

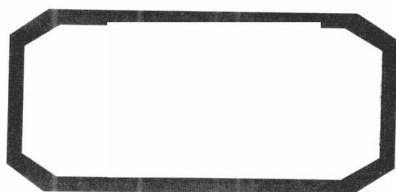
# 统计学原理

TONGJIXUE  
YUANLI  
(第三版)

刘桂荣 编著



华东理工大学出版社  
EAST CHINA UNIVERSITY OF SCIENCE AND TECHNOLOGY PRESS



# 统计学原理 (第三版)

刘桂荣 编著



华东理工大学出版社  
EAST CHINA UNIVERSITY OF SCIENCE AND TECHNOLOGY PRESS

·上海·

## 图书在版编目(CIP)数据

统计学原理 / 刘桂荣编著. —3 版.—上海:华东理工大学出版社,2016.2

ISBN 978-7-5628-4520-1

I. ①统… II. ①刘… III. ①统计学  
IV. ①G8

中国版本图书馆 CIP 数据核字(2016)第 005836 号

---

策划编辑 / 欧聿红

责任编辑 / 陈新征

装帧设计 / 方 雷

出版发行 / 华东理工大学出版社有限公司

地址:上海市梅陇路 130 号,200237

电话:021-64250306

网址:press.ecust.edu.cn

邮箱:press\_zbb@ecust.edu.cn

印 刷 / 上海展强印刷有限公司

开 本 / 787mm×1092mm 1/16

印 张 / 22

字 数 / 534 千字

版 次 / 2016 年 2 月第 1 版

印 次 / 2016 年 2 月第 1 次

定 价 / 49.00 元

---

版权所有 侵权必究

# 第三版前言

统计是作为认识社会和管理国家的工具而产生的,至今已有三百多年的历史。21世纪是信息大爆炸的时代。不管你是否愿意,我们每天都面对大量的信息,而且要在大量的信息中搜寻到对自己有用的信息。而数据信息是信息的主体,绝大多数信息都是以数字的形式存在着的。现代人在工作和生活中,都要从这些大量的、杂乱纷繁的数据中发掘出事物的规律,做出正确的判断,以制订合适的行动方针。

统计学是关于数据的科学和艺术,作为数据处理和分析技术的统计方法已被越来越广泛地应用于科学研究、生产经营管理和普通人的日常生活之中。学习统计知识,掌握统计思想,运用统计方法已是现代社会发展的必然趋势。因此,统计学已经成为我国高校经济管理类各专业的核心课程。

《统计学原理》主要是介绍基本的统计思想、统计理论和统计方法,教会同学们用统计思想思考问题,用统计方法解决某些社会实际问题。文中结合 Excel 软件进行示范,以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。

本教材共分为 8 章。

第一章是**绪论**。介绍统计学的产生和发展、统计学的性质和特点、统计学中的基本概念和中国的政府统计。

第二章是**数据资料的收集**。介绍数据资料收集的方法、收集的数据资料的类型、中国的统计报表制度和四种重要的专项调查的组织方式。

第三章是**数据资料的整理**。介绍如何对大量的数据资料进行加工处理,使数据资料系统化、条理化。本章重点介绍数据整理的技巧和方法以及数据的图形和表格展示方法。

第四章是**数据资料的分析**。介绍描述性统计数据的计算和应用,主要是综合指标、数据的集中趋势和离散趋势的测定。

第五章是**动态数列**。任何现象都是不断发展变化的,本章介绍了对现象发展进行动态水平分析、速度分析、趋势分析和季节变动分析的技巧和方法。

第六章是**指数**。介绍统计指数的种类和意义,指数分析法的应用,常见的重要统计指数。

第七章是**随机抽样方法**。介绍随机抽样调查的种类、意义、组织方式,抽样误差的计算控制,全及指标的推算以及必要样本容量的确定。

第八章是**相关与回归**。介绍相关与回归的概念及种类,重点介绍一元线性相关分析、一元线性回归分析,多元线性相关分析、多元线性回归分析和可线性化的曲线回归分析。

本书作为教材具有以下特点。第一,内容精练、语言通俗、资料翔实。全书介绍了社会经济统计学的基本原理和统计学的入门知识,内容精练。用通俗的语言叙述统计学的概念、

知识点和统计方法,易于理解和掌握。丰富翔实的实证资料可以帮助学习者理解概念、掌握要领、融会贯通。第二,密切联系实际。统计学是一门应用社会科学,只有用好统计分析技术,才是真正理解和掌握了统计学。本教材提供了大量的实例,用发生在身边的事实,解释叙述统计学的概念,理论联系实际,阐述统计思想,提高学习者的学习兴趣。第三,结合 Excel 软件的统计分析功能。Excel 是 Windows 环境下的电子表格系统,它操作简便,具有强大的图表功能和数据分析功能,为复杂的统计分析带来极大的便利。本教材介绍了 Excel 的基本图表功能和相关数据分析功能,使复杂枯燥的数据处理变得生动、直观和形象。

