和企业的生产情况，资产阶级必须对有关的经济活动，进行广泛的统计，于是工业、农业、商业、对外贸易等各方面的统计活动都发展起来。当资本主义发展到垄断资本主义阶段，随着科学技术的进步，统计在现代经济管理和科学技术分析中的作用更为重要，从而得到了更广泛的应用和发展。随着统计实践的发展和经验的积累，各种统计理论和学说应运而生。但是直到17世纪以后统计才逐渐发展成一门科学，逐步形成了不同的学派。从统计学产生和发展的过程来看，大致可以分为古典统计学、近代统计学和现代统计学三个时期。

### （一）古典统计学时期

这是指17世纪中叶至18世纪中叶统计学萌芽时期。当时有国势学派和政治算术学派。

#### 1. 国势学派（记述学派）

这个学派产生于18世纪封建制度的德国，代表人物是康令（Hesmanu Corning, 1606—1681），继承者主要有阿肯瓦尔（Cottfriol Achen Wall, 1719—1772）。《近代欧洲各国国势

学概论》是为代表作。他们在哥丁根大学开设了一门新的课程,最初叫做国势学,因为在德文中“国势”与统计的含义相通,后正式定为统计学。它收集大量的实际资料分门别类记述有关国情国力,包括土地、人口、政治、军事、财政。这个学派主要采用文字记述为主,始终没有把数量对比分析作为这门科学的基本特征,故又称为记述学派。这一学派对统计学的贡献不是很大,只是阿肯瓦尔在 1749 年第一个把国势学定名为统计学,统计学这一名词以后就一直沿用至今。

## 2. 政治算术学派

这一学派产生于 17 世纪资本主义的英国,创始人是威廉·配第(William Petty, 1623—1687)。他在研究各国政治经济情况时不只是用文字表述,还通过大量的数字资料进行对比分析,来反映国情国力。配第在《政治算术》一书中第一次用计量和比较的方法,即大量观察法、分类法以及对比综合法等,从整体上分析了英、法、荷三国的经济军事实力及其内在潜力,论证了英国称雄世界的条件和地位。《政治算术》中提出了一整套比较系统地反映社会经济现象并分析其数量关系的方法,创造性地进行了建立一门统计科学的尝试。马克思称配第在某种程度上也可以说是统计学的创始人。这里的政治是指政治经济学,算术是指统计方法。这本书中采用了不同于过去传统的方法,而是利用数字、重量、尺寸来说话,找出事物变化的规律。

政治算术学派的另一个代表人物是约翰·格朗特(J. Graunt, 1620—1674)。17 世纪上半叶,英国多次发生严重的瘟疫,政府定期公布有关人口出生和死亡的数字。约翰·格朗特利用这些数字资料研究并发表了《关于死亡表的自然和政治的观察》的论著,首次提出通过大量观察,可以发现新生儿性别比例具有稳定性和不同死因的比例等人口规律;并且第一次编制了初具规模的生命表,对年龄性别死亡率与人口寿命做了分析,从而引起了普遍的关注。

## (二) 近代统计学时期

18 世纪末~19 世纪末的 100 多年间,统计学有了长足的发展,其间形成了两大主流学派——数理统计学派和社会统计学派。

### 1. 数理统计学派

数理统计学派产生于 19 世纪中叶,其奠基人是比利时的物理学家和统计学者凯特勒(A. Quetelet, 1796—1874)。他是法国数学家、物理学家拉普拉斯的学生,其代表作是《论人类》、《概率论书简》和《社会物理学》。凯特勒在研究社会现象时主张采用自然科学的方法,把古典概率论引进统计学,发展了概率论,推广了概率论在统计中的应用。他最先用大数定律论证了社会经济现象纷繁复杂的偶然性中隐藏着规律性,并且提出了误差理论,用以解决统计上的准确性问题。他明确地指出统计学是一门既研究社会现象又研究自然现象的独立的方法论科学。随着统计研究的领域不断扩大,概率论方法的运用日趋成熟。到 19 世纪末数理统计学慢慢地独立,自成一派。

凯特勒把德国的国势学派、英国的政治算术学派和意大利、法国的古典概率论加以融合改造为近代意义的统计学。他是数理统计学派的奠基人。

### 2. 社会统计学派

19 世纪后半叶,正当数理统计学派开始发展的时候,在欧洲以德国为中心又兴起了社会统计学派。由大学教授克尼斯(K. G. A. Knies, 1821—1898)首创,主要代表人物有恩格尔

