

普通高等教育经管类专业系列教材

应用统计学

(第六版)(微课版)

卢冶飞 主 编
王治皓 廖科智 副主编

清华大学出版社
北 京



内 容 简 介

本书结合作者多年的教学实践经验和国内外优秀统计学教材的成果,在内容上包括描述统计方法、推断统计方法及社会经济管理中常用的一些统计方法;在写法上与计算机紧密结合,大部分统计方法都给出了 Excel 与 R 语言的具体操作方法与数据分析结果的解析;在编排上把 Excel 与 R 语言在统计学中应用的具体操作集中在第 11 章与第 12 章阐述,保证了应用统计学理论学习的连贯性;将相关章节中演绎统计方法例题的演算与第 11 章中运用 Excel 和第 12 章运用 R 语言进行数据处理的例题进行呼应,既简化了统计方法的演绎过程,又自然而然地引导学生运用 Excel 或 R 语言进行快捷的数据处理。

本书在尽量保持系统性的基础上,结合中国社会经济管理的实际,尽力做到清晰、易懂,使其能适用于非数学专业的学生。

本书配套的教学大纲、电子课件、习题答案、案例库、上机实验数据及处理结果、各章案例及参考答案,以及曾获浙江省统计调查大赛一等奖的若干作品,可以上 <http://www.tupwk.com.cn/download> 网站下载,也可以扫描前言中的二维码获取。扫描前言中的视频二维码可以直接观看教学视频。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。举报:010-62782989, beiqinquan@tup.tsinghua.edu.cn。

图书在版编目(CIP)数据

应用统计学:微课版/卢冶飞主编.--6版.

北京:清华大学出版社,2025.6.--(普通高等教育经管类专业系列教材).--ISBN 978-7-302-69204-1

I. C8

中国国家版本馆 CIP 数据核字第 2025RL7168 号

责任编辑:胡辰浩

封面设计:周晓亮

版式设计:恒复文化

责任校对:成凤进

责任印制:宋林

出版发行:清华大学出版社

网 址: <https://www.tup.com.cn>, <https://www.wqxuetang.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-83470000 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:三河市龙大印装有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:23.5 字 数:558 千字

版 次:2012 年 3 月第 1 版 2025 年 6 月第 6 版 印 次:2025 年 6 月第 1 次印刷

定 价:79.80 元

产品编号:108350-01

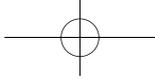
统计信息是社会经济信息的主体，人们从事经济研究和各种社会经济管理活动，都应学会运用统计工具，掌握基本的统计理论和方法，这样才能做好工作。因此，统计学是各大院校经济管理类专业不可或缺的一门专业基础课程。我国传统的统计学教材，往往从统计理论的角度出发，内容抽象，但逻辑推理与论证严密。可这样的统计学教材内容略显枯燥、生硬，而且需要学生具有扎实的数学基础知识，以及一定的抽象思维、逻辑推理和计算能力，学生不容易自学。

应用统计学可以从实际出发，通过案例讲故事式地阐明统计理论和方法。那该如何运用案例，将学生引入应用统计学领域，帮助学生构建统计学的知识体系和领悟统计思想，并掌握基本的统计理论和方法，使学生具有运用统计方法解决实际问题的能力呢？这需要具有统计学知识体系的教师们有的放矢地、耐心地积累我国社会经济实践中的统计案例，然后精心地把抽象、枯燥、生硬的统计理论与方法融入案例中，或者引入来自实践的案例，让学生轻松建立统计思想，掌握如何应用统计学的概念、理论与方法来解决社会经济中的各种问题。换句话说，就是要编写出以学生为中心的应用统计学教材。

当申报的《应用统计学》被立为2009年度浙江省高校重点教材建设项目(项目编号：ZJB2009037)后，我们秉着以学生为中心的目的精心进行编写，于2012年3月付印出版。我们对首版就能有幸入选“十二五”普通高等教育本科国家级规划教材与浙江省“十二五”优秀教材而感到欣慰。

