普通高等教育“十三五”规划教材

应用统计学

王生喜 主编

科学出版社

普通高等教育“十三五”规划教材

应用统计学

主　编　王生喜
副主编　吴丹丹　麻　露

科学出版社
北京

内 容 简 介

本书系统介绍应用统计学的基本理论与基本方法，涵盖描述统计、概率论基础以及推断统计三个经典模块. 内容安排遵从分层递进、由浅入深的原则，章节段落力求条理清晰，文字表达力求简洁明快. 在理论与方法的阐述上，既有严谨规范的专业术语，又有通俗易懂的直观描述，借以帮助不同层次、不同专业背景的读者提高学习效率.

本书以较大的篇幅介绍了学习推断统计所必需的概率论基础知识，以帮助没有系统学习过概率论课程的读者有效掌握推断统计的方法.

本书可作为高等院校经济类、管理类各专业的本科生的应用统计学教科书.

图书在版编目(CIP)数据

应用统计学/王生喜主编. —北京: 科学出版社, 2018.8
普通高等教育“十三五”规划教材
ISBN 978-7-03-057544-9

I. ①应… II. ①王… 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(2018)第 112358 号

责任编辑: 任俊红　李香叶 / 责任校对: 彭珍珍
责任印制: 霍　兵 / 封面设计: 华路天然工作室

科学出版社 出版
北京东黄城根北街 16 号
邮政编码: 100717
http: //www.sciencep.com

三河市宏图印务有限公司　印刷
科学出版社发行　各地新华书店经销
*
2018 年 8 月第　一　版　开本: 787×1092　1/16
2018 年 8 月第一次印刷　印张: 20
字数: 450 000

定价: 49.00 元

（如有印装质量问题，我社负责调换）

前　言

本书是为高等院校经济学类、管理学类各专业的本科生编写的应用统计学教科书.

改革开放以来, 我国高校普遍为经济管理类各专业的本科生开设了统计学课程. 在多年的统计学教学实践中, 曾经涌现出一批具有时代特色的优秀教材, 这些教材对培养合格的财经管理人才发挥了重要作用. 随着我国高等教育的发展变革, 统计学的教学模式、教学方法以及教学内容体系诸方面都在发生着深刻的变化, 改革和创新成为引领统计学课程发展的主线. 如何为经济管理类的本科生编写一本既体现改革精神又符合读者期望, 既传授统计思想又突出实际应用, 既介绍经典理论又贯穿现代理念的统计学新教材, 是本书作者夙愿.

基于上述构想, 作者深入研究了国内外统计学教育教学改革的成果, 借鉴了众多优秀教材的内容结构和处理方法, 并结合作者长期从事统计学的教学实践, 编写了本书.

本书系统介绍应用统计学的基本理论与基本方法, 涵盖描述统计、概率论基础以及推断统计三个经典模块. 具体内容包括数据的搜集、数据的整理与显示、数据分布特征的度量、随机事件及其概率、随机变量及其分布、抽样分布与大样本理论、参数估计、假设检验、方差分析、一元回归分析、多元回归分析、时间序列分析以及统计指数.

本书的特点主要体现在以下三个方面.

一、三位一体有机融合. 在对描述统计、概率论基础、推断统计三个模块之间的衔接及融合方面进行了新的尝试. 突出描述统计思想对推断统计方法的引领作用, 以统计思想及语言阐述概率论的一些基本概念及模型. 例如, 以频率直方图为桥梁, 用统计观点刻画和描述概率密度函数与分布函数; 强化大样本理论, 基于统计思想阐述大数定律与中心极限定理的意义和作用, 并由此导出大样本下样本比例的渐近分布, 等等.

二、严谨与直观互补. 内容、结构与体系安排遵从分层递进、由浅入深的原则, 章节段落力求条理清晰, 文字表达力求简洁明快. 在理论与方法的阐述上, 既有严谨规范的专业术语, 又有通俗易懂的直观描述, 借以帮助不同层次、不同专业背景的读者提高学习效率.

