



普通高等教育“十三五”规划教材



◎ 李晓峰 刘 馨 编著

应用统计学

非
外
借



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>



普通高等教育“十三五”规划教材

统计分析系列

应用统计学

李晓峰 刘 馨 编著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书将统计学原理、统计分析方法与具体的实际问题有机结合起来,强调理论性与实践性并重、知识与可读性并重。书中大部分统计分析方法都给出了 SPSS 软件的具体操作步骤与数据处理结果的解析,强化软件的应用能力及解决实际问题的能力。

本书遵循“循序渐进”和“学以致用”的原则构建知识体系。全书系统地阐述了以下内容:总论、统计数据的搜集、统计数据的整理、统计数据特征的描述、参数估计、参数假设检验、方差分析、相关与回归分析、时间序列分析与预测、统计指数。为方便教师教学、学生学习,本书提供课件、实例数据,读者可登录华信教育资源网 www.hxedu.com.cn 免费下载使用。需要习题参考答案的任课教师,可发邮件至 scuxfl@163.com 或 qinshl@phei.com.cn 申请获得。

全书体系完备,实例丰富,深入浅出,通俗易懂,应用性强,可作为高等院校经济、管理等专业的本科生、研究生的教学用书,也可作为其他对于统计学感兴趣的各界人士的自学参考书。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

应用统计学 / 李晓峰, 刘馨编著. —北京: 电子工业出版社, 2017.8

(统计分析系列)

ISBN 978-7-121-31762-0

I. ①应… II. ①李… ②刘… III. ①应用统计学—高等学校—教材 IV. ①C8

中国版本图书馆 CIP 数据核字(2017)第 124079 号

策划编辑: 秦淑灵

责任编辑: 秦淑灵

印 刷: 北京季峰印刷有限公司

装 订: 北京季峰印刷有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1092 1/16 印张: 16.5 字数: 422.4 千字

版 次: 2017 年 8 月第 1 版

印 次: 2017 年 8 月第 1 次印刷

定 价: 38.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010)88254888, 88258888。

质量投诉请发邮件至 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式: qinshl@phei.com.cn。

前 言

应用统计学属于统计学的一个范畴,它采用统计学中的有关方法来对自然现象、社会经济现象或管理现象中研究对象的数量特征、数量关系、发展变化趋势及规律进行研究,最终解决实际问题。应用统计学以数理统计学的理论和方法为基础,不断吸收信息论、控制论、系统论和决策论等方面的研究成果,使统计职能从反映和监督拓展到推断、预测和决策。目前,应用统计学被广泛应用于企业管理、市场调查、投资评估、风险分析、经济预测等方面,为管理者进行正确决策提供科学的依据。总之,应用统计学是认识自然现象、社会现象及管理现象的有力武器,是宏观管理和微观管理的重要工具。

作为经济、管理等学科的核心课程,国内外目前广泛采用的应用统计学教材的理论体系较为完整,但在实际应用方面所做的工作还比较薄弱,在软件的实际操作中介绍的还不深入。本书试图将统计学原理、统计分析方法与具体的经济、管理问题结合起来,强调理论性与实践性并重;同时,以 SPSS 软件应用为导向,强化解决实际问题的能力。

在这本书中系统地阐述了以下内容:总论、统计数据的搜集和整理、统计数据特征的描述、参数估计、参数假设检验、方差分析、相关与回归分析、时间序列分析与预测、统计指数等。本书在借鉴、总结目前国内外相关教材的基础上,重点突出了以下“特点”:

(1) 运用通俗易懂、深入浅出的语言,对应用统计学的基本思想做了比较完整的介绍。与一般的统计学教材相比,增加了参数假设检验、方差分析等实用性较强的内容。

(2) 结合统计方法的介绍,精心选编了大量的例题,以便学生能够将方法的学习和方法的应用相互结合起来,在进一步熟悉统计方法及其原理的同时,也能学会和掌握统计方法应用的技巧。