我们对本书的前五版得到认可，并获积极的反响而向广大读者表示感谢；对在前五版使用过程中反馈的宝贵意见与建议做了梳理，并进一步思考了许多细节问题。第六版在保留了前五版的叙述风格与可读性的基础上，内容更加系统、紧凑与充实。比如，相关章节中演绎统计方法例题的演算与第11章中运用Excel进行数据处理的例题做了更好的呼应，既简化了统计方法演绎过程，又自然而然地引导学生运用Excel统计软件进行快捷的数据处理；进而继续思考与吸收每一版使用过程中反馈的宝贵意见，努力使本版内容更加完善。本书配套的案例库进一步完善，更好地支持开展案例教学，更有利于开放教学。本书与时俱进，更改了Excel统计软件的使用版本，增加了R语言在统计学中的应用，补充了新近的浙江省统计调查大赛的部分优选作品；为更便捷高效地开展教学活动，充实了电子课件，配备了微课视频。概括而言，本书具有以下几个特点。

第一，本书将统计理论与方法置于社会经济背景下展开讨论，使统计学的应用性得以自然呈现。这有利于学生更切实地理解统计理论与方法，使其在学习统计学的理论与方法



的同时,自然地吸收具体领域的知识,便于提高学生的学习兴趣,避免死记硬背,拓展教学空间。比如,在第1章中,统计指标体系的阐述安排了工业企业生产经营活动过程及结果的描述,使读者可以简约地了解工业企业统计的全貌,使统计指标体系概念具体化,且借机补充了经济、管理类专业学生的知识点,促进不同学科知识的融合。

第二,每章都以经典案例的形式开头,举例有的放矢,将抽象的统计学理论与方法融入来自实践的案例中,同时强调运用Excel来解决这些实际问题的方法。这种方式不仅有利于学生深切体会到统计学源于实践、用于实践,而且有利于他们理解统计理论,举重若轻地构建统计学的知识体系,还有利于提高学生运用统计方法解决实际问题的能力。如统计分组,通过“中国A股房地产上市公司财务负责人年薪”的举例,演绎出两种分组结果,分析其作用与优缺点,总结如何具体问题具体分析,如何恰当取舍,揭示事物的本质,在此基础上归纳统计分组的知识体系;运用Excel或R语言进行分组数据处理,使枯燥、烦琐的数据处理变得生动有趣又便捷,以引导学生积极运用统计软件进行数据处理,提高其数据处理的能力。

第三,运用框架图,帮助学生理解抽象的概念。由于统计学有些概念较抽象,不容易理解,因此本书充分利用图表的优势,构建概念框架图,帮助学生理解抽象、难懂的概念,领会统计思想。

第四,每章内容都设有本章小结。由于本书的部分内容有一定难度,不容易把握重点,因此,为了帮助学生理清思路,我们对每章内容都进行了小结,以便学生理解和掌握。

第五,配套的案例与上机实验数据及处理结果有利于教师在教学中开展案例教学。每章的练习题不仅设计了帮助读者理解、掌握统计基本理论和方法的思考题、选择题与计算分析题,还安排了便于在教学中实践统计学应用的案例分析题,这些练习提高了读者运用统计基本理论与方法来解决实际问题的能力。精选的配套案例题与本书提供的从易到难的三级案例库,能支持教师开展案例教学,尤其是从学生统计调查研究竞赛获奖作品与作者科研成果中精选提炼的案例,极具现实感,能激发学生关注现实社会、经济与生活,引发其思考、调研现实问题。

此外,本书在局部内容上存在亮点。例如,在数据获取上,系统地阐述了获取原始数据的科学方法与次级数据的来源渠道,内容具体实用,引领前沿。如具体给出政府机构与国际组织的次级数据网站;国内外出色的专业研究机构提供次级数据的公司或网站;顺应大数据时代背景,展望革命性的可扩展标识语言,思考部门统计与行政记录孤岛问题等;因素分析方法破解了同度量因素固定期的确定的难题。

在实际使用中,由于微软已终止对Excel 2010的服务支持,考虑到Excel版本的更新与我国用户的实际使用情况,本书选择采用Excel 2016作为实例操作的软件。Excel 2016在功能上相较于早期版本有了显著提升,尤其是在数据处理、图表绘制和动态分析等方面更加高效,同时其与Windows系统和Office套件的兼容性依然保持良好,能够满足绝大多数用户的需求。这一版本目前在我国具有较高的普及率,操作界面更加现代化,适合教学 and 实际工作应用,因此被本书选为主要演示工具。