三、方法与应用并重. 在对较复杂问题的处理上, 着力引导读者理清问题的来龙去脉, 力求做到“知其然且知其所以然”. 本书以较大的篇幅介绍了学习推断统计所必需的概率论基础知识, 以帮助没有系统学习过概率论课程的读者有效掌握推断统计的方法. 对于学习过概率论课程的读者, 运用统计观点重温这部分内容, 也能使他们更好地理解并应用推断统计方法. 书中配备了大量联系经济管理实际的、富有启发性的例题、案例及课后练习, 借以训练读者的统计思维方式、综合应用能力以及开拓创新能力.

本书采用 Office 2010 版中的 Excel 为数据分析的基本软件. Excel 2010 版提供了专门的数据分析工具、丰富的统计函数、易于掌握的编辑方式和强大的图表绘制功能, 能够快速完成各类数据的整理、展示以及对统计模型的计算分析任务, 是非统计专业的读者学习应用统计学的理想软件之一.

本书尝试采用移动式课外学习模式，每章末尾嵌入了二维码课外学习平台，以方便读者利用智能手机的扫码功能获取本书的学习资料，借以巩固学习成果，开拓知识视野.

本书第 1、5~7、10~12 章由王生喜编写，第 2~4、9 章由吴丹丹编写，第 8、13、14 章由麻露编写. 每章末尾的二维码课外学习平台由吴丹丹负责设计嵌入. 全书由王生喜负责总纂、整理和定稿.

刘颖参与了本书初稿的讨论，提出了有益的修改建议，提供了第 10、11 章的部分资料及若干例题.

在本书的策划及编写过程中，朱智强先生、苏锦俊博士、卢晨博士给予了作者热情的鼓励、支持和帮助.

作者由衷感谢关心本书的所有亲朋好友和每一位读者朋友. 由于作者水平所限，书中难免有疏漏之处，欢迎读者批评指正.

作　者

2018 年 6 月

此为试读，需要完整PDF请访问：www.ertongbook.com

目　录

第 1 章　绪论 …… 1
1.1　统计学的概念与性质 …… 2
1.2　统计学的产生与发展 …… 3
1.3　统计学的基本方法与应用 …… 5
1.4　统计学的基本术语 …… 7
1.5　变量与数据的分类 …… 10
1.6　统计软件简介 …… 12
本章小结 …… 13
思考与练习 …… 13
第 2 章　数据的搜集 …… 15
2.1　数据的来源 …… 15
2.2　调查方案与组织形式 …… 16
2.3　随机抽样与非随机抽样 …… 18
2.4　数据的搜集方法 …… 22
2.5　调查问卷编制 …… 23
2.6　数据的误差与控制 …… 26
2.7　Excel 操作 …… 27
本章小结 …… 28
思考与练习 …… 28
第 3 章　数据的整理与显示 …… 30
3.1　数据的审核与预处理 …… 30
3.2　数据的分组 …… 32
3.3　数据的显示 …… 38
3.4　Excel 操作 …… 44
本章小结 …… 49
思考与练习 …… 50
第 4 章　数据分布特征的度量 …… 52
4.1　集中趋势的度量 …… 52
4.2　离散程度的度量 …… 59
4.3　数据分布几何特征的度量 …… 67
4.4　数据分布特征的综合分析 …… 68
4.5　Excel 操作 …… 70
本章小结 …… 73
思考与练习 …… 74

