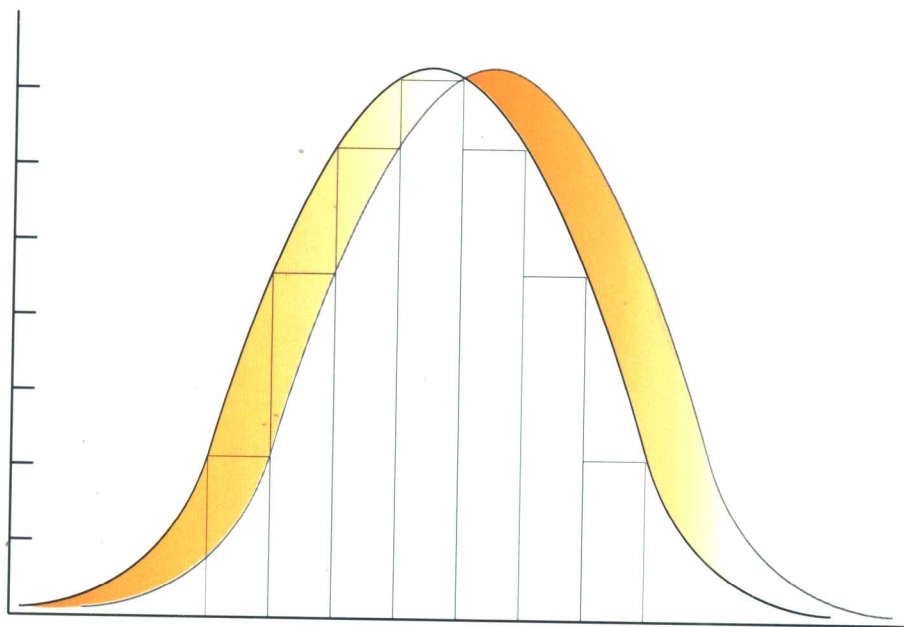


21世纪高等院校规划教材

# 统计学原理

王长江 郝华荣 主编



国防工业出版社

National Defense Industry Press

21 世纪高等院校规划教材

# 统计学原理

王长江 郝华荣 主编

国防工业出版社

·北京·

**图书在版编目(CIP)数据**

统计学原理/王长江,郝华荣主编. —北京:国防工业出版社,2005.8

21 世纪高等院校规划教材

ISBN 7-118-03989-6

I. 统... II. ①王...②郝... III. 统计学  
IV. C8

中国版本图书馆 CIP 数据核字(2005)第 066211 号

**国防工业出版社** 出版发行

(北京市海淀区紫竹院南路 23 号)

(邮政编码 100044)

新艺印刷厂印刷

新华书店经售

\*

开本 787×1092 1/16 印张 25¼ 580 千字

2005 年 8 月第 1 版 2005 年 8 月北京第 1 次印刷

印数:1—4000 册 定价:33.00 元

---

(本书如有印装错误,我社负责调换)

国防书店:(010)68428422

发行邮购:(010)68414474

发行传真:(010)68411535

发行业务:(010)68472764

## 内 容 简 介

本书在详细介绍统计学基本理论的同时,讨论了日常统计工作中常用的统计方法。全书共分 13 章,包括描述统计、推断统计和常用统计方法三部分内容。前 4 章为描述统计部分,内容包括统计数据的收集方法、数据的加工整理方法、数据的显示方法、数据分布特征的概括与分析方法等;第五章到第十章为推断统计部分,内容包括概率基础理论、抽样与抽样分布理论、参数估计、假设检验、方差分析、列联表分析等内容;第十一章到第十三章为常用统计方法部分,详细介绍了相关与回归分析方法、时间数列分析方法、指数分析方法。

本书文字简洁,层次清晰,基本理论与实例分析相结合,具有很强的指导性和实用性。本书既可作为高等院校相关专业师生教学用书,也可作为实际统计工作者的参考书。

# 前 言

统计学是教育部规定的高等院校财经类专业的一门专业基础课,本教材就是为了满足上述专业教学的需要而编写的。

本书包括描述统计、推断统计和经济管理中常用统计方法三部分内容。第一章至第四章为描述统计部分,第五章至第十章为推断统计部分,第十一章至第十三章为常用统计方法部分。本书在认真总结多年教学经验的基础上,努力将数理统计学和社会经济统计学有机融合于编写体系中,并力求在体系和内容上做到重点突出、通俗易懂,注重对统计方法的阐述,注意理论联系实际,多举实例。为便于读者复习和加深理解,本书各章后均配有适量的习题。