此外,本书特别加入了第12章“R语言在统计学中的应用”。选择R语言主要基于其开

源性质和广泛的应用场景。R语言作为一种专为统计学和数据科学设计的开源工具，在数据分析、可视化以及统计建模方面具有极强的灵活性和适用性。企业信息系统中开源ERP的广泛使用也说明了开源技术的巨大潜力，而R语言作为一种典型的开源统计工具，既契合了开源技术的趋势，又为用户提供了深入探索和自定义分析的可能性。

在本次修订中，邀请了承接我课的同事——王治皓、廖科智两位博士来担任本书的副主编。王治皓、廖科智具备深厚的经济学与统计学理论基础，能够娴熟运用多种统计软件，在数据处理和分析领域积累了丰富的经验。他们的专业背景和实践能力，使他们成为本书副主编的理想人选。王治皓修订了第8章“相关分析和回归分析”，增加了回归模型的稳健性检验，使本章内容更趋完善，还撰写了第12章“R语言在统计学中的应用”；廖科智修改了第11章“Excel在统计学中的应用”，使得本书能够跟得上时代的技术发展。两位副主编的加入，不仅提高了本书内容的深度与广度，还使其更能适应当下统计学教学和技术发展的需求，为读者提供更全面、更实用的学习资源。

本书共分12章，第1、2、3、10章由卢冶飞编写；第4、5、6、7章由孙忠宝编写；第8、9章由卢冶飞、蒋烨编写，第8章由王治皓修订；第11章由朱海华编写，廖科智修改；第12章由王治皓编写、廖科智审核。全书由卢冶飞担任总纂并定稿。本版配备的微课视频由王治皓、廖科智制作完成。在本书的修改过程中，得到了金鑫、傅阳、卢狄的帮助，在此深表谢意。本书在编写过程中参阅了大量的统计学教材及有关论著，在此谨向文献的作者表示衷心的感谢！

由于我们水平有限，书中难免有不足之处，恳请专家和广大读者批评指正。我们的电话是010-62796045，信箱是992116@qq.com。

本书配套的教学大纲、电子课件、习题答案、案例库、上机实验数据及处理结果、各章案例及参考答案，以及曾获浙江省统计调查大赛一等奖的若干作品，可以上<http://www.tupwk.com.cn/downpage>网站下载，也可以扫描下方二维码获取。扫描下方二维码可以直接观看教学视频。

扫描下载



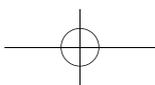
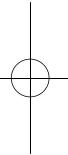
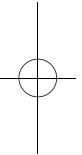
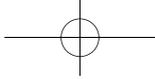
配套资源

扫一扫



看视频

卢冶飞
2025年3月



目 录

第 1 章 绪论	1	第 3 章 数据整理	44
1.1 统计的含义	2	3.1 数据整理概述	44
1.2 统计学的研究对象及其特点	3	3.1.1 数据整理的步骤	45
1.3 统计学在经济研究和管理中的 应用	4	3.1.2 数据分组	45
1.4 统计学的基本范畴	7	3.1.3 数据汇总	48
1.4.1 统计学的基本概念	8	3.2 数据整理结果的描述：统计指标	57
1.4.2 统计思想的概念框架	10	3.2.1 总量指标	58
1.4.3 统计数据的研究过程	11	3.2.2 平均指标	59
1.5 描述统计与推断统计	13	3.2.3 相对指标	60
1.5.1 描述统计	13	3.3 数据整理结果的描述：统计表和 统计图	65
1.5.2 推断统计	15	3.3.1 统计表	65
1.6 计算机在统计中的应用	16	3.3.2 统计图	66
1.6.1 计算机在描述统计中的应用	17	本章小结	69
1.6.2 计算机在推断统计中的应用	17	练习题	70
1.6.3 计算机在统计分析中的应用	17	第 4 章 数据分布特征的度量	74
本章小结	17	4.1 集中趋势的度量	74
练习题	18	4.1.1 众数	75
第 2 章 数据与数据收集	21	4.1.2 中位数	76
2.1 数据	21	4.1.3 算术平均数	78
2.1.1 数据的测量尺度	21	4.1.4 调和平均数	81
2.1.2 统计数据的常用类型	23	4.1.5 几何平均数	82
2.2 数据的收集	25	4.1.6 众数、中位数和平均数的比较	83
2.2.1 原始数据的收集	25	4.2 离散程度的度量	84
2.2.2 次级数据的收集	34	4.2.1 极差	85
本章小结	39	4.2.2 异众比率	85
练习题	39	4.2.3 四分位差	85
		4.2.4 平均差	87
		4.2.5 方差和标准差	88