(3) 详细而又清晰地介绍了 SPSS 软件在相关统计方法中的应用,强调应用 SPSS 软件来处理统计数据,并能正确地理解分析结果。

(4) 各章都附有相应的习题,以帮助读者更好地理解和应用统计学中的方法,拓展视野,启迪思维。

为方便教师教学、学生学习,本书提供课件、实例数据,读者可登录华信教育资源网 www.hxedu.com.cn 免费下载使用。需要习题参考答案的任课教师,可发邮件至 scuxfl@163.com 或 qinshl@phei.com.cn 申请获得。

本书体系完备,布局合理,实例丰富,应用性强,可作为高等院校相关专业的本科、研究生等的教学用书,也可作为从事统计、市场调查等实际工作的各类经济管理人士的参考用书或培训教材。

本书为四川大学立项建设教材,获得了相关经费的资助,并得到了四川大学教材建设委员会有关同志的大力支持,在此表示衷心的感谢!同时,在本书的编写过程中,四川大学商学院的多位教师和研究生付出了大量的劳动,在此一并致谢。他们是:钟胜、刘海涛、田旺等。

由于作者水平有限，本书中难免有疏漏和差错的地方，恳请读者给予谅解并批评指正，以便进一步修改提高。

编著者
2017年8月于四川大学

目 录

第一章 总论	1
第一节 统计学的产生和发展	1
一、统计与统计学	1
二、统计学的产生	2
三、统计学的发展	3
四、统计学和其他学科的关系	4
第二节 统计学的研究对象	5
一、研究对象是群体现象	5
二、探索群体现象数量表现的内在规律性	5
三、对群体现象数量特征进行计量描述和统计推断	7
第三节 统计学分类及性质	7
一、描述统计学和推断统计学	7
二、理论统计学和应用统计学	9
第四节 统计学研究方法	10
一、大数定律的方法论	10
二、统计研究基本方法	11
第五节 统计学的基本范畴	12
一、总体和单位	13
二、标志与变量	14
三、统计指标	16
第六节 统计学的应用领域	18
一、市场调查	18
二、人口管理	18
三、企业管理	19
四、投资分析	19
五、期货交易	19
六、风险管理	19
思考与练习	20
第二章 统计数据的搜集	21
第一节 数据的计量和类型	21
一、数据的计量方法	21
二、数据的分类	23
第二节 统计调查的方法	23

一、统计数据来源	23
二、统计调查方案	24
三、统计调查方式	27
四、统计调查误差	30
第三节 问卷设计	31
一、问卷基本结构	31
二、问卷设计原则	32
三、问卷设计步骤	33
四、问卷设计技术	34
第四节 用 SPSS 建立数据文件	36
一、SPSS 的启动	36
二、数据的编辑窗口	37
三、数据文件的建立	37
四、数据文件的打开	41
五、数据文件的编辑	42
思考与练习	43
第三章 统计数据的整理	44
第一节 统计数据整理概述	44
一、统计数据整理的意义	44
二、统计数据整理的内容	44
三、统计数据整理的程序	44
第二节 统计分组	45
一、统计分组的概念与作用	45
二、统计分组的种类	46
三、统计分组的方法	47
四、频数与频率	50
第三节 统计图表	52
一、统计表	52
二、统计图	54
第四节 用 SPSS 整理数据	58
一、给变量值排序	59
二、给变量值排名次	59
三、频次及频率分析	60
四、变量重新赋值	62
思考与练习	65
第四章 统计数据特征的描述	66
第一节 总量指标	66
一、总量指标的作用	66
二、总量指标的分类	67