第 5 章　随机事件及其概率 …… 77
5.1　随机事件 …… 78
5.2　随机事件的概率 …… 83
5.3　古典概型 …… 85
5.4　条件概率 …… 88
5.5　事件的独立性 …… 94
本章小结 …… 98
思考与练习 …… 99
第 6 章　随机变量及其分布 …… 101
6.1　随机变量 …… 101
6.2　离散型随机变量 …… 103
6.3　随机变量的分布函数 …… 109
6.4　连续型随机变量 …… 112
6.5　随机变量的数字特征 …… 122
6.6　Excel 操作 …… 132
本章小结 …… 137
思考与练习 …… 138
第 7 章　抽样分布与大样本理论 …… 140
7.1　总体与样本 …… 141
7.2　样本统计量 …… 142
7.3　抽样分布 …… 143
7.4　大样本理论 …… 149
本章小结 …… 155
思考与练习 …… 155
第 8 章　参数估计 …… 157
8.1　点估计 …… 158
8.2　区间估计 …… 165
8.3　一个总体参数的区间估计 …… 167
8.4　两个总体参数的区间估计 …… 174
本章小结 …… 180
思考与练习 …… 181
第 9 章　假设检验 …… 183
9.1　假设检验思想与表述 …… 184
9.2　检验形式与检验过程 …… 188
9.3　一个总体参数的假设检验 …… 192
9.4　两个总体参数的假设检验 …… 197
9.5　Excel 操作 …… 203
本章小结 …… 206

思考与练习 …… 207
第 10 章　方差分析 …… 209
10.1　基本概念和术语 …… 209
10.2　单因素方差分析 …… 212
10.3　无交互作用的双因素方差分析 …… 214
10.4　存在交互作用的双因素方差分析 …… 216
10.5　Excel 操作 …… 218
本章小结 …… 220
思考与练习 …… 221
第 11 章　一元回归分析 …… 223
11.1　线性相关分析 …… 223
11.2　一元线性回归模型 …… 228
11.3　回归直线的拟合优度 …… 232
11.4　一元回归模型的统计检验 …… 236
11.5　回归统计报告解析及模型评价 …… 239
11.6　估计与预测 …… 242
11.7　非线性回归模型的线性化 …… 243
11.8　Excel 操作 …… 246
本章小结 …… 247
思考与练习 …… 248
第 12 章　多元回归分析 …… 250
12.1　多元线性回归模型 …… 250
12.2　多元回归模型的统计检验 …… 252
12.3　统计报告解析及模型评价 …… 255
12.4　Excel 操作 …… 257
本章小结 …… 258
思考与练习 …… 258
第 13 章　时间序列分析 …… 260
13.1　时间序列的描述性分析 …… 261
13.2　时间序列的构成因素 …… 266
13.3　时间序列的预测 …… 279
本章小结 …… 282
思考与练习 …… 283
第 14 章　统计指数 …… 284
14.1　统计指数概述 …… 285
14.2　总指数的编制 …… 287
14.3　指数体系与因素分析 …… 292
14.4　几种常见的经济指数 …… 294

本章小结 …… 296
思考与练习 …… 297
参考文献 …… 299
附录　常用统计表 …… 300
附表 1　标准正态分布表 …… 300
附表 2　t 分布临界值表 …… 302
附表 3　χ^2 分布临界值表 …… 304
附表 4　F 分布临界值表 …… 305
附表 5　泊松分布概率值表 …… 309

第1章 绪 论

《无处不在的统计》

《无处不在的统计》是中国统计学会于2011年组织出版的统计科普丛书. 该书既有从不同角度解读GDP、CPI、PPI、ICP等宏观经济指标的文章, 也有从实际生活小事中引发的对概率问题、抽样问题的探讨, 更有对统计实际工作的全面展现, 为读者打开一扇了解统计、发现统计的窗口. 统计科普丛书涉及诸如“透过指标看经济”“大数据时代的政府统计”“统计改变人生”“统计学是一门艺术”“你有多幸福——漫谈幸福、幸福感及幸福指数”等各个方面的话题. 通过阅读这类书籍我们能够深切地感受到, 从关乎国计民生的大事, 到家庭生活的一些细节, 都会不同程度地涉及统计数字、统计思想和统计方法.

下面的案例印证了“统计无处不在”的断言.