(C. L. E. Engel, 1821—1896)、梅尔 (G. V. Mayr, 1841—1925) 等人。他们认为统计学是一门社会科学, 是研究社会现象变动原因和规律性的实质性科学, 以此同数理统计学派的通用方法论相抗衡。这是社会统计学派的“实质性科学”的显著特点。他们强调统计学是研究社会现象的科学, 包括统计资料的搜集、整理和分析研究, 目的是要揭示现象内部的联系。强调在统计研究中必须以事物的质为前提和认识质的重要性。从学术渊源上看, 他们融合了记述学派和政治算术学派的观点, 又继承和发扬了凯特勒强调研究社会现象的传统, 把政府统计和社会调查结合起来形成自己的特点。德国的社会统计学派在国际学术界占有一席之地, 日本等国深受其影响。

### (三) 现代统计学时期

这是指 20 世纪初至今的时期。这一时期统计科学出现了新的趋势。

由于自然科学和工程技术科学的发展日新月异, 数理统计学与其紧密结合, 获得了广泛的应用并进入了鼎盛时期。首先它在随机抽样的基础上建立了推断统计的理论和方法。统计学的主流从描述统计学转向推断统计学。

20 世纪 30 年代 R·费希尔的推断统计理论标志着现代数理统计学的确立。所谓推断统计也就是通过随机样本来推断总体数量特征的方法。这种方法源于英国数学家格赛特 (N. S. Gosset, 1876—1936) 的小样本  $t$ -分布理论, 其后美国学者费希尔 (R. A. Fisher)、波兰统计学家内曼 (J. Neyman, 1894—?) 以及 E. S 比尔森等人进一步发展, 建立了统计假设理论。后来美国统计学家瓦尔德 (A. Wald, 1902—1952) 又将统计学中的估计和假设理论加以归纳, 创立了决策理论; 美国的威尔克斯 (S. S. Wilks, 1906—1964)、英国的威史特 (J. Wishart, 1898—1956) 等学者对样本分布理论又加以充实和发展; 美国的科克伦 (W. G. Cochran, 1909—1980) 等在 1957 年提出了实验设计的理论和方法, 大大拓宽了统计学的范围。

20 世纪 60 年代以后数理统计学发展有三个明显的趋势: (1) 统计学对数学的依赖和吸收数学更多, 大量的采用数学方法; (2) 数理统计学的新分支和以数理统计为基础的边缘学科不断形成; (3) 与电子计算机技术相结合, 应用范围更广, 作用更大, 已成为现代统计学的主流。

而同一时期的社会统计学发展则比较缓慢, 其基本趋势是由实质性科学向方法论转变。重要的人物如二战后德国法兰克福大学教授弗拉斯克姆波 (P. Flaskamper, 1886—?)。

在其他国家, 如前苏联的一些统计学者也在理论上做出了一定的贡献, 但囿于意识形态和其他方面因素的束缚, 没能及时吸取各国统计学发展的最新成果, 反而一味地加以排斥以至落后于西方学者。中国的统计学发展也深受影响进步迟缓, 经济统计和数理统计之争一直没有间断。进入到 20 世纪 80 年代以后, 局面有所改观, 大统计格局 (不分经济统计和数理统计) 深入人心。

## 第二节 统计学的特点与方法

从统计学的发展史来看, 统计学是从研究社会经济现象开始的, 经过 300 多年的演变和发展, 统计学趋于成熟, 成为一门研究客观事物数量方面的独立的方法论科学。从研究对

象、研究方法和学科体系上看,统计学有以下三个特点:

(1) 它是研究客观事物总体数量方面的方法论科学。

任何事物都是质和量的统一,事物的变化也是由量的渐变达到一定程度后才引起质的根本变化的。科学本身的发展程度也决定于它对客观世界事物数量方面的认识和数学工具的掌握。由此可见,统计学是从数量方面来认识客观世界,是人类认识深化的客观要求,也是科学发展的必然产物。

统计研究客观事物的数量方面,指的是总体的数量方面,而不是个别事物的数量方面。最初的统计学是同人口问题的研究结合在一起的。人口就是一个群体。一定范围的人口出生率、死亡率、平均寿命等都是人口总体的数量特征。虽然统计必须对人口总体中每一个人的数量特征和属性进行登记,但这不是统计研究的目的所在,而是统计研究的必要阶段。为的是从个体特征向总体数量方面过渡。