三、总量指标的计算	67
第二节 相对指标	68
一、相对指标的作用	68
二、相对指标的分类	68
三、相对指标的计算	69
第三节 平均指标	71
一、平均指标的作用	71
二、平均指标的分类	71
三、平均指标的计算	72
第四节 变异指标	81
一、变异指标的作用	81
二、变异指标的种类	82
三、变异指标的计算	82
第五节 偏态与峰态的度量	86
一、偏态及其度量	86
二、峰态及其度量	87
第六节 用 SPSS 描述统计数据特征	88
一、频率分析模块介绍	88
二、举例说明	90
思考与练习	91
第五章 参数估计	93
第一节 参数估计基本概念	93
一、全及总体和抽样总体	93
二、总体指标和样本指标	93
三、重复抽样和不重复抽样	94
第二节 统计量与抽样分布	94
一、统计量	94
二、抽样分布	94
三、与抽样分布有关的定理	96
第三节 点估计	97
一、估计量的概念	97
二、评价估计量的标准	98
第四节 区间估计	99
一、区间估计的基本问题	99
二、总体均值的估计区间	99
三、总体成数的区间估计	102
四、总体方差的区间估计	103
第五节 样本容量的确定	104
一、测定平均数的样本单位数	104

二、测定成数的样本单位数	105
第六节 用 SPSS 作参数估计	106
一、 μ 和 σ^2 的点估计	106
二、总体均值的置信区间的计算	108
思考与练习	112
第六章 参数假设检验	113
第一节 假设检验的基本概念	113
一、假设检验的基本问题	113
二、假设检验的一般步骤	115
第二节 总体均值的假设检验	116
一、单个总体均值的假设检验	116
二、两个总体均值差异的假设检验	119
三、配对样本的 T 检验	122
第三节 总体成数的假设检验	123
一、单个总体成数的假设检验	123
二、两个总体成数之差的假设检验	124
第四节 总体方差的假设检验	125
一、单个总体方差的假设检验	125
二、两个总体方差比的假设检验	127
第五节 用 SPSS 作假设检验	128
一、均值比较过程	128
二、单样本的 T 检验	130
三、独立样本 T 检验	131
四、配对样本 T 检验	135
思考与练习	137
第七章 方差分析	138
第一节 方差分析引论	138
一、问题的提出	138
二、方差分析的基本思想	139
三、方差分析的基本假定	139
第二节 单因素方差分析	140
一、样本数据结构	140
二、单因素方差分析的步骤	140
三、关系强度的度量	143
四、单因素方差分析中的多重比较	144
第三节 双因素方差分析	144
一、双因素方差分析及其类型	144
二、无交互作用的双因素方差分析	145
三、有交互作用的双因素方差分析	147

第四节 用 SPSS 作方差分析	150
一、单因素方差分析的实现步骤	150
二、单因素方差分析的结果及其解释	151
三、无交互作用的双因素方差分析的实现步骤	153
四、无交互作用的双因素方差分析的结果及其解释	155
五、有交互作用的双因素方差分析的实现步骤	155
六、有交互作用的双因素方差分析的结果及其解释	157
思考与练习	159
第八章 相关与回归分析	161
第一节 相关分析	161
一、相关关系的概念	161
二、相关关系的种类	161
三、相关表与相关图	162
四、相关系数	163
第二节 一元线性回归分析	167
一、回归分析与相关分析的关系	167
二、一元线性回归模型	167
三、一元线性回归模型的估计	168
四、一元线性回归模型的检验	170
五、一元线性回归模型预测	171
第三节 多元线性回归分析	172
一、多元线性回归模型	172
二、多元线性回归模型的估计	173
三、多元线性回归模型的检验	174
四、多元线性回归模型预测	175
第四节 用 SPSS 作相关与回归分析	175
一、相关分析	175
二、绘制相关散点图	177
三、一元线性回归分析	178
四、多元线性回归分析	180
思考与练习	184
第九章 时间序列分析与预测	186
第一节 时间序列概述	186
一、时间序列的概念和作用	186
二、时间序列的种类	186
三、编制时间序列的原则	188
第二节 发展水平指标	188
一、发展水平和增减水平	189
二、平均发展水平和平均增减水平	190