总裁与出租车司机的聊天记录

2014年的一天, 某跨国公司驻中国分支机构的总裁从上海徐家汇搭乘出租车赶往浦东机场. 这位总裁和健谈的司机路上聊过天. 到达机场后, 总裁留给司机一张名片, 说: “你有没有兴趣这个星期五到我办公室, 给我的员工们讲一讲你是怎么开出租车的? 你就当打着里程表, 60公里一小时, 你讲多久, 我就付你多少钱. ”总裁在飞行途中记录了这次聊天的内容并将其发到个人微博, 吸引了众多的读者. 以下是聊天记录中出租车司机的几段精彩对话:

“开出租车也要用科学的方法, 要懂得统计. ……我每天要交380元的份子钱(管理费), 油费大约210元. 我每天开17个小时的车, 平均下来, 每小时成本为34.70元, 其中固定成本(份子钱)22.35元, 油费12.35元.”

“成本是不能按公里数计算的, 只能按时间计算. 计价器有一个检查功能, 可以看到每天的详细记录. 我做过数据分析, 每两次载客之间的平均空驶时间为7分钟. 上来一个客人起步价10元, 大约要走10分钟, 也就是每一个10元的客人要花17分钟的成本, 就是9.83元, 不赚钱啊! ”

“有一次一位客人打车去火车站, 我问怎么走. 他说了一条路线, 我说慢, 于是我说了一条路线, 他又说绕远了. 我告诉他, 按他的路线打表要50元, 按我的走法, 等里程表显示超过50元了, 我就翻表, 他只给50元就好了, 多出来算我的. 最后按我的路线多走了4公里, 却快了25分钟, 我只收了50元, 乘客很高兴, 省了10元左右. 这4公里对我来说就是1元多的油钱, 我相当于用1元多买了25分钟. 我刚才说了, 我一小时的成本是34.70元, 你看我多合算啊!”

上述案例也印证了著名学者威尔斯(H.G.Wells)的名言: “统计思维总有一天会像读与写一样成为一个有效率公民的必备能力.”

本章简要介绍统计学的概念与性质、统计学的产生与发展、统计学的基本方法与应用，以及统计学的一些基本术语.

1.1　统计学的概念与性质

1.1.1　统计与统计学

“统计”一词源于德语“statistik”. 在不同的场合，“统计”具有三种不同的含义，即统计活动、统计资料和统计学.

1. 统计活动

统计活动又称为统计工作，是指收集、整理和分析统计数据，并探索数据的内在数量规律性的过程.

2. 统计资料

统计资料又称为统计数据，即在统计活动过程中获得的所有资料的总称. 统计资料表现为各种反映自然现象和社会经济现象数量特征的试验数据、调查数据、统计图表、数据分析报告、政府统计公报、统计年鉴等各种数字、文字等资料. 广义的统计资料还包括可数字化的数据，包括声音、图像、视频等音像资料.

3. 统计学

统计学是在统计实践活动的基础上产生并逐步发展起来的一门科学. 在统计学的发展和演变过程中，一些研究者对统计学给出了多种不同版本的定义. 韦伯斯特国际词典第 3 版对统计学概念作了如下简明的阐述：“统计学是一门收集、分析、解释和提供数据的科学.”换言之，统计学研究如何测定、收集、整理、归纳和分析反映客观现象总体数量的数据，以帮助人们正确认识客观世界的数量规律，是一门方法论科学.

统计工作、统计资料和统计学相互依存，相互联系，共同构成了完整的“统计”体系. 统计资料是统计工作的成果，统计学是对统计工作实践过程的理论概括和科学总结，统计学的发展和研究成果是支持和推动统计工作创新的原动力.

1.1.2　统计学的性质和特点

如上所述，统计学是研究客观现象总体数量方面的方法论科学. 作为方法论科学的统计学，不仅适用于对自然现象数量方面的研究和分析，同样适用于对社会经济现象数量方面的研究和分析.