本书第三版保留了第二版的基本理论内容和编写体例,更新了各章的统计实例,新增了我国政府统计改革取得的实践成果,体现了较强的时代特征,同时补充完善了各章的练习题和附录内容,学习内容更加系统。

本书可作为高校经济管理类各专业的教材和参考书,也可作为统计知识爱好者、统计研究工作者的修读书。

本书的编写,参考了大量的现有文献,无法逐一列出这些文献的作者姓名。在此,对这些文献作者表示深深的谢意!

本书的顺利完成得益于以下同仁的大力帮助和支持,他们是马艳梅、陈艳华、刘桂芳。袁国华、倪坚、刘天和袁泉协助完成了本书的相关工作,作者对他们的辛勤劳动表示感谢。

由于时间和水平所限,书中难免存在疏漏和不妥之处,敬请各位同仁、读者给予批评指正,不吝赐教!

编者

2015年10月

# 目 录

<b>第一章 绪论</b> .....	1
<b>第一节 统计学的产生与发展</b> .....	1
一、统计学的产生 .....	1
二、统计学的发展 .....	2
三、统计学的分科 .....	4
<b>第二节 社会经济统计学的性质和特点</b> .....	6
一、统计的含义 .....	6
二、社会经济统计学的性质 .....	7
三、社会经济统计学的特点 .....	8
四、社会经济统计学的研究对象 .....	9
<b>第三节 社会经济统计的作用</b> .....	9
一、社会经济统计的作用 .....	9
二、统计工作的职能 .....	11
三、统计工作过程 .....	12
<b>第四节 统计学中的几个基本概念</b> .....	13
一、总体和总体单位 .....	13
二、标志与变量 .....	14
三、统计指标与指标体系 .....	16
四、流量与存量 .....	19
<b>第五节 中国的政府统计</b> .....	20
一、中国统计的组织体制 .....	20
二、中国的统计公报制度 .....	21
三、统计信息共享和发布制度 .....	22
本章小结 .....	25
练习题 .....	26
<b>第二章 数据资料的收集</b> .....	29
<b>第一节 统计数据的量化尺度和种类</b> .....	29
一、统计数据的量化尺度 .....	29
二、统计数据的类型 .....	32
<b>第二节 统计数据的收集</b> .....	33
一、统计调查的意义 .....	33
二、统计调查的种类 .....	34

第三节	统计调查方案设计	37
一、	明确调查的目的	37
二、	确定调查对象、调查单位和填报单位	38
三、	确定调查项目	39
四、	确定调查时间和调查期限	40
五、	制订调查的组织实施计划	41
第四节	统计报表	42
一、	统计报表的意义和作用	42
二、	统计报表的种类	42
三、	企业一套表联网直报	44
四、	统计报表制度的意义和作用	45
第五节	专项调查	46
一、	普查	46
二、	重点调查	48
三、	典型调查	49
四、	抽样调查	49
第六节	次级资料的收集方法	50
本章小结		51
练习题		51
附录 2-1	2015 年全国 1% 人口抽样调查方案	54
附录 2-2	2015 年全国 1% 人口抽样调查表	57
附录 2-3	2015 年全国 1% 人口抽样调查死亡人口调查表	64
附录 2-4	2015 年全国 1% 人口抽样调查流程	65
<b>第三章</b>	<b>数据资料的整理</b>	<b>66</b>
第一节	数据资料整理的基本问题	66
一、	数据资料的整理	66
二、	数据资料整理的基本步骤	67
第二节	统计分组	68
一、	统计分组的意义	68
二、	统计分组的作用	68
三、	统计分组的方法	70
第三节	频数分布	73
一、	频数分布的概念和种类	73
二、	频数分布表的编制	76
三、	累计频数分布	80
四、	统计频数分布的主要类型	81
第四节	统计资料的图形展示	83
一、	定类数据的图形展示	84
二、	定序数据的图形展示	85
三、	定距数据和定比数据的图形展示	88
四、	数据可视化	91
第五节	统计表	92