计算机在现代统计分析中的作用越来越突出,目前产生了大量的专门分析软件,本书限于篇幅,对此问题并未涉及,但建议读者能自行学习这些内容。

本书由王长江、郝华荣任主编,郝晓东任副主编。具体编写工作如下:第一、二、四、五、八章由郝华荣编写;第三章及附录由王继仓编写;第六、七、九章及小部分附录由郝晓东编写;第十、十一、十二、十三章由王长江编写。全书由王长江进行总纂定稿,在编写过程中得到了课程组各位老师的热心帮助。

对给予帮助的各位老师表示深刻感谢。

编著者

2005年4月

# 目 录

|                            |     |
|----------------------------|-----|
| 第一章 绪论 .....               | 1   |
| 第一节 统计与统计学 .....           | 1   |
| 第二节 统计学的产生和发展 .....        | 4   |
| 第三节 统计学学科及与其他学科关系 .....    | 7   |
| 第四节 统计学中的几个基本概念 .....      | 9   |
| 练习 .....                   | 13  |
| 第二章 统计数据的收集 .....          | 14  |
| 第一节 统计测量尺度 .....           | 14  |
| 第二节 原始数据的收集 .....          | 18  |
| 第三节 次级资料的收集 .....          | 28  |
| 第四节 统计调查误差 .....           | 30  |
| 练习 .....                   | 31  |
| 第三章 统计数据的整理 .....          | 33  |
| 第一节 数据整理的程序 .....          | 33  |
| 第二节 数据分组 .....             | 35  |
| 第三节 统计表与统计图 .....          | 45  |
| 练习 .....                   | 54  |
| 第四章 综合指标 .....             | 57  |
| 第一节 总量指标 .....             | 57  |
| 第二节 相对指标 .....             | 60  |
| 第三节 平均指标 .....             | 66  |
| 第四节 变异指标 .....             | 85  |
| 第五节 标准差应用 .....            | 92  |
| 练习 .....                   | 97  |
| 第五章 概率及其分布 .....           | 103 |
| 第一节 概率 .....               | 103 |
| 第二节 概率分布 .....             | 112 |
| 第三节 常用的离散型随机变量及其概率分布 ..... | 117 |
| 第四节 常用的连续型分布——正态分布 .....   | 121 |

|                                                 |            |
|-------------------------------------------------|------------|
| 练习.....                                         | 128        |
| <b>第六章 抽样与抽样分布 .....</b>                        | <b>130</b> |
| 第一节 有关抽样的基本概念、意义及其作用.....                       | 130        |
| 第二节 基本抽样设计.....                                 | 136        |
| 第三节 样本统计量和抽样分布.....                             | 140        |
| 第四节 样本平均数的抽样分布.....                             | 141        |
| 第五节 中心极限定理.....                                 | 146        |
| 第六节 有限总体校正系数.....                               | 149        |
| 第七节 比率的抽样分布.....                                | 150        |
| 第八节 总体标准差 $\sigma$ 不明确时的抽样分析 .....              | 152        |
| 练习.....                                         | 153        |
| <b>第七章 参数估计 .....</b>                           | <b>156</b> |
| 第一节 基本概念 几个重要的点估计与点估计的优良性.....                  | 156        |
| 第二节 母体均值 $\mu$ 的区间估计 .....                      | 164        |
| 第三节 比例 $p$ 的区间估计 .....                          | 171        |
| 第四节 正态母体 $\sigma^2$ 的区间估计 .....                 | 173        |
| 第五节 两母体均值差的区间估计.....                            | 175        |
| 第六节 两比例 $p_1$ 、 $p_2$ 之差 $p_1 - p_2$ 的区间估计..... | 180        |
| 第七节 样本容量的决定.....                                | 182        |
| 练习.....                                         | 185        |
| <b>第八章 假设检验 .....</b>                           | <b>188</b> |
| 第一节 假设检验的基本原理.....                              | 188        |
| 第二节 总体均值的假设检验.....                              | 194        |
| 第三节 总体比例的假设检验.....                              | 201        |
| 练习.....                                         | 203        |
| <b>第九章 方差分析 .....</b>                           | <b>206</b> |
| 第一节 方差分析的基本思想.....                              | 206        |
| 第二节 单因素方差分析.....                                | 207        |
| 第三节 两因素方差分析.....                                | 217        |
| 练习.....                                         | 230        |
| <b>第十章 列联分析 .....</b>                           | <b>232</b> |
| 第一节 列联表.....                                    | 232        |
| 第二节 $\chi^2$ 分布和 $\chi^2$ 检验 .....              | 235        |