第三节 发展速度指标	195
一、发展速度和增减速度	196
二、平均发展速度和平均增减速度	198
三、应用平均速度指标应注意的问题	200
第四节 时间序列的趋势分析与预测	200
一、时间序列变动因素的分解	201
二、长期趋势变动的测定	202
三、季节变动的测定	209
四、常用的季节预测模型	211
第五节 用 SPSS 作时间序列分析与预测	212
一、长期趋势分析	212
二、季节变动分析	214
思考与练习	219
第十章 统计指数	221
第一节 统计指数概述	221
一、统计指数的概念	221
二、统计指数的作用	221
三、统计指数的种类	222
第二节 综合指数	223
一、综合指数的概念	223
二、综合指数的编制	223
第三节 平均数指数	227
一、平均数指数的概念	227
二、平均数指数的编制	228
第四节 指数体系和因素分析	231
一、指数体系概述	231
二、因素分析法概述	233
三、总量指标的因素分析	234
四、平均指标的因素分析	236
第五节 指数序列	238
一、指数序列的概念和种类	238
二、数量指标指数序列	239
三、质量指标指数序列	239
四、定基指数和环比指数之间的联系	239
思考与练习	240
附录 常用统计分布表	242
参考文献	252

第一章 总 论

统计学是通过搜索、整理、分析、描述数据等手段,以达到推断所测对象的本质,甚至预测对象未来的一门综合性学科。目前,统计学中大量的定量分析方法,被广泛地应用在各门学科之中,从自然学科、社会科学到人文科学,甚至被用在工商业及政府的情报决策之中。本章将阐述统计和统计学的含义、统计学的发展简史、统计学的研究对象和研究方法、统计学的基本范畴及主要应用领域等内容。

第一节 统计学的产生和发展

一、统计与统计学

统计作为一种社会实践活动已有悠久的历史。在英语中“统计”一词与“国家”(state)一词来自同一词源。因此,可以说自从有了国家就有了统计实践活动。最初,统计只是一种为满足统治者管理国家的需要而搜集资料,弄清国家的人力、物力和财力,作为国家管理依据的计数活动。而在现代社会中,“统计”已不再是统治者的专利,从国家管理者到一般平民百姓,都已把“统计”以及由此而得到的各种结果——统计信息运用到国家管理、社会经济活动及日常生活中,成为参与上述各种活动的工具和依据。因此在今天,“统计”已被赋予多种含义,很难给出一个简单明确的定义。在不同场合,“统计”一词可以有不同的含义。它可以指统计数据的搜集活动,即统计工作;也可以指统计活动的结果,即统计数据或统计信息;还可以指分析统计数据的方法和技术,即统计学。

统计学(statistics)是随着人类社会的发展和社会管理的需要而发展起来的。随着社会管理的日趋复杂,仅用数字计量客观现象已不能满足社会的需要,人们试图对客观现象进行定量分析,与此相适应,研究数据计量和分析方法的统计学诞生了。目前,随着统计方法在社会科学、自然科学和工程技术等领域的广泛应用,统计学已发展为具有众多分支学科的大家族,因而人们也从不同的角度对统计学的概念给出许多大同小异的解释,略举如下:

- 统计学是关于如何搜集、整理和分析客观现象数量方面的一门方法论科学。
- 统计学是关于数据资料的搜集、整理、显示、分析和推测的科学,目的在于帮助人们做出更有效的决策。
- 统计学是对被观察数据进行整理、描述和分析,进而探索数据内在规律的学科。
- 统计学是面对不确定、不充分的信息,为人们提供避免做出错误决策信息的科学方法。
- 统计学是对令人困惑费解的问题做出数字设想的艺术。

.....

面对众多的解释,要给统计学下一个能被普遍接受的定义是十分困难的。在这里,综合各种解释,在本书中我们说:统计学是一门关于客观现象数据搜集、整理、归纳、分析的方法论学科,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。

二、统计学的产生

统计学的产生与统计实践活动是密不可分的,统计作为一种社会实践活动,已有四五千年的历史。据历史记载,在远古时代有“结绳记事”、“刻木记数”之说,其实这可以说是最原始的统计。在我国,夏王朝“平水土,分九州,计民数”,进行了我国最早的人口调查,分中国为九州,人口数大约 1355 万。在国外,埃及、希腊和古罗马的历史中也有类似记载。公元前 3000 年左右,埃及为了规划金字塔的建筑和建立大型农业灌溉系统,曾先后调查了全国的人口状况。由此可以看出统计工作有着数千年的历史。