一、统计表的作用 .....	92
二、统计表的结构 .....	92
三、统计表的种类 .....	93
四、编制统计表的规则 .....	94
本章小结 .....	95
练习题 .....	95
附录 学习 Excel 软件的基本统计功能 .....	99
<b>第四章 数据资料的分析</b> .....	104
<b>第一节 总量指标</b> .....	104
一、总量指标的概念和作用 .....	104
二、总量指标的分类 .....	106
三、总量指标的计量单位 .....	108
<b>第二节 相对指标</b> .....	109
一、相对指标的概念和表现形式 .....	109
二、相对指标的种类和计算方法 .....	111
三、计算和运用相对指标应注意的问题 .....	121
<b>第三节 平均指标</b> .....	122
一、平均指标的概念、特点和作用 .....	122
二、算术平均数 .....	124
三、调和平均数 .....	128
四、几何平均数 .....	130
五、众数 .....	131
六、中位数 .....	133
七、各种平均指标之间的关系 .....	135
<b>第四节 标志变动指标</b> .....	136
一、标志变动指标的概念 .....	136
二、全距 .....	137
三、四分位差 .....	138
四、平均差 .....	140
五、标准差 .....	142
六、离散系数 .....	144
本章小结 .....	145
练习题 .....	146
附录 4-1 利用 Excel 软件计算描述统计指标 .....	150
附录 4-2 利用 Excel 的函数功能计算各种描述统计指标 .....	152
<b>第五章 动态数列</b> .....	155
<b>第一节 动态数列的概念、种类和编制原则</b> .....	155
一、动态数列的概念 .....	155
二、动态数列的种类 .....	156
三、动态数列的编制原则 .....	158
<b>第二节 动态分析的水平指标</b> .....	159
一、发展水平 .....	159

二、平均发展水平 .....	160
三、增长量 .....	166
四、平均增长量 .....	168
第三节 动态分析的速度指标 .....	168
一、发展速度 .....	168
二、增长速度 .....	170
三、平均发展速度 .....	171
四、平均增长速度 .....	173
第四节 动态数列的长期发展趋势分析 .....	174
一、动态数列的影响因素 .....	174
二、长期趋势的测定 .....	174
三、时距扩大法 .....	175
四、移动平均法 .....	176
五、最小平方法 .....	178
第五节 季节波动的测定 .....	181
一、季节波动的概念 .....	181
二、按月平均法 .....	181
三、移动平均趋势剔除法 .....	184
本章小结 .....	186
练习题 .....	187
附录 5-1 利用 Excel 软件制作动态数列数据的折线图 .....	192
附录 5-2 利用 Excel 软件的数据分析功能计算移动平均值 .....	193
附录 5-3 利用 Excel 软件进行季节变动分析 .....	195
附录 5-4 利用 Excel 软件计算按月平均法的季节指数 .....	197
附录 5-5 利用 Excel 软件的公式复制功能进行简便计算 .....	198
<b>第六章 指数</b> .....	<b>200</b>
第一节 指数的概念 .....	200
一、指数的概念 .....	200
二、指数的作用 .....	202
三、指数的分类 .....	203
第二节 个体指数 .....	205
一、个体数量指数 .....	205
二、个体质量指数 .....	205
第三节 综合指数 .....	207
一、综合指数的概念和特点 .....	207
二、数量指标综合指数的编制 .....	208
三、质量指标综合指数的编制 .....	210
第四节 平均数指数 .....	212
一、平均数指数的概念 .....	212
二、加权算术平均数指数 .....	213
三、加权调和平均数指数 .....	214
第五节 平均指标对比指数 .....	216