|             |                                 |            |
|-------------|---------------------------------|------------|
| 第三节         | 列联表中的相关测量.....                  | 240        |
| 第四节         | 列联分析中应注意的问题.....                | 244        |
| 练习          | .....                           | 247        |
| <b>第十一章</b> | <b>相关关系分析</b> .....             | <b>249</b> |
| 第一节         | 相关关系分析的概念和任务.....               | 249        |
| 第二节         | 直线相关的测定.....                    | 251        |
| 第三节         | 回归分析——简单直线回归.....               | 260        |
| 第四节         | 回归分析——曲线回归及复回归.....             | 270        |
| 练习          | .....                           | 285        |
| <b>第十二章</b> | <b>时间数列</b> .....               | <b>291</b> |
| 第一节         | 时间数列的种类和编制方法.....               | 291        |
| 第二节         | 动态分析指标.....                     | 296        |
| 第三节         | 时间数列的测定——长期趋势.....              | 303        |
| 第四节         | 时间数列的测定——季节变动、循环变动和剩余变动的测定..... | 312        |
| 练习          | .....                           | 317        |
| <b>第十三章</b> | <b>指数</b> .....                 | <b>320</b> |
| 第一节         | 指数的概念和作用.....                   | 320        |
| 第二节         | 综合指数.....                       | 323        |
| 第三节         | 平均数指数.....                      | 330        |
| 第四节         | 平均指标指数.....                     | 334        |
| 第五节         | 指数体系与因素分析法.....                 | 338        |
| 第六节         | 指数数列.....                       | 344        |
| 第七节         | 指数的应用.....                      | 346        |
| 练习          | .....                           | 351        |
| <b>附录</b>   | <b>常用统计表</b> .....              | <b>357</b> |
| 附表 1        | 二项分布表.....                      | 357        |
| 附表 2        | 泊松分布表.....                      | 369        |
| 附表 3        | 标准正态分布表.....                    | 371        |
| 附表 4        | 正态分布分位数表.....                   | 372        |
| 附表 5        | $t$ 分布表.....                    | 375        |
| 附表 6        | $\chi^2$ 分布表 .....              | 377        |
| 附表 7        | $F$ 分布表 .....                   | 380        |
| 附表 8        | 随机数表.....                       | 393        |
| 附表 9        | 相关系数 $\rho=0$ 的临界值表 .....       | 396        |
| <b>参考文献</b> | .....                           | <b>397</b> |

# 第一章 绪论

## 第一节 统计与统计学

### 一、统计的涵义

对于“统计”一词，人们并不陌生，在日常工作及生活中，人们经常用到、看到和听到这个词。例如：开会时要统计出席会议的人数；球赛中教练员要统计各位队员的命中率、犯规次数；在有关媒体中也经常会看见一些报道使用统计数据、图表等。因而人们一般提到统计首先想到的是统计工作，想到政府统计机构以及从事统计工作的人等。事实上这种理解是不全面的。统计作为一种社会实践活动，自人类文明开始就有了，而今，“统计”一词已被人们赋予多种含义。通常，统计一词包含以下三种含义：

一是统计工作 (statistical work)，它是指为了取得和提供统计数据资料而进行的各项工作，包括对研究目标进行数据收集、整理、分析和提供的全部工作。我们的各级政府机构基本上都有统计部门，如统计局，它们的职能主要就是从事统计数据的收集、整理、提供。大多数企业也都有专门从事统计工作的人员，负责企业生产和销售数据的记录、积累以及向上级部门报送数据、分析数据的任务。如果对一个从事统计工作的人来说，他们提到的“统计”一词也许单指统计工作。

二是统计数据 (statistical data)，它是指统计工作的成果，包括调查得到的经过整理具有信息价值的各种统计数据和分析报告。我们经常看到专门出版统计数据的出版物，如各类统计年鉴，在报纸、杂志、网络及其他媒体上都会见到大量的统计数据。这些统计数据就是统计工作成果的体现。当看到或听到“据统计……”这样的说法时，这里的“统计”一词特指统计数据。

三是统计学 (statistics)，它是指分析统计数据的方法和技术，是一门对群体现象数据特征进行计量描述和分析推论的科学。从事统计研究或统计教学的人，他们心目中的“统计”一词多指统计科学。人们在日常的表达中，习惯地将统计学简称为统计。

可见，在不同场合，“统计”一词可以具有不同的含义。但统计的三种含义是密切联系的。第一，统计工作和统计资料是实践和成果的关系，一方面统计工作的实施受统计资料需求的支配，另一方面统计工作的质量又直接决定着统计资料的数量和质量。第二，统计工作与统计学是实践和理论的关系，一方面，统计理论是统计实践活动的经验总结和理论概括，另一方面，统计工作的发展又需要统计理论的指导，统计工作现代化是和统计科学的发展相关联的。