随着人们在这数千年的统计活动中的不断探索,对统计规律的认识也逐渐加深,这就产生了统计学。一般认为,统计学产生于 17 世纪中期,至今已有 300 多年的历史。统计学的产生与形成是从几个不同领域开始的,主要源自于英国的政治算术、德国的国势学和概率论。并在此基础上产生和形成了“政治算术学派”、“国势学派”、“数理统计学派”和“社会统计学派”等。

政治算术产生于 17 世纪中期的英国,其代表人物有约翰·格朗特(John Graunt)和威廉·配第(William Petty)。约翰·格朗特于 1662 年出版了《对死亡表的自然观察和政治观察》一书。书中他通过大量观察,研究并发现了一系列人口统计的规律,例如男性多于女性,在伦敦其比例为 14:13,在地方其比例为 15:14;地方的出生率大于死亡率;一般疾病和事故的死亡率比较稳定,而传染病的死亡率波动较大等。与格朗特同时代的英国经济学家威廉·配第的《政治算术》(1676)和其对国民收入估算的方法,对经济学、统计学的发展具有重大意义。《政治算术》中用于探索社会和经济现象的数量规律方法,最早地体现了统计学的基本思想。

从 17 世纪中期起,在德国兴起的国势学派用记述的方法来描述和研究一国的人口、财政、军事地理、政治和法律制度等。到 18 世纪,哥廷根大学教授阿亨瓦尔(Gottfried Achenwall)首次提出了“统计学”[Statistik, (英) Statistics]这一名词,并将其定义为关于“国家显著事项”的学问,通过这门科学可以清楚地了解国家的运行情况,这与当今的《官方统计》的内容基本一致。国势学派最初的“统计”只用文字记载,在后来发展中使用了数字和图表。

统计学的另一个重要源头是以概率论的研究为开端,并以其为基础形成和发展起来的以随机现象为主要研究对象的数理统计学。

概率论的产生起源于对赌博中掷骰子输赢问题的研究,其目的是找出一套办法,合理算出取胜的机会,早在但丁(Alighieri Dante)的《神曲》中就记载了投掷三个骰子所可能出现的结果,后来到 16 世纪中期,意大利数学家卡达罗(Kirolama Cardano)写出《论赌博》一书,为运用数学理论研究概率开辟了道路。几乎在同一时期,意大利天文学家伽利略(Galilei Galileo)也进行了相同的研究,提出了概率论的基本原理。但概率论的真正历史始于 17 世纪中期,法国数学家帕斯卡尔(Blaise Pascal)和费尔马特(Pierre de Fermat)通过通信方式,将赌博中出现的具体问题归纳为一般的概率原理,为以后的概率论和统计学的发展奠定了基础。直到 18 世纪初,瑞士数学家贝努里(Jakob Bernoulli)的《推算法》问世,概率论作为一门理论才初步形成。在此之后,概率的乘法运算法则、正态分布、先验概率等理论相继问世,使概率论得以快速发展起来。

概率论的发展和在社会现象研究中的应用促进了具有科学推论特色的近代统计学的形成。到 19 世纪,用概率论研究社会各种现象的人日益增多。在这方面对近代统计学的形成做

出重大贡献的有法国数学家拉普拉斯(Pierre Simon Laplace)、德国数学家高斯(Carl Friedrich Gauss)和比利时统计学家凯特勒(Adolphe Quetelet)。特别是凯特勒,他以大量观察和综合平均的方法对社会现象进行研究,从而将概率论(正态分布等)、国势学和政治算术中用于观察群体现象并进行数量分析的方法融合为统计学。