一、平均指标对比指数的概念 .....	216
二、平均指标对比指数的应用 .....	217
第六节 指数体系及因素分析法 .....	219
一、指数体系的概念 .....	219
二、总量指标的两因素指数分析法 .....	220
三、总量指标的多因素指数分析法 .....	222
四、平均指标的因素分析法 .....	224
五、指数推算 .....	227
第七节 几种常用的价格指数 .....	228
一、商品零售价格指数 .....	228
二、居民消费价格指数 .....	229
三、制造业采购经理指数 .....	231
四、中国非制造业商务活动指数 .....	232
五、上海证券交易所的证券交易价格指数 .....	233
本章小结 .....	239
练习题 .....	239
<b>第七章 随机抽样方法</b> .....	245
第一节 抽样调查的基本问题 .....	245
一、抽样调查的概念和特点 .....	245
二、抽样调查的应用范围 .....	247
三、抽样调查的几个基本概念 .....	248
四、抽样调查的理论依据 .....	252
第二节 抽样误差 .....	253
一、抽样误差的概念 .....	253
二、抽样平均误差的意义和作用 .....	254
三、抽样平均误差的计算 .....	254
四、影响抽样平均误差的因素 .....	258
五、抽样极限误差 .....	259
第三节 抽样估计 .....	259
一、抽样估计的概率度和可靠程度 .....	259
二、抽样估计的方法 .....	261
三、全及总体指标的推算 .....	262
第四节 随机抽样的组织方式 .....	264
一、抽样调查的组织程序 .....	264
二、简单随机抽样 .....	265
三、类型抽样 .....	266
四、等距抽样 .....	268
五、整群抽样 .....	271
六、多阶段抽样 .....	273
第五节 必要样本容量的确定 .....	277
一、必要样本容量的含义 .....	277
二、必要样本容量的计算公式 .....	278

本章小结 .....	281
练习题 .....	282
<b>第八章 相关与回归</b> .....	<b>286</b>
<b>第一节 相关分析</b> .....	<b>286</b>
一、现象之间的关系 .....	286
二、相关关系的特点 .....	288
三、相关关系的种类 .....	288
四、相关分析的主要内容 .....	290
<b>第二节 简单线性相关分析</b> .....	<b>291</b>
一、线性相关关系的判定 .....	291
二、线性相关系数 .....	293
三、线性相关系数的性质和应用 .....	295
<b>第三节 简单线性回归分析</b> .....	<b>295</b>
一、回归分析的概念 .....	296
二、相关分析与回归分析的关系 .....	296
三、一元线性回归分析 .....	297
四、回归系数的计算 .....	297
五、估计标准误差 .....	299
六、判定系数 .....	300
七、相关系数 $r$ 和估计标准误差 $S_y$ 的关系 .....	300
八、相关系数 $r$ 与回归系数 $b$ 之间的关系 .....	301
九、回归方程的应用 .....	302
<b>第四节 多元线性回归分析</b> .....	<b>303</b>
一、多元线性回归模型 .....	303
二、多元线性回归估计标准误差 .....	304
三、复相关系数 .....	304
<b>第五节 可线性化的曲线回归</b> .....	<b>306</b>
一、二次曲线(抛物线) .....	306
二、指数曲线 .....	307
三、幂函数 .....	307
四、双曲线 .....	308
五、对数函数 .....	308
六、多项式回归 .....	309
本章小结 .....	311
练习题 .....	311
附录 8-1 利用 Excel 软件计算相关系数 .....	315
附录 8-2 利用 Excel 软件计算回归系数 .....	317
附录 8-3 利用 Excel 软件制作散点图、判断相关性 .....	318
<b>附录 常用英汉统计词汇</b> .....	<b>320</b>
<b>参考答案</b> .....	<b>326</b>
<b>参考文献</b> .....	<b>341</b>

## 绪 论

### 08 本章学习目的

理解“统计”这个名词的含义。了解统计学的产生和发展过程。学习三个重要的统计学流派。掌握社会经济统计学的研究对象和特点。熟练掌握统计中的几个基本概念。了解中国的政府统计。

### 第一节 统计学的产生与发展

#### 一、统计学的产生

统计学(Statistics)作为一门社会学科距今已有三百多年的历史。同任何一门学科的发展规律一样,必须先有社会实践活动,然后才有对社会实践活动的理论概括和经验总结。统计的社会实践活动最早产生于还没有书面文字的原始社会。原始社会是人类社会发展的第一个阶段,人们过着简单的部落群居生活,共同劳作,共同分享劳动成果。在对劳动成果进行记录和分配时,群居生活的部落首领必须知道本部落的人数、待分配的各种食物以及劳动成果,这就需要知道事物“量”的特征。在没有文字和数字用以表示事物“量”的特征时,古人想出了比较形象和直观的“量”的表示方式,出现了“结绳记数”和“刻石记数”。如《周易正义》中写道:“事大,大结其绳;事小,小结其绳;结之多少,随物众寡”,这就是原始社会人们对社会现象“量”的方面进行的简单的记录和表达,于是产生了最早的统计萌芽。