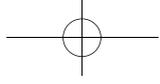
4.2.6	离散系数	90
4.3	偏态和峰态的度量	90
4.3.1	矩的基本形式	91
4.3.2	偏度系数	92
4.3.3	峰度系数	92
	本章小结	94
	练习题	94
第5章	参数估计	98
5.1	抽样调查的一般问题	98
5.1.1	抽样调查的概念、特点和作用	98
5.1.2	抽样推断的几个基本概念	100
5.1.3	抽样分布理论	102
5.2	参数估计	111
5.2.1	参数估计的一般问题	112
5.2.2	一个总体参数的区间估计	115
5.2.3	两个总体参数的区间估计	120
5.2.4	样本容量的确定	128
5.3	抽样设计	131
5.3.1	抽样设计的基本原则	131
5.3.2	抽样组织设计	132
	本章小结	138
	练习题	139
第6章	假设检验	144
6.1	假设检验的基本问题	144
6.1.1	假设的陈述	144
6.1.2	两类错误与显著性水平	147
6.1.3	检验统计量与拒绝域	148
6.1.4	利用 P 值进行决策	150
6.2	一个总体参数的检验	152
6.2.1	总体均值的检验	153
6.2.2	总体成数的检验	157
6.2.3	总体方差的检验	158
6.3	两个总体参数的检验	160
6.3.1	两个总体均值之差的检验	160
6.3.2	两个总体成数之差的检验	165
6.3.3	两个总体方差比的检验	167
	本章小结	170
	练习题	170

第7章 方差分析 175

7.1	方差分析的一般问题	175
7.1.1	方差分析的基本概念	176
7.1.2	方差分析的基本思想	177
7.1.3	方差分析检验的一般形式及基本假定	177
7.2	单因素方差分析	178
7.2.1	分析步骤	178
7.2.2	方差分析表	180
7.2.3	方差分析中的多重比较	181
7.3	双因素方差分析	182
7.3.1	双因素方差分析及其类型	182
7.3.2	无交互作用的双因素方差分析	183
7.3.3	有交互作用的双因素方差分析	185
	本章小结	188
	练习题	188

第8章 相关分析和回归分析 193

8.1	相关分析	194
8.1.1	相关关系的概念和种类	194
8.1.2	相关分析的图表和意义	195
8.1.3	简单线性相关	198
8.2	线性回归分析	200
8.2.1	回归分析	200
8.2.2	一元线性回归模型	203
8.2.3	多元线性回归模型	207
8.3	回归模型的统计检验和预测	210
8.3.1	模型的拟合优度检验	210
8.3.2	模型的显著性检验	212
8.3.3	解释变量的显著性检验	213
8.3.4	预测	214
8.4	回归模型的稳健性检验	216
8.4.1	回归模型的多重共线性检验	217
8.4.2	回归模型的异方差检验	218
8.4.3	回归模型的自相关检验	219
8.4.4	其他模型设定与数据问题	221
8.5	非线性回归分析	222
8.5.1	双曲线	222
8.5.2	幂函数曲线	222
	本章小结	223