## 二、统计学的研究对象

统计学的研究对象是指统计研究所要认识的客体,只有明确了研究对象,才可能根据它的性质特点指出相应的研究方法,达到认识对象客体规律性的目的。

我国成立五十多年来,关于统计学的研究对象、统计学是一门还是两门等统计学中的一些重大理论问题,在统计学界曾展开几次比较热烈的争论。几十年的争论促进了我国统计学的发展。自1992年1月我国国家技术监督局颁布GB/T14745-92《学科分类与代码》中,将统计学与数学、经济学等学科并列上升为一级学科。自从把包括原属社会科学领域和自然科学领域的各种统计学归并为一门统计学以来,在我国的统计学界基本达成了共识,接受了大统计思想,即统计学不仅包括社会经济统计学,也包括数理统计学及自然领域中的各种统计学。统计学是个多学科的大家族,它的内容遍及自然现象及社会现象的各个领域,因而有着不同门类的统计学,在各门统计学之间存在着共性,但由于研究对象的内容和性质不同,又产生各自不同的理论和方法。因而在各门统计学之间也存在显著的差异,形成不同领域的各门独立的统计学,如天文统计学、物理统计学、经济统计学等。本书的统计学并不是作为一级学科的大统计学,它仅仅是大统计学中的一门课程,其内容主要阐述大统计学的基本的理论知识 and 最基本的统计方法,即统计学原理。

统计学的研究对象是在质与量的相互联系中研究一切现象总体的数量特征及其相互关系。

对统计学的研究对象进行认识,须明确以下几点:

第一,统计学的研究对象是群体现象。

统计学中有两个基本概念:总体和样本。总体是指调研者研究对象的集合;样本是指来自总体的部分对象的集合。统计学要研究的是这些集合,而不是组成集合的个体。如果要知道一件物体有多重,只要把它称一称就可;如果要知道一个人的身高,只要测量一下就可知。而若要知道一批物体的重量,一群人的身高,就需要汇总平均,采用一定的统计方法。若期望通过小小的样本信息了解总体,就需要抽样调查,估计检验,进行统计分析推论。虽然总体或样本的信息都表现在一个个元素或个体上,研究总体不能脱离个体,但统计学研究的不是个体现象,而是通过个体所载有的信息来研究、说明群体现象。

第二,统计学的研究对象是群体现象的数量表现。

统计学的英文是“statistics”。这个英文名词有两个含义:作为复数,它表示统计数据;作为单数,则表示作为一门科学的“统计学”。从中可以看出,统计学和统计数据之间有着密不可分的关系。数据是统计的语言,统计学是用数据来说明总体现象特征的。作为特征,可以是数量特征,如人的身高、年龄等;也可以是属性特征,如人的性别、民族等。统计学研究群体现象的特征,总是用数字来计量、说明的。例如,统计学所探索的某人口群体的性别特征表现为不同性别的人口数量、比例;民族特征为各民族人口数量、所占比例等。离开了统计数据,统计就失去了用武之地,统计学也就失去了它存在的意义。

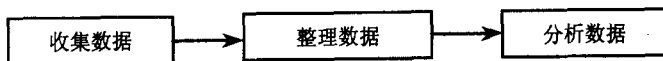
第三, 统计学所探索的是群体现象数量表现的内在规律性。

现实中群体现象总是由许多数量特征各异的个体组成, 而这些千差万别的个体数量特征下却掩盖着群体现象的某一数量规律性。例如, 就单独一个家庭来观察, 每个家庭的新生婴儿的性别可能是男性, 也可能是女性。从表面上看, 新生婴儿的性别比例似乎是没有规律可循的。但如果就大量的家庭新生儿进行观察, 就会发现新生儿中男女比例约为 107:100。这个比例就是新生儿这个群体性别特征的数量规律。这是由人类自然发展的内在规律所决定的。又如, 投掷硬币或掷骰子的游戏, 随机地投掷一次硬币或骰子, 出现正面、反面或某个点数是不确定的, 完全是偶然的。但当我们进行多次的重复投掷, 就会发现投掷一枚均匀硬币出现正反面的次数大体相同, 即比值接近于  $1/2$ 。投掷的次数越多, 就越接近于  $1/2$  这一稳定的数值。同样, 在掷骰子时, 出现 1 点~6 点的比例也逐渐接近于  $1/6$ 。这里的  $1/2$  和  $1/6$  就是掷硬币和掷骰子出现某一特定结果的概率, 也就是投掷硬币或骰子时所呈现的数量规律性。再如, 在进行农作物试验时, 如果其他试验条件相同, 我们会发现某种粮食作物的产量会随着某种肥料施肥量的增加而变化。当最初增加施肥量时, 产量增加较快, 以后增加同样的施肥量, 粮食产量的增加逐渐减少, 当施肥量增加到一定数值时, 产量不再增加, 这时如果再增加施肥量, 产量反而会减少。粮食产量与施肥量之间的这种数量关系, 就是我们所要探索的数量规律性, 从而有助于确定最佳的施肥量, 以求得最大的效益。