在统计实践活动中，人们不仅要面对确定性现象的数量问题，还要面对不确定现象(即随机现象)的数量问题. 随机现象的特点是：人们虽然知道这种现象过去的状况，但其未来的发展不能完全确定. 例如，某种新产品投放市场后受消费者欢迎的程度、对某

此为试读，需要完整PDF请访问：www.ertongbook.com

个房地产项目投资后的获利情况、股票的未来价格等，这些都是不确定的随机现象. 所谓不确定，只是对一次或少数几次观察或试验而言. 当我们在相同条件下进行大量观察时，随机现象就会呈现出某种规律，即随机现象具有统计规律性. 揭示随机现象的统计规律是统计学研究的核心问题.

从对统计学研究对象的考察分析，可对统计学的主要性质和特点概括如下.

1. 数量性

数字是统计的语言，统计数据是客观事物量的反映. 统计学的主要任务是处理和分析数据，所以数量性是统计学研究对象的基本性质和特点之一.

2. 群体性

群体性也称为总体性，即研究对象应包含不止一个个体(元素)，单个个体不需要做统计分析. 统计学的数量研究是对客观现象的群体中普遍存在的事实进行观察和分析，从而得出反映现象总体的数量特征. 例如，要了解某地区贫困人口的生活现状，就不能局限于对个别贫困户的访问，要着眼于该地区贫困人口这个群体，继而对整个地区居民的人均收入、贫困人口的比例以及贫富差距等方面进行系统的考察.

3. 变异性

变异性也称为随机性，即随机现象的不确定性及群体中个体之间的差异性. 一般情况下，统计学研究面对的群体(总体)各单位的特征表现存在差异，这些差异往往并不是事先可以预知的，例如股票的价格和成交量每天都不相同. 面对这样的随机现象，如何通过存在差异的数据揭示出总体数据的统计分布规律，归纳和概括出其共同特征，就成为统计学研究的重要任务.

1.2 统计学的产生与发展

1.2.1 统计学的产生

统计学形成于19世纪中叶，它的产生主要源于国势学、政治算术和概率论三个学术流派.

国势学产生于17~18世纪的德国. 从17世纪中叶起，在德国兴起的国势学派用记述的方法研究一国的地理、人口、财政、军事、政治、法律制度等方面的问题. 18世纪中期，哥廷根大学教授戈特弗里德·阿亨瓦尔(G.Achenwall)首次提出了“统计学”这一名词，并将其定义为国家显著事项的学问. 最初的“统计”主要是文字记载，后来才涵盖了数字和图表.

政治算术学发源于英国伦敦，产生于17世纪中叶. 英国是较早利用数字对人口、经济等方面进行记载和推断的国家. 约翰·格朗特(J. Graunt)于1661年首次发表了对人口

出生率研究的结果，并观察到一切疾病和事故在全部死亡原因中占有稳定的百分比. 与格朗特同时代的英国经济学家威廉·配第(W. Pett)在其著作《政治算术》中，运用观察和数量分析等方法对社会经济现象进行研究，为统计学的发展开辟了广阔的前景. 威廉·配第等学者引入的国民收入估算方法，对经济学和统计学的发展具有重大的意义. 这一派学者没有使用“统计学”这一术语，而是用“政治算术”来表明其研究的特色，因而被称为政治算术学派.

概率论起源于对赌博中输赢问题的研究. 1714 年，瑞士数学家雅各布·伯努利(J.Bernoulli)的《推测术》问世，标志着概率论作为一种理论初具雏形. 之后，概率的古典定义、乘法法则、正态分布、先验概率等理论相继问世，概率论逐渐发展成一门独具特色的数学分支.

将上述三方面成果融为一体而创立了统计学的学者是比利时统计学家阿道夫· 凯特莱 (A. Quetelet). 他首次研究了天文与气象方面的统计资料，其后又利用统计数字研究植物界和人类社会. 他以概率论作为理论基础，用大量观察和综合平均的方法来研究众多现实问题，进而将概率论、国势学、政治算术融合为具有现代意义的数理统计学.

1.2.2　统计学的发展

统计学的发展大致划分为如下三个阶段.