练习题·····	224	10.4 指数分析法·····	276
第 9 章 时间序列分析 ·····	229	10.4.1 指数体系及其作用·····	276
9.1 时间序列的编制·····	230	10.4.2 因素分析法·····	277
9.1.1 时间序列的概念和意义·····	230	本章小结·····	286
9.1.2 时间序列的种类·····	231	练习题·····	286
9.1.3 时间序列的编制原则·····	232	第 11 章 Excel在统计学中的应用 ·····	290
9.2 时间序列的对比分析·····	233	11.1 Excel统计功能介绍·····	290
9.2.1 发展水平与平均发展水平·····	233	11.1.1 统计函数·····	290
9.2.2 增长量和平均增长量·····	236	11.1.2 数据分析工具·····	290
9.2.3 发展速度和增长速度·····	237	11.1.3 统计图表·····	292
9.2.4 平均发展速度与平均增长速度·····	239	11.2 用Excel进行统计数据的 整理与显示·····	292
9.3 时间序列的成分及其分析·····	240	11.2.1 数据排序·····	293
9.3.1 时间序列的成分·····	240	11.2.2 频数分布函数·····	293
9.3.2 时间序列的分解模型·····	241	11.2.3 数据透视表·····	293
9.3.3 长期趋势分析·····	242	11.2.4 统计图·····	294
9.3.4 季节成分分析·····	247	11.3 用Excel工作表函数描述 统计量·····	296
9.3.5 循环成分与不规则成分分析·····	250	11.3.1 用Excel工作表函数描述 集中趋势·····	298
9.4 时间序列的预测方法·····	250	11.3.2 用Excel工作表函数描述 离中趋势·····	298
9.4.1 趋势外推法·····	251	11.3.3 用Excel工作表函数描述 分布形态·····	299
9.4.2 移动平均预测·····	252	11.4 抽样推断·····	300
9.4.3 指数平滑预测·····	254	11.4.1 抽样与抽样分布·····	300
9.4.4 预测误差·····	257	11.4.2 参数估计·····	301
本章小结·····	258	11.4.3 假设检验·····	302
练习题·····	258	11.5 方差分析·····	304
第 10 章 统计指数 ·····	262	11.5.1 单因素方差分析·····	304
10.1 统计指数的概念和类别·····	262	11.5.2 双因素方差分析·····	306
10.1.1 统计指数的概念·····	262	11.6 用Excel进行相关与回归分析·····	308
10.1.2 统计指数的分类·····	263	11.6.1 用Excel进行相关分析·····	310
10.2 总指数的编制·····	264	11.6.2 用Excel进行回归分析·····	310
10.2.1 个体指数与总指数性质 及其关系·····	264	11.7 时间数列分析·····	313
10.2.2 总指数的编制方法·····	265	11.7.1 用Excel作趋势图直接预测·····	314
10.2.3 直接影响统计指数功能发挥的 基本要素·····	272	11.7.2 利用移动平均分析工具 进行预测·····	316
10.3 几种重要的常用指数·····	273		
10.3.1 零售价格指数·····	273		
10.3.2 消费价格指数·····	274		
10.3.3 股票价格指数·····	275		

11.7.3	使用直线函数LINEST和趋势函数TREND进行线性预测	317
11.7.4	指数平滑法与预测	319
	本章小结	320
	练习题	321
第12章 R语言在统计学中的应用 325		
12.1	R语言与RStudio初步介绍	325
12.1.1	R语言的安装与使用	326
12.1.2	RStudio的安装与使用	326
12.1.3	R语言脚本文件的新建、编辑与保存	327
12.1.4	在RStudio中设定工作文件夹	327
12.1.5	程序包的源设置与加载	328
12.1.6	在RStudio中查看帮助文档	329
12.2	用R语言进行数据的初步处理与显示	330
12.2.1	数据的读取与保存	330
12.2.2	数据的图形化展示	335
12.3	用R语言描述统计量	337
12.3.1	用R语言描述集中趋势	337
12.3.2	用R语言描述离中趋势	339
12.3.3	用R语言描述分布形态	341
12.4	抽样推断	342
12.4.1	抽样与抽样分布	342
12.4.2	参数估计	343
12.4.3	假设检验	346
12.5	方差分析	348
12.5.1	单因素方差分析	348
12.5.2	双因素方差分析	350
12.6	用R语言进行相关与回归分析	351
12.6.1	用R语言进行相关分析	351
12.6.2	用R语言进行回归分析	353
12.6.3	用R语言进行回归预测	355
12.6.4	用R语言进一步检验分析	356
12.7	时间数列分析	357
12.7.1	用R语言绘制趋势图直接进行预测	357
12.7.2	利用移动平均分析工具进行预测	359
12.7.3	利用自回归分析工具进行预测	360
	本章小结	361
	练习题	362
	参考文献	363
	附录A 标准正态分布表	364
	附录B χ^2 分布表	364
	附录C t 分布表	364
	附录D F 分布表	364

第1章

绪论

【案例】为了落实教学组织纪律，提高教学质量，××大学教务处要求，课程主讲教师平时记录每位同学听课、作业等学习情况，在每门课程结束后，填写教学质量分析表。以下是一份××大学应用统计学课程考试质量分析表。