以上说明, 通过多次观察或试验得到大量的统计数据, 利用统计方法可以探索出其内在数量的规律性。统计学正是研究各种各样的统计方法, 这些方法是分析认识客观现象的有力工具, 帮助人们在各种不确定的或复杂的情况下作出明智的判断和决策。

### 三、统计研究工作过程

统计学是一套处理数据的方法和技术, 它告诉我们, 统计离不开数据。统计研究的过程首先要有数据, 在拿到数据后, 为满足分析的需要, 还要对数据进行一定的整理, 而后再对数据进行分析。因此统计研究的过程可作如下描述。



数据收集是取得统计数据的过程, 它是进行统计分析的基础。如何取得较为可靠的统计数据是统计学研究的内容之一。数据收集方法可归为两大类: 一是观察法, 通过调查或观测而收集数据; 二是实验方法, 在实验中控制试验对象而收集数据。

数据整理是统计研究的中间阶段, 它是对统计数据进行科学的加工处理, 使统计数据系统化、条理化, 以便进一步开展统计分析。数据整理的内容包括数据的审核, 以保证数据的质量, 为进一步的整理与分析打下基础; 数据分组, 将数据按需要进行分门别类; 数据的表述, 将数据用图表等形式展示出来, 以便找出数据的初步特征, 或者是方便别人看懂数据所要表达的问题。

数据分析是统计学的核心内容, 它是通过统计方法对加工整理后的统计数据进行

多方面的分析研究,发现问题,揭示矛盾,寻找被研究现象的特征和规律,形成统计研究的最终成果。这一阶段是统计研究的决定性阶段。

## 第二节 统计学的产生和发展

从人类文明开始起就有了统计活动,即统计工作有着数千年的历史。而统计学成为系统和独立的科学至今只有 300 多年的历史,它是数千年的统计活动中,人们对统计规律的认识逐渐加深、并不断总结规范出来的。一般认为,统计学产生于 17 世纪中叶,其形成过程是从几个不同的领域开始的,因而在统计学的产生和发展过程中,产生了许多不同的统计学派。

### 一、古典统计学时期

17 世纪中叶欧洲各国相继进入资本主义工场手工业的经济迅速发展阶段,但是某些国家的封建制度尚未解体,这时的欧洲处于思想活跃的社会变革时期。为了适应各国经济发展的不同需要,欧洲各国不约而同地从不同领域开始了统计学的奠基工作,并相继形成了统计学的政治算术学派和国势学派。

#### (一) 政治算术学派

政治算术学派产生于 17 世纪中叶的英国,创始人是威廉·配第(William Petty, 1623—1687)。威廉·配第在其代表作《政治算术》一书中,运用大量的数字资料,用计算和对比的方法对英、法、荷三国的经济实力进行了比较,论证了英国称雄世界的条件和地位。他在书中运用数字、重量、尺度等进行数量对比分析,其所采用的方法是前所未有的,为统计学的产生奠定了基础。马克思称他为“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”。

政治算术学派的另一个代表人物是约翰·格朗特(John 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·格朗特利用这些资料研究并发表了《关于死亡表的自然观察和政治观察》(1662 年)的论著,首次通过大量观察的方法,研究并发现了一系列人口统计规律,如男婴出生多于女婴,基本上为 14:13;男性的死亡率高于女性;新生儿在大城市的死亡率较高;一般疾病和事故的死亡率较稳定,而传染病的死亡率波动较大;等等。此外,约翰·格朗特还在研究中运用各种方法对统计资料进行间接的推算,相互印证。由于约翰·格朗特的这些研究成果,该书被许多统计学家誉为“真正统计科学的肇端”。

政治算术学派在当时的欧洲广泛传播,并逐渐形成了两大分支,即以信奉配第为主的经济统计派,以信奉格朗特为主的人口统计派。政治算术学派为后来的社会经济统计的发展奠定了基础。

#### (二) 国势学派

国势学派又称记述学派,产生于 18 世纪的德国。所谓国势学就是以文字记述国家

显著事项的学说。由于当时德国许多大学设有国势学这门课程,故国势学派亦称德意志大学教授派。其创始人是海尔曼·康令(Hermann Coning, 1606—1681),他第一个在德国赫尔莫斯达德大学讲授《欧洲最近国势学》,奠定了国势学的基础。