三、统计学的发展

从统计学的发展历史可以看出,统计学的发展过程基本上沿着两条主线展开的:一是以政治算术、国势学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统计学和经济统计学;二是以概率的研究为开端,并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学。在很多时期,上述两个方面又交织在一起,相互影响并促进了统计学的发展。从统计学发展的时间和内容上来划分,大致可以分为三个阶段。

(一) 统计学的产生和形成阶段

统计学的产生和形成阶段从 17 世纪中期至 19 世纪末,历时 200 余年。本阶段统计学以政治算术,国势学为开端,沿着由格朗特开创的人口统计、配第开创的经济统计和国势学的记述统计,针对不同的研究对象出现了诸如人口统计、经济统计、犯罪统计、道德统计、社会统计等各方面内容的“社会统计学”。与之相适应的社会调查与社会研究方法也取得了较大的发展。人们通过社会调查、搜集、整理、分析资料,揭示社会现象和问题,并提出相应解决问题的具体办法。这些方法成为当今我们称为描述统计学的基本内容。另外,由于概率论的基本内容在这段时期逐渐形成,借助于概率方法,统计研究的基本问题成为如何用一套通用的数学公式来描述研究现象的整体。

(二) 统计推断方法和理论体系确立的阶段

从 20 世纪初英国统计学家古塞特(William Gosset)建立小样本(t 分布)理论,到 40 年代末美国瓦尔德(Abraham Wald)的统计决策与序贯抽样的提出。估计理论、样本分布理论、方差分析、实验设计、非参数估计、时间序列、随机过程等方面的研究获得重大进展。统计推断方法和理论的快速发展,使统计学进入到现代统计学阶段。

(三) 全面发展的阶段

从 20 世纪 50 年代起,统计理论、方法和应用进入了一个全面发展的阶段。在本阶段统计学受计算机、信息论等现代科学技术的影响,新的研究领域不断出现,如多元统计分析、现代时间序列分析、贝叶斯统计、非参数统计、线性统计模型等。据美国学者统计,现代统计学以指数式加速发展,新的研究分支不断增加,统计应用的领域不断扩展,几乎所有的科学研究都要用到统计学。统计学在各学科领域的应用又进一步促进了统计方法研究的深入和发展。可以说,在现代的科学研究中,统计学已成为一门基础性的方法论学科。

进入 21 世纪,随着科学技术的持续发展以及对统计方法和技术的依赖性越来越强,统计方法与技术应用将越来越重要。从当前统计学理论和应用研究的发展趋势看,统计学理论方面的研究将会更依赖于现代数学理论,统计学应用研究的分支越来越多,一些过去与统计学毫无关系的学科,如历史学、心理学等都已开始利用统计方法来进行研究分析。这些表明,

统计学的发展前景十分宽广,统计学科除了自身的独立性发展外,也将越来越具有交叉学科的性质。

四、统计学和其他学科的关系

统计学是一门应用性很强的基础学科,必然与其他学科有密切的联系。由于其他学科在研究中都要研究和分析数据,而统计方法可以帮助各学科探索和揭示其内在的数量规律,而对数量规律性的解释则由各学科的研究来完成。由此可以看出统计学不能替代其他学科,只能为其他学科提供一种有用的定量分析工具。

(一) 统计学与马克思主义哲学的关系

统计学作为一门社会科学,它是在现象的质与量的密切联系中,来研究客观现象总体的数量特征和数量关系,以及通过这些数量方面来反映客观现象发展变化的规律性。它以马克思主义哲学作为理论和方法论基础。马克思主义哲学是关于自然、社会和思维发展的最一般规律的科学,它既是正确的世界观,又是科学的方法论,对于统计学的理论和方法具有普遍的指导意义。比如,哲学中关于人类认识事物的规律总是由认识个别的特殊事物,逐步扩大到认识一般的事物,即由具体到抽象,由特殊到一般,当人们已经认识了事物的共同本质之后再用它指导具体事物,统计学中的大量观察法正是运用了这种思想的统计分析方法;统计学中的大数定律的意义在于使统计总体不受偶然因素的影响而表现出集体的规律性,这正是哲学中的偶然性与必然性、个体与一般的对立统一规律在数量关系上的反映;哲学中强调人们认识事物、分析事物必须运用科学的方法从各个层次、多个角度去全面深入地把握,统计学中涉及众多的统计分析指标,形成庞大的指标体系,指导人们全面分析问题;哲学中认为事物是发展变化的,这种变化是有规律可循的,统计学中的动态数列反映现象变化的全过程,为人们研究分析事物的发展变化规律提供依据,这正是对这一辩证思想的反映;哲学中认为任何事物和现象都不可能孤立存在和发展,而统计学中的相关分析正是运用这种思想从定性和定量两个角度帮助人们判断事物之间的联系方式和变化规律。