××大学课程考试质量分析表

教师姓名		教师职称		所在学院	经贸学院
课程名称	应用统计学		考试学期		
课程总学时	40	本学期学时	40	考试形式	闭卷笔试
授课各班级	财管、旅管、国贸等专业			学生总人数	422

一、试卷分析：

试题来源	自拟	试题题量	较大	试题难度	中上
------	----	------	----	------	----

二、考试成绩及分布：

考试卷面成绩			学期总评成绩		
分数/分	人数/人	百分比/%	分数/分	人数/人	百分比/%
100~90	23	5.45	100~90	23	5.45
89~80	96	22.75	89~80	98	23.22
79~70	116	27.49	79~70	115	27.25
69~60	113	26.78	69~60	158	37.44
59~40	61	14.45	59~40	17	4.03
39~0	13	3.08	39~0	11	2.61
合计	422	100.00	合计	422	100.00

三、考试成绩结果分析和意见：

最终成绩以期终考试成绩(占60%)为基础，结合平时听课(占20%)与平时作业(占20%)情况综合评定。期终考试采用闭卷笔试的方式，试卷题量较大，涉及概念、基本知识与综合分析题，计算题难度适中。422位学生卷面平均成绩仅70分，成绩分布处于右偏(分布图略)，且有17.53%的学生不及格。卷面失分主要分布在概念、基本知识与综合分析3类题型上，尤其是概念表达的严谨性、准确性不甚理想。卷面成绩处于右偏分布的主要原因是，文科生的计算题得分不尽如人意，也因此财管(2)、旅管(1)、国贸(3)及国贸(4)班的成绩相对差些

(续表)

课程考试质量分析提醒教师至少有3个问题要引起注意:①概念题的失分,反映学生对概念的理解欠严谨、准确,这要求教师研究如何让学生严谨、准确地理解概念;②文科生在计算题得分上不尽如人意,这需要教师认真研究如何教授文理兼招、数学基础参差不齐的学生的应用统计学课;③比较卷面成绩的分布与学期总评成绩的分布,教师需要研究试卷内容的恰当性与教学组织的合理性

教师签名: 日期:

四、基层教学负责人意见:

签 名: 日 期:

五、教学院长意见:

签 名: 日 期:

- 注: 1. 此表一式两份, 学期考试结束后由课程主讲教师填写;
2. 经基层教学负责人签署意见后交学院教务秘书;
3. 再经教学院长签署意见后由教务秘书汇总, 一份交还教师保存, 另一份由学院统一装订存档, 以备教务处或督导组检查。

1.1 统计的含义

本章的开篇案例中课程主讲教师所做的教学活动, 如平时对每位学生的听课、作业等学习情况进行记录, 登记卷面考试成绩, 统计评定每位同学的课程总成绩, 分类汇总学生的卷面考试成绩与课程总成绩, 分析成绩得失分的情况与原因, 揭示教学中存在的问题等, 为进一步采取措施来提高教学质量提供了切实有效的信息。课程考试质量分析表上报学院教务管理部门, 为各级教务管理部门了解教学情况、开展有效的教学管理提供了基础资料。这项教学活动过程的组织工作、数据处理都很简单, 形成的活动成果(统计资料)——课程考试质量分析表也直观、简单明了。不难想到, 课程主讲教师对如何评定每位同学的课程总成绩、填写课程考试质量分析表等一系列的教学活动, 事前一定有所安排, 整个过程按照一定的程序、方法进行。其程序与方法虽然简单, 但蕴含着一定的统计学问。

据2024年3月9日国家统计局发布的数据: 2024年2月份, 全国居民消费价格指数(CPI)同比上涨0.7%, 环比上涨1.0%, ……。我们自然会联想到, 全国居民消费品如此之多, 许多消费品的价格随时在变化, 统计部门是怎么收集这繁多且千变万化的消费品价格数据以计算CPI的呢? 显然, 这比本章开篇案例的统计复杂得多, 每位消费者每次发生的交易事件是无法被一一记录的, 因此我们需要接受统计学理论和方法的指导, 在全国布点定时地跟踪记录部分有代表性的消费品价格数据, 再进行汇总计算推得结果(具体参见第10章的内容)。