国势学派最重要的继承人是高特弗瑞德·阿痕瓦尔(Gottfried Achenwall, 1719—1772),其主要著作是《欧洲各国国势学概论》,书中讲述“一国或多数国家的显著事项”,主要是用对比分析的方法研究关于国家组织、人口、军队、领土、财产等国情国力,比较各国实力的强弱,为德国的君主政体服务。他在1749年确定了统计学(statistics)这一学科的名称及有关统计学的一些术语。

国势学派只是对国情的记述,未能进一步揭示社会经济现象的规律,也不研究事物的计量分析方法,只是用比较级和最高级的词汇对事物的状态进行描述。所以,人们也把它叫做记述学派,并认为国势学派有统计学之名,而无统计学之实。但统计学之名就此沿用下来。至于确认以现象的数量方面为研究内容才是统计,是在19世纪中叶,即当代著名的德国经济学家和统计学家克尼斯(K. G. A Knies, 1821—1897)于1850年发表了《独立科学的统计学》的论文,它综合了当时各国大多数经济学家和统计学家的意见,提出了将国家论作为国势学的科学命名,而将统计学作为政治算术的科学命名,从此统计学才名副其实,也标志着两大统计学派争论的结束。

## 二、近代统计学时期

18世纪末到19世纪末的一百多年中,统计学有了很大发展,其间的主要贡献是建立和完善了统计学的理论体系,并逐渐形成了以随机现象的推断统计为主要内容的数理统计学和以传统的政治经济现象描述为主要内容的社会统计学两大学派。

### (一) 数理统计学派

数理统计学派产生于19世纪中叶,创始人是比利时的物理学家和统计学家凯特勒(L. A. J. Quetelet, 1796—1874)。凯特勒最主要的贡献是把概率论正式引进统计学,他师承法国数学家、统计学家拉普拉斯(P. S. Laplace, 1794—1827),主张用研究自然科学的方法研究社会现象,使统计学的理论、内容和方法都发生了质的飞跃;他最先用大数定律论证了社会生活现象纷繁复杂变化不定的偶然性中存在着规律性,并提出了误差理论,用来解决统计上的准确性问题;在学科性质问题上,他明确地认为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。可以认为他是古典统计学的完成者,近代统计学的先驱,也是数理统计学派的奠基人。同时,他还是第一届国际统计会议(1853年)的召集人,因此,他被称为“近代统计学之父”。由于数理统计学主要是在英美等国家发展起来的,故又称英美数理统计学派。

### (二) 社会统计学派

19世纪后半叶,正当英美数理统计学派开始发展的时候,在欧洲又兴起了社会统计学派。

社会统计学派产生于德国,创始人是克尼斯,主要代表人物有恩格尔(C. L. E. Engel,

1821—1896)以及梅尔(C.G.V.Mayer, 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,并认为统计学是以社会总体的规律性为其独立的研究对象;由于社会现象的复杂性和整体性,必须对总体进行大量观察和分析,研究其内在联系,才能揭示社会现象的规律。这是社会统计学派的实质性科学的显著特点。从学术渊源上看,他们融合了国势学派和政治算术学派观点,又继承和发扬了凯特勒强调研究社会现象的传统。由于当时数理统计学派尚未充分发展,社会统计学派便在欧洲大陆占有优势地位,并对日本等国的统计学界都有一定影响。

社会经济的发展,要求统计学提供更多的统计方法,社会科学本身不断地向细化和定量化发展,也要求统计学能提供更有效的调查整理、分析资料的方法。因此社会统计学派也日益重视方法论的研究,出现了从实质性科学向方法论转化的趋势。但是,社会统计学派仍然强调在统计研究中必须以事物的质为前提和认识事物的重要性,这同数理统计学的计量不计质是有本质区别的。

### 三、现代统计学时期

这是指20世纪初至今的统计学发展时期。这一时期科学技术迅猛发展,社会生产发生巨大变化,人类社会经历了两次世界大战,国际政治风云几番突变。统计科学在这一时期也出现了新的分化和组合。

这一时期,数理统计学由于同自然科学、工程技术科学紧密结合,被广泛应用而迅速发展,进入鼎盛时期。到20世纪50年代这一阶段是推断统计学发展最迅速的时期。这期间有影响的理论和大师很多。如英国数学家哥塞特(William Sealy Gosset, 1876—1936)的小样本 $t$ -分布理论;波兰统计学家尼曼(Jerzy splawa Neyman, 1894—1981)以及E. S. 毕尔生等人的假设检验理论及置信区间估计等理论;20世纪40年代的瓦尔德(美A. Wald, 1902—1952)等学者的统计决策理论、多元分布理论等。到了20世纪50年代,经过几代大师的努力,推断统计的基本框架已经建成,并逐渐成为20世纪的主流统计学。