随着人类社会的发展,人们的社会实践活动日益丰富,社会生产关系和生产力得到了极大的发展,人类社会随之进入了奴隶社会和封建社会。原始社会的部落首领发展成奴隶社会的奴隶主和封建社会的封建主,奴隶主和封建主们为了加强自己的统治和扩大自己的势力,开始了对人口、土地和财产的基本记录和计量,统计实践活动的内容更加丰富。对事物“量”的表达和掌握事物“量”的特征已经成为对社会现象进行认识和加强管理的手段之一。

同时,国家为了加强对内统治和对外防御战争,需要进行征兵、征税等,因而统计已成为

一项国家管理工具。例如,早在四千多年前的夏禹王朝,为了治国治水的需要,曾进行初步的国情统计,查明当时我国有人口约 1 355 万人,土地约 2 431 万顷,并将全国分为九州。约公元前 3000 年,埃及人为了建造金字塔,在全国对人口、财产也进行过统计。

这个时期的统计活动,范围相对狭窄,仅限于对人口、土地、财产等重要社会财富的记录和计量,统计活动是人类社会生产实践活动的附加活动,统计工作仅仅是生产者的附加职能,社会分工中还没有出现专职的统计工作人员,更没有专业的统计研究机构,公开发布各种统计数据信息。这是由当时的社会生产力和生产关系的发展程度决定的。

## 二、统计学的发展

统计活动产生于人类的社会生产实践活动,统计活动的理论经验在社会实践活动中经过提炼和升华,发展成统计学。统计学是随着人类社会的发展和国家管理的需要而发展起来的。

统计取得广泛迅速的发展是在资本主义社会。资本主义社会的社会分工更加细化,生产力得到快速发展,生产关系更加复杂,资本家追求利润最大化的动机更强。社会对情报、信息和统计数据的需求日益增长。对统计的需求已不限于人口、土地、财产等社会财富范畴,它逐步扩展到了更为广泛的领域,如贸易、交通运输、工业生产、商业流通、教育、银行、保险、劳动、就业、医疗、文化等,生产者需要收集社会生产和社会生活等各个方面的数据信息,作为制定生产经营计划的依据。由于社会对统计需求的增长,统计最终成为一种专门的社会职业,统计从生产者的附加职能中分离出来,形成了各种专业的社会经济统计。

1830—1849 年,是欧洲的“统计狂热”时期,各国相继成立了统计机关和统计研究机构。1853 年,在布鲁塞尔举办了首届国际统计会议(International Statistical Congress)。这种有组织的国际统计会议逐渐发展成目前的国际统计学会(International Statistical Institution)。此外,17 世纪之后,在统计学的发展历史上出现了几个比较重要的统计学流派,主要有政治算术学派、国势学派和数理统计学派。

**政治算术学派。**政治算术学派产生于 17 世纪中叶的英国,主要代表人物是约翰·格朗特(John Graunt, 1620—1674)和英国著名的古典经济学家威廉·配第(William Petty, 1623—1687)。16 世纪末到 17 世纪初,英国发生了几次鼠疫大流行,伦敦市民对出生、死亡、结婚、洗礼等含有大量数字的报告变得关心起来。约翰·格朗特根据寺院提供的关于死亡和洗礼的资料,编制了世界上第一张“死亡率”统计表,并出版了《关于伦敦死亡表的观察》。威廉·配第继承和发展了格朗特的研究工作,根据对人口、土地、财政、经济等各方面的大量观察,完成了《政治算术》(1676)一书。这里的“政治”是指政治经济学,“算术”是指统计方法。在这部著作中,配第第一次运用数字、重量、尺度等计量和比较的方法,运用大量的统计资料,将英国的国力与法国、意大利、荷兰等国家进行了比较研究,提出了一套在当时来说是较为系统的数量对比分析方法,用于对社会经济现象进行数量性的描述和比较分析,创立了政治算术学派。经过几个世纪的发展和完善,政治算术学派发展成现代的社会经济统计学。

政治算术学派是统计学发展的开端。此后,在人口统计学方面,由于研究对象的不同,出现了保险统计、卫生统计等统计学的分支;在经济统计方面,出现了农业统计、工商统计和物价指数计算方法的研究等。