(二) 统计学与经济学的关系

统计学以经济学所阐述的经济关系和经济规律作为理论指导。统计学中所涉及的概念、分组和计算范围都必须以经济学所确定的经济范畴和理论为依据,例如,总产值、净产值、工资、利润、劳动生产率等具体统计指标,都反映一定的经济范畴,必须以经济学的定义为基础。另外,统计分析也必须根据经济学所研究的经济规律来确定现象之间的本质联系,才能进一步分析现象之间的变动数量关系,作出科学的判断,例如,要分析原材料支出总额的变化情况,必须先分析产量、产品的原材料单耗和原材料单价变动情况,因为原材料支出总额是产量、产品的原材料单耗和原材料单价乘积的结果。

(三) 统计学与数学的关系

统计学与数学有密切的关系,在现代统计学理论和方法的研究中用到了比较多的数学方法,从事统计理论研究和统计应用的人也须具备相应的数学知识,但不能由此认为统计学就是数学的一个分支,实际上,数学只是为统计理论和统计方法的发展提供了数学基础,统计学的主要特征是研究表现客观现象的数据,而不是数学。另一方面,统计方法与数学方法一样,并不能独立地直接研究和探索客观现象的规律,而是给其他学科提供一种研究和探索客

观规律的数量方法,但统计学与数学是有本质区别的。尽管从表面上看两学科都是研究数量规律,都在与数字打交道,但实际上有差别。首先,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;数学研究的是无量纲和单位的抽象的数,而统计学研究的是有具体实物或计量单位的数据。其次,两者在研究中所使用的逻辑方法也不相同,数学研究使用纯粹的演绎,而统计学研究则是演绎与归纳相结合。另外,数学研究可以与实际无关,而统计学研究则需要深入实际搜集数据,并与具体问题相结合,经过科学的演绎和归纳后才可能得出有益的结论。

(四) 统计学与计算机科学的关系

统计数据处理作为统计学的主要运作手段,历经了手工、机械、机电、电子等循序渐进的阶段,数据处理手段的每一次发展变迁,都给统计实践带来革命性的飞跃。随着信息化、网络化的发展,统计学本身也发生了巨大的变化,主要体现在两个方面:一是“多”,即数据量内容大大增加;二是“快”,即要求处理数据的速度又快又好。由于此时计算工作量大,使得传统统计分析方法在一开始并没能够在实践中很好地推广开来。而计算机科学的诞生与发展,使得复杂的数据处理工作变得非常容易,那些计算繁杂的统计方法的推广与应用,由于相应统计软件的开发与商品化而变得更加方便与迅速,非统计专业的理论工作者可以直接凭借商品化统计分析软件来处理各类现实问题的多变量数据分析,而无须对有关统计方法的复杂理论背景进行研究。计算机运行能力的提高,使得大规模统计调查数据的处理更加准确、充分与快捷。

第二节 统计学的研究对象

由统计学的概念可知,统计学的研究对象是具有某种相同属性的群体现象,以及探索群体现象数量表现的内在规律性及对群体现象进行计量描述和分析推断的方法。它包含了统计学的三个核心要点。