20世纪中期至今的几十年中,是统计学全面发展的阶段。由于受计算机和新兴科学的影响,统计学越来越依赖于计算技术,成为数量分析的方法论科学。这一时期统计学的研究和应用范围越来越广,使得人们在现代统计学史中很难找到权威性的代表人物。当今的统计学家只能限制在有限的专业领域内从事某方面的研究,这是现代统计学的主要特点。如科克伦(W. G. Cochran, 1909—1980)的实验设计理论、安得森(Th. W. Anderson)的复变数分析等。

这一时期,以社会现象为研究对象的社会统计学依然在许多国家存在,并且有所发展,其基本趋势是由实质性科学向方法论转变。如第二次世界大战后社会统计学派的重要人物德国法兰克福大学教授弗拉斯卡姆波(P. Flaskamper, 1886—?),他吸收了英国数理统计学派的通用方法论,把自然科学中的方法应用于社会现象的研究。但是,总的来看,社会统计学发展比较缓慢,这不仅是因为社会现象本身更为复杂,而且作为社会科学它还受到社会政治变故的影响。

特别应该指出的是,这一时期,由于俄国十月社会主义革命胜利,在苏联以及第二次世界大战后的其他社会主义国家逐步建立和发展起来的社会经济统计学。社会经济统计学是以辩证唯物主义和历史唯物主义以及马克思主义政治经济学作为理论指导的,其学说渊源来自古典统计学和凯特勒确立的近代统计学,而且深受德国社会统计学派的影响。社会经济统计学在它产生后的半个多世纪里,实践上曾经为社会主义国家高度集中的计划经济服务,在理论上如分组理论、指数理论等也有不少建树,被认为是统计学史上又一次质的飞跃。

综观统计学的发展历史可以看出,无论是古典统计学、近代统计学,还是现代统计学,其发展过程始终是沿着两条主线展开的:一是以“政治算术学派”为开端形成和发展起来的、以社会经济问题为主要研究对象的社会经济统计;二是以概率论为基础形成和发展起来的、以方法和应用研究为主的数理统计。统计学发展到今天,其应用领域不断扩展,几乎所有的研究领域都要用到统计学。尽管应用领域不同,统计学在不同领域的发展具有各自的特点,但所用的统计方法基本上是相同的。可以说,现代统计学已发展成为一门基础性的方法科学。

### 第三节 统计学分科及与其他学科关系

#### 一、统计学分科

目前,统计方法已被应用到自然科学和社会科学的众多领域,统计学已发展成为由若干分支学科组成的学科体系。一般而言,大致有两种基本的分类:描述统计学和推断统计学;理论统计学和应用统计学。

##### (一) 描述统计学和推断统计学——按统计方法的特点分

描述统计学(descriptive statistics)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,进而通过综合计算及分析等形式来反映客观现象的规律性数量特征。其内容包括统计数据的收集方法、数据的加工整理方法、数据的显示方法、数据分布特征的概括与分析方法等。

统计学研究的另一个方面是利用样本数据推断总体特征。这部分内容则形成了推断统计学(inferential statistics)。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,推断统计学则是现代统计学的主要内容。在对现实问题的研究中,如果收集到的是总体数据,经过描述统计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,则必须运用推断统计。由于我们所获得的数据大多是样本数据,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内容。当然,它仍然必须以提供可靠有效的样本信息的描述统计为基础。

## (二) 理论统计学和应用统计学——按统计研究和应用分

理论统计学主要研究统计学的一般理论和统计方法的数学原理,其理论基础是概率论,它所包括的主要内容有:概率理论、抽样理论、实验设计、估计理论、假设检验理论、决策理论、非参数统计、序列分析、随机过程等。理论统计学是统计方法的理论基础。

将理论统计学的基本原理应用于各个学科领域,就形成了各种各样的应用统计学。如社会统计学、经济统计学、农业统计学、体育统计学、医疗卫生统计学、生物统计学等。应用统计学着重研究统计方法如何应用于某个具体的领域内,如何用统计方法去解决实际问题。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,因而统计方法在各领域的运用中又有不同的特点。