同时,统计研究客观事物的数量方面还必须是有对象的具体的数量,而不是抽象的量。而数学则是研究抽象的数量关系和空间几何形式,是舍弃了具体对象质的规定性的抽象的量。因此,它不完全同于数学,尽管统计学中有许多数学公式和应用各种数学方法。它必须在质与量的紧密结合中研究总体现象的量的方面。

(2) 统计研究强调对客观事物总体进行大量的观察,通过归纳推理来获得总体数量方面的综合认识。

所谓大量观察,就是对所研究的总体的全部或足够数量的单位进行观察。只有对总体的全部或足够数量的单位进行观察,并加以综合分析,才能使事物中非本质的偶然因素的影响相互抵消而获得关于总体本质特征或规律性的认识。300多年前从统计资料中发现新生婴儿的男女比例为105:100,就是大量观察的结果。又如,商品价格有高有低,经过大量观察发现总价格和总价值、平均价格和价值基本上趋于一致,在价格的偶然性波动中存在着价值规律的必然性的支配作用。

这种大量观察的方法虽然是由社会统计学派创立,但也是数理统计的基本方法原则。凯特勒曾经明确指出过:统计学是对性质相同的事物进行大量观察,从而探索出社会现象相续不绝之理的一门学问。凯特勒把概率论引进统计学,进一步从随机现象的角度论证了大量观察方法的客观依据,即大数定律。它表明:在观察过程中因为有偶然性,每次取得的结果不同;但大量重复观察所得结果的平均值却几乎接近确定的数值。由于大数定律具有普遍性,所以统计学家把大数定律称为统计规律。

大量观察,既可以是对总体的全部单位观察,也可以是对其中足够数量的单位进行观察。也就是说,可以对部分单位观察,经过归纳推理而获得对总体数量特征的综合认识。这在统计方法论上具有特殊的意义。因为需要认识的对象范围往往相当大,有时甚至是无限的以至于只能从中观察部分单位进行计算和分析,并据以推断总体。例如:通过部分产品进行破坏性或耐用性试验来推断整批产品的平均使用寿命;通过对部分居民的收支的调查来推断全部居民的平均生活水平等等。这种根据样本数据来推断总体数量特征的归纳推理称为统计推断,是现代统计学方法体系中一个重要组成部分。

(3) 统计学是一门多科性的科学。

统计学经过300多年的发展,与各学科之间既有融合又有分化,特别是在现代自然科学和社会科学相互影响和渗透、边缘学科不断涌现的大趋势下,统计学作为一门方法论科学应

用于各学科,同它们结合,产生了一系列专门的或专业的统计学分支,从而使统计学形成了一个多学科的家族。对其主体部分,以研究领域为标志进行划分,统计学体系可分为两个部分:数理统计学和社会统计学。数理统计学,它是以概率论为基础,以抽样为核心,研究随机变量的数量规律,是在社会经济和自然科学及应用技术统计实践基础上抽象和概括出来的具有普遍性的纯方法论科学。数理统计学又分为理论数理统计学和应用数理统计学,其分支如抽样理论、非参数统计、多变量分析和时间序列分析等。社会统计学是以人口、社会、经济、科技、文化等人类自身及其活动为研究对象的统计方法论,包括众多学科分支,如人口统计学、管理统计学、商业统计学、金融统计学、教育统计学等。还有一些交叉学科如生物统计学、心理统计学、气象统计学、医疗卫生统计学、统计物理学,这些学科都是统计学与自然科学和应用技术科学相结合产生的新学科。此外还有统计史、比较统计学等。

### 第三节 统计学基本概念

#### 一、统计总体和总体单位

“统计总体”就是统计研究的对象的全体,简称“总体”。它是由统计研究目的所决定的、具有某种共同性质的多个个体构成的集合体。统计总体随研究的目的要求及范围的变化而改变。例如,目的是研究某市今年国民生产总值状况,全市所有区县的集合就是统计总体。因为各区县在一定时空条件下是客观存在的,尽管它们之间存在着经济规模、结构、水平和效益等方面的差异,但每个区县都有国民生产总值,在这方面性质是相同的。

“总体单位”就是指构成总体的各个基本单位,简称为“单位”。根据研究问题的要求不同,总体单位可以是一个人、一事物、一个班组或一个企业。如果要研究沈阳市工业生产情况,全市所有工业企业的集合体就是统计总体,而每个工业企业就是总体的单位。又如,要进行工业生产设备普查,所有工业企业生产设备构成总体,每台生产设备则是总体单位。

统计总体具有以下几个基本特征:

(1) 大量性。统计总体是由各个总体单位组成的。由于各个总体单位的特征互有差别,特别是在各单位之间差异很大时,要求组成总体的总体单位更应该多一些,在大量事物的普遍联系中体现出总体特征。如某市所有职工组成的一个总体,每个职工是一个总体单位,要了解职工工资一般水平,只调查少数几个职工的工资情况是不行的。因为所调查的那几个职工的工资可能偏高或偏低,不能准确代表全体职工工资的一般水平。所以调查的职工越多越能准确反映总体特征。

(2) 同质性。同质性是指各单位在某一方面共同具有的性质。它是形成统计总体的一个必要条件。如果我们要研究各工业企业的生产情况,就把全国工业企业作为一个总体。每一个工业企业都是从事工业生产活动、制造工业产品的基层单位,在这方面每个工业企业具有相同的性质。通过对总体进行研究,才能说明我国工业生产活动的状况特征以及存在的问题。相反构成总体各单位如不具备相同性质的基础,如在研究上述工业生产情况时,把商店、农场也包括进来,其研究的结果就不能说明全国工业生产活动的情况。

(3) 变异性。构成统计总体的各个单位在某一方面是同质的,但是在其他方面又是有所差别的,在统计上把这种差别称为变异。如在工人总体中每个工人是同质的,但是他们的工

资、性别、工龄、技术等级又都有差别。统计总体中各单位之间存在着差别是统计研究的前提条件,如果没有差别统计研究也就没有意义,只有对这些差异进行研究才能揭示总体的特征。

在统计研究中总体和总体单位不是一成不变,它是根据研究目的变化而变化。它们的关系是总体与个体的关系,是通过总体单位具体数量的计算来反映总体特征,实现统计研究的目的。统计总体与总体单位是相对而言的,并非固定不变,它们随着研究的目的不同而相互转化。统计总体可以转化为总体单位,总体单位也可以转化为统计总体。既然统计总体与研究的目的直接相关,那么,当目的发生了改变,统计总体也会改变,构成总体的总体单位也将随之而改变。假如目的不是全省而是研究全国工业企业生产经营情况,则全国所有工业企业为统计总体,各省、市、自治区的工业企业就由原来的统计总体转化成了总体单位;如果目的是研究某地区工业企业生产经营情况,则地区的工业企业由原来的总体单位转化成了统计总体,而构成该地区的各县工业企业又上升为总体单位。

根据统计总体中包含单位数的多少,可以分为有限总体和无限总体。有限总体是指统计总体中包含的总体单位数是有限的。如企业个数、职工人数等。无限总体是指统计总体中包含的总体单位数是无限的。如工厂批量流水线生产的产品、织物纤维、病虫细菌繁殖等。对有限总体可以对其全部单位计数,故可以进行全面调查,也可以进行非全面调查;对无限总体的单位无法对其全部单位计数,故只能采取抽样调查,并据此推算总体状况。

从总体中抽取出来代表这一总体的部分单位组成的集合称为“样本”。虽然样本中的单位数相对于总体而言只是很少一部分,但样本是从总体中产生并代表它的。基于这种关系,总体也可以称为“母体”,而样本则又称为“子样”。

## 二、标志

标志也称为“标识”。标志是说明总体单位所具有特征的名称。按表示特征不同分为品质标志和数量标志。

### 1. 品质标志

品质标志是表明总体单位品质和属性特征的标志。它一般用文字来表示。如工人的性别、工种,企业的经济类型、规模大小等。这类标志只是从性质的差别上来表现总体特征和属性的。如果说某公司下属各企业为总体单位,则各企业生产经营方向、所有制性质是企业质的属性,这些品质标志一般用文字说明,不用数值表示。实践中也有以量来反映总体单位质的属性的,如反映产品质量的“一等品”,劳动效率为“人均年销售额 200 万元”等。

### 2. 数量标志

数量标志是表明总体单位数量特征的标志,即变量。它是用数值表示的。如工人的工龄、年龄、工资等,又如作为总体单位的各工业企业的资本总额、资本综合成本、工业总产值、净产值、商品产值、成本总额、期间费用、实现利税总额等,这些都是以具体的绝对数值来表示的。又如企业的工人数、总产值、设备原值等,这类标志只是从数量上的多少来表现总体的特征,所以称为数量标志,数量标志的具体表现值就称为“标志值”。

## 三、统计指标与指标体系

统计指标也称为“指标”,反映了总体数量特征的概念和具体数值。它是由总体中各单