第一阶段：统计学的初创阶段. 从17世纪中叶开始到19世纪末，经历了200多年的漫长时间，逐渐形成了描述统计和概率论的基本体系.

第二阶段：统计推断方法体系基本确定的阶段. 从20世纪初“小样本 t 分布” 方法的问世，到40年代末“统计决策函数”与“序贯抽样”的提出，经历了近30年的时间. 在这个阶段，估计理论、大样本理论、方差分析、非参数方法、时间序列分析、随机过程等方面的研究都获得了重大进展，推断统计理论迅速发展，概率论体系日臻完善.

第三阶段：统计方法与应用研究全面发展的阶段. 从20世纪50年代起，统计学在计算机科学和信息论等现代科学技术的影响下，新的研究领域层出不穷，产生了多元统计分析、现代时间序列方法、人工智能等众多统计学新分支. 在这个阶段，统计应用领域的不断拓展进一步促进了统计方法的创新.

从理论研究的发展趋势看，统计学日益凸显出交叉学科的性质，统计学将会不断出现新的分支. 在理论与方法的创新上，统计学对现代数学的需求日益增大. 在数据处理的手段上，统计学对计算机和数字网络技术的依赖日益增强.

从方法应用与技术推广的未来趋势看，统计学应用的范围将日益广泛，研究的对象将日益繁多. 政治学、历史学、法学、文学等“非数据”人文学科，也已经在逐步地应用统计方法进行研究和分析所面临的问题. 数据挖掘技术以及大数据分析技术的应用前景也很广阔. 与经典统计学并驾齐驱的贝叶斯统计学在统计推断、经济预测、管理决策等领域将继续显示其强大的生命力，必将成为统计实践中的有力工具.

1.3 统计学的基本方法与应用

1.3.1 统计工作的基本环节

统计工作的全过程包括以下四个基本环节.

1. 统计设计

根据所要研究问题的性质, 基于相关学科理论的指导, 制定统计指标、指标体系和统计分类, 给出统一的概念和标准, 提出收集、整理和分析数据的方案与工作计划等. 统计设计是整个统计研究的前期工程. 做好统计设计不仅需要统计学一般理论和方法的指导, 而且要求设计者具备相关的学科知识, 对所要研究的问题本身具有深刻的认识. 例如, 要设计一套能够较好地评价企业经营状况的统计体系与方案, 设计者不仅要掌握一些统计方法, 还必须具备企业经营管理的经验和一定的理论素养.

2. 数据收集

经过统计设计, 形成方案之后, 就可以进入收集统计数据的程序. 统计数据的收集有两种基本方法: 实验方法和社会调查方法. 对于大多数自然科学和工程技术的研究, 一般通过有控制的科学实验获得数据, 即可以采用实验法. 对于社会经济现象的研究, 一般无法进行重复试验, 要取得相关数据, 就必须在总体中选取足够多的单位进行调查或观察. 如何科学地进行调查是统计学研究的重要内容.

3. 数据整理与分析

统计数据的收集完成之后, 还必须对原始数据进行整理、加工和分析.

数据整理是指根据统计研究的任务与要求, 对搜集来的各种原始资料进行科学的分类和汇总, 为统计分析提供系统化、条理化的综合资料的工作过程. 统计资料整理在统计工作中处于中间阶段, 它是统计调查的继续, 是统计分析的前提, 起着承前启后的作用.

数据分析是指应用适当的统计方法对经过整理的数据进行系统的研究, 从中获得有价值的信息或形成某些结论的研究过程. 统计分析能够有效地将数据、情况、问题、建议等融为一体, 是发挥统计的信息、咨询、管理、监督和决策功能的重要内容.

4. 统计资料的积累、开发与应用

统计数据经过整理和分析, 就形成了相关的统计资料. 统计研究的目的在于认识客观现象的数量规律, 所以统计资料的提供并不是统计研究的终结. 对于已经公开的统计资料, 可以进一步加工, 结合相关的实质性学科的理论知识进行分析和利用. 如何更好地将统计资料和统计方法应用于各个研究领域, 也是统计学研究的一个重要方面.