## 二、统计学与其他学科的关系

### (一) 统计学与数学的关系

数学是研究现实世界的空间形式和数量关系的科学,它属于自然科学,是其他科学的数理基础。统计学是计量分析的工具,它离不开数学方法。首先,统计学对客观事物的数量认识要遵循各类事物间的关系的客观性,而数学方法正是对客观数量关系的规律性认识,所以采用数学方法是对数量关系处理和认识的捷径;其次,统计研究经常利用数学模型来进行,通过模型对事物数量关系进行本质的反映,抛开了杂乱的次要因素及随机因素的影响,属于科学的认识方法;再次,统计学中各种特征值的计算都是数学方法的具体体现,所以说统计研究中注意数学方法的应用是科学发展和进步的客观要求。统计学与数学又有着本质的区别。首先,虽然统计学与数学都是研究数量规律的,但数学研究的是抽象的数量规律,而统计学则是研究具体的、实际现象的数量规律;数学研究的是没有量纲或单位的抽象的数,而统计学研究的是有具体实物或计量单位的数据。其次统计学与数学研究中所使用的逻辑方法也是不同的,数学研究所使用的是纯粹的演绎,而统计学则是演绎与归纳相结合,占主导地位的是归纳,统计必须深入实际搜集可靠的数据,并与实际问题相结合,经过科学的归纳才能得出有益的结论。

### (二) 统计学与其他实质性学科的关系

所谓实质性学科,是指这类科学的内容与任务在于揭示客观事物发展变化的规律,以指导人们按客观规律的要求去改造世界,如人口学、财政学、市场营销学、医学等,都属于实质性科学。统计学是一门方法论科学,而不是实质性科学,它研究的是事物普遍存在的数量关系的计量和数量分析的方法,并通过数量分析来认识特定事物的内在规律,但不是研究规律本身。这使得统计学必须与实质性学科相联系,帮助这些学科探索学科内在的数量规律性,而对这些数量规律性的解释并进而研究各学科内在的

规律,只能由实质性学科研究来完成。比如,大量观察法已经发现了新生婴儿的性别比是107:100,但为什么会是这样?形成这一比例的原因应由人类遗传学或医学来研究和解释,而非统计方法所能解释的。由此可以看出,统计方法仅仅是一种有用的定量分析的工具,它不是万能的、是不能解决各学科所有的问题的,但统计方法在这些学科的研究中仍然发挥着越来越重要的作用。

### (三) 经济统计学与经济核算的关系

经济统计学是以经济生活中大量存在的数量特征及数量关系为对象的统计学科,它以经济现象指标体系的设计与核算及指标间的数量关系为主要研究内容,为经济管理和经营分析提供依据和方法。这样,经济统计学就必然以经济核算为其主要的研究内容。而属于经济核算范畴的并不只是统计核算,它主要包括会计、统计和业务三种核算,这三种核算各自独立、相互联系,共同构成经济核算的有机整体。会计核算是以企业的资金运动为对象的微观经济具体的价值量核算;业务核算是以各类业务工作为对象的具体实物量核算;统计核算是以经济生活的总体行为为对象的;不同于前两种具体核算的综合性核算。

经济统计核算主要包括两部分内容,一是国民经济核算,即以一定的经济理论为指导,综合运用统计、会计和数学等方法,对某地区的国民经济各类总量指标及其构成在特定时刻的存量及一定时期内的各类经济流向和流量进行的综合核算。二是企业经济统计核算,即在会计和业务核算的基础上,为满足宏观经济核算和企业自身经营管理的需要,而对企业经济总量及其构成进行的各类存量和流量的核算。

## 第四节 统计学中的几个基本概念

统计学也像其他各门科学一样,在叙述本门科学的理论和方法时,采用一些专门的范畴,这些范畴也是从研究对象所具有的特性中总结出来,并且在研究工作中经常应用到。这里介绍的是统计中常用到的几个基本概念。

### 一、统计总体和总体单位

统计总体是根据一定的目的和要求所确定的研究事物的全体,它是由客观存在的、具有某种共同性质的许多个别事物构成的整体。例如我们要研究全国乡镇工业企业发展情况,那么全国乡镇工业企业就组成一个总体。尽管这些乡镇工业企业所处的地域、生产和销售的产品、生产规模、组织形式等各不相同,但都是从事工业生产活动的乡镇企业,至少在这一方面具有共性。这种共性,就是同质性,这是构成一个总体的必要条件。再如,为研究生产工人的工资水平,则总体中的所有单位必须都是生产工人,而不能包括任何其他成分的人员,否则,研究结果就不能反映生产工人的工资状况。同质性是构成统计总体的基础。所以,统计总体也有同质总体之称。

总体单位是指构成总体的各个单位,它是总体的基本单位,是各项统计数字的原始承担者。我们要了解总体的数量特征,就是从一个个的总体单位调查登记开始的。