一、研究对象是群体现象

统计学有三个最基本的概念:总体、样本和个体。总体是指研究对象全体的集合;样本是指来自总体中的部分对象的集合,个体是总体中的任何一个元素(称为总体单位)。统计学的研究对象是这些集合而不是组成这些集合的一个个元素。例如,如果要知道某一物体的重量,只要把它称一称即可得知,无须进行统计分析。然而,若要知道一批物体的总重量与平均重量,一群人的最高者、最低者和这群人的平均身高,就需要测量汇总,并采用一定的统计方法进行平均。若期望通过其中的小部分,即样本的信息来了解总体,则需要抽样、估计检验、统计分析推论的一套过程,而这些正是统计学的内容。尽管总体或样本的信息都表现在一个个元素或个体上,研究总体离不开个体或部分个体,但统计学研究的不是个体现象,而是通过个体或部分个体所载有的信息来研究、认识和说明群体现象。

二、探索群体现象数量表现的内在规律性

统计探索的是群体现象数量表现的内在规律性,这里有两个要点:数量性和规律性。

（一）数量性

统计学是一门有关统计数据的科学，统计学与统计数据有着密不可分的关系。英文中，“Statistics”一词有两个含义：当其以单数名词出现时，表示作为一门学科的“统计学”；当其以复数名词出现时，则表示“统计数据”或“统计资料”。统计学的这一名词特性，直观地说明了统计学与统计数据之间密不可分的关系。统计学由一套搜集和处理统计数据的方法所组成，这些方法源自于对统计数据的研究，目的也在于对统计数据进行研究。所以，离开了统计数据，统计方法也就失去了用武之地，统计学也就失去了其存在的意义。而统计数据如果不用统计方法进行分析研究，也就仅仅是一堆数字而已，得不出任何有益的结论。

此外，统计数据在英文中是以复数形式出现的，这就表明统计数据不是单个的数字，而是由多个数据构成的数据集。对单个数据既用不着统计方法进行分析，也不可能探索事物的规律性，只有对同一特性事物进行反复多次的观察或计量并获得大量的数据，才有可能利用统计方法探索其内在的规律性。

可见，统计学是用数字来说明群体现象特征的，作为特征，既可以是其数量方面的特征，如人的身高、年龄等；也可以是属性方面的特征，如人的性别、民族等。统计学所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，则表现为不同性别的人口在总人口中的数量、比例；民族特征则表现为各民族人口在总人口中的数量、比例等。

（二）规律性

统计学研究的是各个体之间在数量上存在差异的群体现象。如果一批物体中的每件重量都一样，一群人中的每个人身高都相同，也是用不着进行统计研究的。然而，现实中群体现象总是由多数量特征各异的个体所组成。而这些有差别的个体数量特征中却存在着某一数量规律性。统计学提供了探索数据内在规律性的一套方法。什么是统计数据的内在数量规律性？为什么统计方法能够通过对数据的分析找出其内在的数量规律性？下面的例子可以说明这些问题。

【例 1-1】 在新生婴儿的性别问题上，如果就单独的一个家庭来观察，每个家庭的新生婴儿的性别可能是男性，也可能是女性。如果不对出生人口进行任何限制，有的家庭几个孩子都有可能是男性，而有的家庭几个孩子都有可能是女性。从个别看，新生婴儿的性别比例似乎没有什么规律可循。但是，如果对大量家庭的新生婴儿进行观察，就会发现新生儿中男孩略多于女孩，大致比例为 107: 100。也就是说每出生 100 个女婴，相应地就有 107 个男婴出生。男、女婴的这个性别比例就是新生儿性别比的数量规律。古今中外这一比例大致相同，这是由人类自然发展的内在规律所决定的。人类社会要发展，就要保持男女人数大致相等。尽管以新生婴儿来看，男婴略多于女婴，似乎并不平衡，但由于男婴的死亡率高于女婴，到了中年时，男女人数就大体上相等了。进入中老年后，男性的死亡率仍然高于女性，导致男性的平均寿命低于女性（约 4~6 岁），老年男性反而少于女性。生育人口在性别上保持大体平衡，保证了人类社会的进化和发展。对人口性别比例的研究是统计学的起源之一，也是统计方法探索的数量规律性之一。