1.3.2　统计学的基本方法

1. 描述统计方法

描述统计也称统计描述，是指对采集的数据进行登记、审核、整理、归类及图表显示，在此基础上计算出各种能反映研究对象数量特征的综合指标，经过归纳分析得到各种有用的统计信息. 通过统计描述，可以对所研究对象的数据分布规律有一个基本的认识和直观的了解.

统计描述是统计研究的基础，它为统计推断、统计咨询及统计决策提供必要的依据. 统计描述也是对客观事物认识不断深化的过程，通过对分散无序的原始数据的整理归纳，揭示客观事物内在的数量规律性，以达到认识事物本质的目的.

2. 推断统计方法

推断统计方法也称归纳统计学，是研究根据样本数据推断总体特征的方法.

20 世纪以来，随着科技的进步和社会的发展，社会经济领域和自然科学技术领域都对统计学提出了更高的要求. 仅靠描述统计方法既不能解决现象间复杂的数量关系，也不能对未来做出估计和预测. 在这样的背景下，推断统计方法应运而生. 自 20 世纪下半叶，统计学的推断理论和方法逐渐成为现代统计的核心内容.

推断统计学方法的核心是“利用样本信息推断总体”. 该方法主要包括估计和检验两个方面，即利用获得的样本数据，对总体分布或总体参数进行科学估计，进而进行统计检验. 人们根据样本数据的分析结果，对研究对象的总体进行统计推断，包括推断研究对象的分布规律、推断不同因素间的相关性、判断多个总体间的差异性等.

1.3.3　统计学的应用

科学技术和社会经济的绝大多数领域都要用到统计方法. 经济学、金融学、管理学、工程学、气象学、生物学、医药学、遗传学、地质学、教育学等，这些学科都在大量使用统计方法. 随着大数据分析技术的发展，数据驱动管理决策的重要性日益凸显.

随着统计学在不同领域的应用和发展，先后出现了一些统计学的专业分支，例如“生物统计学”“教育统计学”等. 统计学在应用于社会经济领域的过程中，逐渐形成了“社会经济统计学”这一统计学分支. 本书主要从社会经济及管理科学的视角介绍统计学的基本概念、基本方法和基本应用，因此也可以称作“经济与管理统计学”.

统计学在社会经济及管理领域中的应用主要包括以下几个方面.

1. 企业发展战略规划

制定企业发展战略就是对企业的长远发展作出规划. 一方面，需要及时了解和把握整个宏观经济的状况及发展变化趋势，了解市场的变化；另一方面，企业要进行合理的市场定位，把握自身的优势和劣势. 这一决策过程需要统计工作提供可靠的数据，利用统计方法对数据进行科学的分析和预测等.

2. 产品质量管理

产品质量是企业可持续发展的基础，质量管理中离不开统计方法的应用. 质量控制已成为统计学在生产领域的一项重要应用，各种统计质量控制图广泛用于监测生产过程.

3. 市场调研

企业要在激烈的市场竞争中取得优势，就必须了解市场，进行市场调查，取得所需的信息，并对这些信息进行科学的分析，以此作为生产和营销的依据，这些都需要统计方法的支持.

4. 财务分析

企业的投资离不开对财务数据的分析. 投资管理者通过项目投资回报率及其风险的分析，确定是否对某个项目进行风险投资，其中要用到大量的统计方法. 上市公司的财务数据是股民投资的重要参考依据. 投资咨询公司根据上市公司提供的财务数据进行统计分析，为股民提供翔实的参考信息.

5. 经济预测

企业要对未来的市场状况进行预测，经济学家也要对宏观经济的某一方面进行预测. 在进行预测时要使用各种统计信息和统计方法. 比如，企业要对产品的市场潜力作出预测，以便及时调整生产计划，这就需要利用市场调查取得数据，并对数据进行统计分析. 经济学家在预测通货膨胀时，要利用有关生产价格指数、居民消费价格指数、失业率、生产能力利用等统计数据，然后通过统计模型进行预测.