政治算术学派虽然被称之为统计学的真正起源,但是,这一学派没有把对社会客观现象

的数量研究方法叫作“统计学”，它“有实无名”。“统计学”这个名称实则源自“国势学派”。

**国势学派(又称记述学派)**。国势学派产生于18世纪的德国,创始人是赫尔曼·康林(Herman Conring, 1606—1681)。他对欧洲许多国家的人口、国土面积、政体、财政、军备等方面进行了文字性的记述,并在大学开设“国势学”课程。18世纪德国政治学教授格特弗里德·阿亨瓦尔(Gottfried Achenwall, 1719—1772)继承和发展了康林的“国势学”,1749年,阿亨瓦尔出版了其代表作《近代欧洲各国国势学概论》。康林和阿亨瓦尔都比较全面地记述了欧洲各国的基本情况,如人口、领土、政治结构、议会、军队、财政、经济、宗教、地理、风俗、货币、艺术等,并且用对比分析的方法研究了解国家组织、领土、人口、资源财富和国情国力,比较了欧洲各国实力的强弱,为德国的君主政体服务。因“国势”与“统计”在外文中词义相通,后来正式命名为“统计学”。该学派在进行国势比较分析中,偏重事物性质的解释,而不注重数量对比和数量计算,但却为统计学的发展奠定了经济理论基础。“统计学”的名字一直沿用至今。

**数理统计学派**。数理统计学派产生于19世纪的比利时,主要代表人物是凯特勒(A. Quetelet, 1796—1874),他的代表作是《社会物理学》。凯特勒发展了政治算术学派,在应用数量观察分析方法的同时,将统计方法应用于社会生活的各个方面,可以说是开创了统计学的新纪元。此外,凯特勒还率先将概率论引入到社会现象的统计研究之中,提出了社会现象的发展并非偶然,而是具有内在规律性的观点,并且提出了关于统计学的新概念。通过他的努力,统计学的方法获得了普遍应用。19世纪60年代,凯特勒又进一步将国势学、政治算术、概率论的科学方法结合起来,使之形成近代应用数理统计学,创立了“数理统计学派”,被后人称为“现代统计学之父”。凯特勒还是比利时国家统计工作的领导人,是国际统计学术会议的倡导者和组织者,在统计学的发展中,做出过巨大贡献,产生了重要影响。

在社会发展进程中,各种统计学派相互影响,相互汲取其他学派的有益之处,相互促进,共同发展,不断完善。同时概率论等数学学科的发展,也丰富和发展了统计学的数学方法,使统计学逐渐发展成一门社会科学。前文三个主要的统计学流派分别发展成现代的社会经济统计学、自然技术统计学和数理统计学。

20世纪50年代之后,随着数学和计算机技术的发展,统计理论、统计方法和统计学的应用都进入了一个全面发展的新阶段。特别是计算机技术的发展与普及,为统计学在应用上的普及开拓了广阔的前景,新的研究领域层出不穷。无论在自然科学、社会科学还是在农业、林业、医药、教育等领域,统计学都成为不可缺少的分析工具和管理工具。

21世纪是信息大爆炸的时代,政府、企业、家庭和个人即是信息生产者,也是信息需求者,互联网加快了信息的传播速度,互联网和云技术改变了信息的存储方式,加大了信息存储量。信息是有经济价值的,在瞬息万变的信息时代,最先掌握信息的人具有信息优势,可以在市场赢得先机,实现收益最大化或利益最大化,而信息最主要的表现方式即是各种统计数字,比如国内生产总值GDP、居民消费价格指数CPI、生产者价格指数PPI、制造业采购经理指数PMI、股票价格指数等。

在现代经济金融社会中,统计学更是一门非常重要的管理工具和分析工具,掌握基本的统计分析技术和了解统计数据的真实意义,对于国家管理、企业管理和个人及家庭生活都将产生重要的影响。我国著名经济学家马寅初先生曾经说过:“学者不能离开统计而研究,政治家不能离开统计而施政,事业家不能离开统计而执业。”统计是窥测奥秘的工具,是使人聪明的技术。对统计学的一知半解常常造成不必要的上当受骗,对统计学的一概排斥往往造

成不必要的愚昧无知。现代社会,人人都要有统计思维,都要学习一些基本的统计方法和统计分析技术,统计无处不在。

### 三、统计学的分科

统计学产生于人们的统计社会实践活动,是对人类社会统计实践活动经验的高度总结和概括,是对统计实践活动经验的归纳和提炼,用于指导人们更好地进行各种统计社会实践活动。《不列颠百科全书》把统计学定义为“收集和分析数据的科学和艺术,其中标出统计方法的艺术”。维基百科(Wikipedia)对统计学的解释是:统计学是在资料分析的基础上,自17世纪中叶产生并逐步发展起来的一门学科。它是研究如何测定、收集、整理、归纳和分析反映数据资料,以便给出正确讯息的科学。

我们认为,统计学是一门关于数据的科学和艺术,研究有关数据的收集、整理和分析的基本方法,探究数据的内在规律和数量关系,研究数据的展示方法,从数量方面达到对客观事物的科学认识。

统计广泛地应用在各门学科,从自然科学、社会科学到人文学科,甚至被用在工商业及政府的情报决策之上。随着大数据(Big Data)时代的来临,统计的面貌也逐渐改变,与资讯、计算等领域密切结合,是数据科学(Data Science)中的重要主轴之一。统计学在发展过程中,经过不断的进化和完善,吸收了数学和概率论的基本方法,研究对象和研究方法也不断细化,形成了不同的分支学科。

#### (一) 描述统计学(Descriptive statistics)和推断统计学(Inferential statistics)

根据统计学的研究方法构成不同,统计学可分为描述统计学和推断统计学。描述统计学产生于20世纪20年代。生物学家达尔文在研究生物的遗传变异等进化理论时,采集了生物、地质、地理等方面的大量标本和化石,在收集、整理和权衡事实的过程中,借助了简单的统计方法,引起了很多人的兴趣。此后,一些生物学家兼科学家,将生物进化和统计研究相结合,使生命科学的研究从定性分析迈向定量分析,开创了生物统计学派。

生物统计学派的主要代表人物是高尔顿(F. Galton, 1822—1911)和皮尔逊(K. Pearson, 1857—1936)。高尔顿是达尔文的表弟,达尔文《物种起源》这一巨作的问世,触动和激发了高尔顿用统计方法去精确地研究智力遗传的进化问题,奠定了其生物统计学派创始人的地位。高尔顿努力探索那些能把大量数据加以描述与比较的方法和途径,结果出色地引入了诸如中位数、百分位数、四分位数、四分位差以及分布、相关、回归等重要的统计学概念与方法,对变异、相关、回归有独到的见解,为统计学的发展做出了巨大的贡献。皮尔逊是生物统计学派的主要干将,他继承和发展了高尔顿的统计思想。皮尔逊致力于生物统计学、优生学和遗传学的统计方法研究,对一般生物现象进行数量描述,创造了许多统计学的用语。例如,频数分布表、频数分布图、相关系数的计算公式、回归系数的计算公式、卡方检验等。可以说,统计中的大部分内容都是由皮尔逊整理出来的,大部分统计术语也是由他命名的,是他使描述统计学不断地发扬光大。

以皮尔逊为首的生物统计学派在研究进化遗传等生物现象时,他们的许多统计方法也能适用于社会、经济、人口等方面的统计研究。因此,生物统计学逐步被统计学家们抽象化

而成为在自然科学、社会科学的广泛领域中都能通用的方法,并构成了统计学的一个极重要的组成部分。因此,现代西方统计学界把 20 世纪 20 年代样本推断统计学以前的、以大样本观察为基础的一整套统计理论与方法称为“描述统计学”。

从生物统计学派发展而来的描述统计学是以大样本近似分布为基础,重点研究数据的收集整理和对数据进行直观的描述。它的基本研究方法即是大量观察法和综合归纳法。把大量的繁杂的统计数据进行归纳整理,提炼出说明现象本质特征及规律的有用信息,是描述统计学的中心内容。描述统计学用诸如排序、分组、列表、制图等整理手段以及百分比、平均数、标准差、相关系数、回归系数等计算手段,把大量的数据简捷地浓缩到可以利用的程度。对收集到的大量数据进行加工、整理、列表、图示以及计算综合指标,用以探索数据内在的数量规律性是描述统计学的基本特征。

20 世纪 30 年代,中国的英美留学生相继回国,带回了国外先进的统计思想和统计方法,描述统计学也随之传入我国。我国最早接触描述统计学的是北洋政府时期的顾澄教授,他于 1913 年翻译出版了皮尔逊的《统计学原理导论》。二十世纪二三十年代后,传入我国的描述统计学的译作日渐增多,为描述统计学在我国的传播和应用打下了基础。我国较早以描述统计学家闻名的有刘大钧、朱君毅、金宝国三人,他们是我国最早参加国际统计学会并作为代表参加国际统计学术会议的学者。

推断统计学诞生于 1920 年前后,主要代表人物是戈瑟特(W. S. Gosset, 1876—1937)和费希尔(R. A. Fisher, 1890—1962)。推断统计学主要研究如何根据样本数据去推断总体未知数量特征的方法,它是在对样本数据进行描述的基础上,对总体未知的数量特征做出以概率形式表述的推断。特别是建立在小样本基础上的推断统计学方法成为推断统计学的主流。