6. 人力资源管理

企业利用统计方法对员工的年龄、性别、受教育程度、工资等进行分析，并作为企业制订工资计划及奖惩制度的依据.

7. 商业智能

对数据进行抽取、转换和装载，合并到数据库中，利用合适的数据挖掘技术以及联机处理工具等进行分析和处理，将数据转化为知识，为管理决策服务.

1.4 统计学的基本术语

1.4.1 总体与样本

1. 总体与总体单位

总体(population) 也称为母体(parent population)，是指在一定的研究目的下所要研

究的、具有共同属性的事物的集合. 或者说, 总体是指满足同质性、群体性和变异性的全部元素的集合体. 例如, 我国城镇居民的家庭可以集合在一起构成总体, 这是因为它们具有共同的性质(同质性), 即它们都是我国的城镇居民, 都有一定的收入和支出, 都要消费一定的商品和服务. 有了这个总体, 我们就可以研究全国城镇居民的各种数量特征, 如人均收入、人均消费、低收入群体的比例等. 根据不同的研究目的, 可以确定不同的总体. 例如, 某专业同一年的所有考生, 一条生产线上生产出来的零件, 由多个企业构成的集合, 多个居民用户构成的集合等, 都可以看作不同的总体.

组成总体的每个元素称为总体单位(overall unit)或个体. 例如, 在由多个企业构成的总体中, 每个企业就是一个个体; 由多个居民构成的总体中, 每个居民就是一个个体等.

总体根据其所包含的单位数目可以分为有限总体和无限总体. 有限总体是指总体的范围能够明确界定, 而且元素的数目是有限可数的. 比如, 由若干个企业构成的总体就是有限总体.

无限总体是指总体所包括的元素是无限的(无限可列或不可列). 例如, 要检验一批灯泡的寿命, 这批灯泡构成的集合就是总体. 在这个总体中, 我们只关心每个灯泡的寿命, 而不是灯泡本身, 所以也可以把这批灯泡的寿命的集合作为总体. 如果我们要全面考察某企业生产的灯泡寿命, 可能的寿命是一个无限区间$[0,+\infty)$, 这是一个无限总体. 此时该企业生产的灯泡寿命总体是取值于区间$[0,+\infty)$上的一个变量.

2. 样本

样本(sample)是从总体中抽取的一部分元素(个体)的集合, 构成样本的元素的数目称为样本容量(sample size). 统计研究的主要任务之一是推断总体的数量特征. 在大多数情况下, 要么总体中的单位数量很大(甚至面对无限总体), 要么调查成本过高, 导致人们无法对总体做全面调查, 这时就需要借助样本来研究总体特征.

来自总体的样本包括随机样本(random sample)和非随机样本(unrandom sample). 推断统计学所使用的样本都是随机样本. 使用随机样本的目的是根据样本提供的信息推断总体的本质特征. 比如, 从一批灯泡中随机抽取100个, 就构成了一个容量为100的样本, 然后根据这100个灯泡的平均使用寿命去推断这批灯泡的平均使用寿命.

1.4.2 参数与统计量

1. 总体参数

总体参数简称参数(parameter), 是用来描述总体特征的概括性数字度量, 它是研究者想要了解的总体数量方面的某种特征值. 总体均值、总体标准差、总体比例(概率)是三个最常用的总体参数. 总体参数通常用希腊字母表示. 本书中, 总体均值、总体标准差、总体比例(概率)分别用μ, σ, π表示.

如上所述, 在实际中, 由于我们难以对总体进行全面调查, 所以总体数据及其分布状态通常是不知道的, 所以总体参数通常是一个未知的常数. 比如, 我们不知道某一地

此为试读,需要完整PDF请访问: www.ertongbook.com