描述统计学和推断统计学的划分,既反映了统计方法发展的两个不同阶段,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。由于推断统计学主要应用概率论和数理统计的基本方法,相对比较复杂,因此本书主要介绍描述统计学的基本方法和应用。

## (二) 理论统计学(Theoretical statistics)和应用统计学(Applied statistics)

根据统计学研究方法和统计方法的应用范围不同,统计学可分为理论统计学和应用统计学。理论统计学主要研究统计学的基本原理和基本方法,研究如何将数学原理和计算机技术应用于统计学,以发展出新的统计学方法和技术。理论统计学的发展推动和完善了统计学学科的发展。理论统计学是统计方法论的理论基础,没有理论统计学,统计学科的发展将止步不前。

应用统计学主要是研究如何将统计学的方法和原理与实际问题相结合,使用统计学的方法解决实际问题。统计学是一门关于数据的科学,现实世界中任何现象都表现为“量”和“质”两个方面,其“量”的规定性可以通过统计学的方法进行认识和把握,统计学方法的应用几乎扩展到了所有的科学研究领域。例如,将统计学的方法应用于生物学的研究,形成了生物统计学;应用于医学研究,形成了医疗卫生统计学;应用于商业,形成了商务统计学;应用于旅游业,形成了旅游统计学;应用于现代金融服务业,形成了金融统计学等。作为现代经济社会中的各个主体,都必须学习一定的统计学知识,而掌握一定的统计学知识将使你受益无穷。

此外,统计学作为一门方法论科学,能够和其他学科交叉结合,形成新的统计学分支,如

社会经济统计学、农业统计学、管理统计学、商业统计学、贸易统计学、人口统计学等。

## 第二节 社会经济统计学的性质和特点

### 一、统计的含义

18世纪,出现了“统计学”这个名词,现代社会中“统计”一词随处可见,每个人都离不开统计,从国家行政机关的构成来看,世界各国都设有专业的统计机构。中国最高的统计研究机构和统计数据信息发布机构是中华人民共和国国家统计局,各级地方政府都设有专业的统计工作机构,基层单位设有专职的统计工作人员。为了加强企业管理,各企业内部也设有统计部门或专职统计人员。“统计”一词到底有什么含义呢?根据国际统计惯例和规则,“统计”一词有三种含义——统计活动、统计资料和统计科学。

#### (一) 统计活动

统计的三种含义是密切联系的,统计活动或统计工作是指统计工作人员对社会经济现象的数量方面进行记录、收集、整理和分析活动的过程,它是一种社会调查研究活动。例如,基层单位的统计工作人员对本单位工人出勤率的记录和分析活动;基层单位的统计工作人员对本单位的产品生产、库存、销售所进行的记录和分析工作;各级政府统计研究机关的各项工作活动;社会调查机构开展的各种社会调查工作等。统计活动或统计工作是统计分析和统计研究的基础,是开展各项统计工作的第一步。

统计工作的成果即是统计资料,包括各种原始台账、调查数据、调查表格,根据调查数据绘制的各种图表及对社会经济现象进行的定性分析和定量分析等。

#### (二) 统计资料

统计资料是反映社会经济现象发展变化规律的重要的数据资料,对于描述现象发展的现状和发展规律、进行发展趋势的预测、开展科学研究、加强管理和国际交流、制定宏观经济政策和企业发展战略具有非常重要的意义。统计资料是统计活动或统计工作的最终成果,数据或图表是现代统计资料的主要表现形式,借助于计算机技术和统计分析软件,可以对统计资料进行深入的分析,并撰写统计分析报告。我国各级政府统计部门每年都要定期发布反映国民经济和社会发展基本情况的统计公报,国家统计局每月都要发布中国宏观经济运行的统计数据,为个人消费和投资、企业投资、生产经营管理和银行信贷的发放提供依据。

统计资料还是统计信息的主要内容,在当今信息化时代,信息的表现形式主要是各种统计数据。数据是一种资源,据互联网数据中心预计,到2020年,全世界所产生的数据规模将达到现在的44倍,无处不数据,大数据时代已经来临。

政府机关通常是国家统计数据信息的发布者,有关国家和地区国民经济和社会发展变动的统计信息数据是一种公共产品。例如,中国国家统计局发布的数据表明:2014年,我国城镇新增就业1322万人,比2013年增加12万人;城镇登记失业率为4.09%,比2013年上升0.04个百分点;城镇居民人均可支配收入为28844元,扣除价格因素实际增长6.8%,增速比2013年