联系以上的事实, 我们很容易理解“统计(statistics)”一词有统计活动(统计工作)、统计资料(统计数据)和统计学三重含义。

统计活动是对各种统计数据进行分析、整理、推断，并加以描述和显示的活动。统计资料是通过统计活动所获得的、能够说明现象总体某种特征的数据，以及与之相联系的文字、图表等资料的总称，是统计活动的成果。统计学则是指导统计活动的理论和方法，是关于如何收集、整理、描述和显示数据的特征，分析和探索(或推断)客观现象总体数量特征、数量关系与数量规律性的科学。由于统计学的研究领域或对象的内涵极为丰富，客观的事物又难以全面笃定把握，因此，从广义上讲，统计学是收集、分析、描述和解释数据的科学与艺术。

统计的三重含义是密切联系的。首先，统计活动与统计资料的关系是统计工作与成果的关系。一方面，统计资料的需求支配着统计活动的布局；另一方面，统计活动的好坏又直接影响着统计资料的数量和质量。其次，统计活动与统计学的关系是统计实践与统计理论的关系。统计理论是统计活动经验的总结，只有当统计活动发展到一定程度，才可能形成独立的统计学；统计活动的发展又需要统计理论的指导，统计学研究大大促进了统计工作水平的提高。

1.2 统计学的研究对象及其特点

统计学的研究对象是指统计研究所要认识的客体，其客体是大量现象的数量方面的总体特征。现象有自然现象和社会经济现象，所以，统计学的研究对象是社会经济、自然现象总体的数量特征。可见，统计学的研究领域或对象的内涵极为丰富。

不是任何客体都可以运用统计学的方法加以认识，也不是任何大量现象的数量方面的事物都必须运用统计学的方法加以认识。统计学的研究对象具有数量性、总体性、具体性和变异性四大特点。

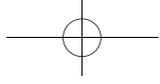
数量性指统计研究的是客观事物数量方面的特征，包括如下几方面。

- (1) 数量多少。
- (2) 现象之间的数量关系。
- (3) 质量互变的数量界限。

统计学属于定量分析的范畴。定量分析是认识客观事物不可或缺的方面，它可以使我们更精确、更具体、更深刻地把握事物的性质、特征及其变化规律。如证券投资，谁都知道其风险高，收益大；风险低，则收益小，可是又都希望能够在较低的风险程度下取得较高的投资回报率，那么风险和收益之间究竟具有什么样的联系呢？通过观察，获取大量的数据资料，建立一定的统计模型可以找出它们之间数量关系的规律性。如股价走势问题，可以通过构建资本资产定价模型研究股价波动规律，投资者可以根据股价波动规律，发现股价升、降的转折点，进行低吸高抛，实现较低的风险程度下取得较高的投资回报率的期望。

总体性指统计学研究的是由许多各不相同的个别事物组成的具有某一共同特征的整体。总体性表明统计学研究的对象不是某一个个体，而是由大量个体组成的整体性事物的某些特点。

具体性表明统计学研究的数量都是客观存在的数量。即具体事物在一定时间、地点、



条件下的数量表现,是具体的、实实在在的、有具体实物内容或计量单位的数据,不是抽象的量。这是统计学与数学的一个重要区别。因此,人们说,数学家可以端坐家中,凭借纸、笔和聪明的大脑,从假设的命题出发而推导出漂亮的结果。而统计学家则必须深入实际收集数据,才能有所作为,没有大量数据的归纳,统计学家就得出任何有益的结论。

变异性指统计学研究的是大量个体之间存在数量差异的整体性事物。如果一批物件的重量都一样,一群人的学历都相同,用不着统计就能知道这批物件的重量情况,这群人的学历状况。然而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。如掷硬币或骰子,谁都知道随机地掷一次是无法确定结果的,即每一次抛掷都结果各异,但如果我们反复不断地掷,当抛掷次数足够多时就会发现规律,即硬币出现正面或反面、骰子出现任一点数的机会都是均等的,这就是掷硬币和骰子的数量规律。统计学就是要揭示大量有差别个体数量特征下掩盖着整体性事物的某一数量特征及其规律性。

1.3 统计学在经济研究和管理中的应用

最成功的管理者和决策制定者是那些能够理解信息并有效利用信息的人。统计学的应用极为广泛,为了大致了解统计学在经济管理中的应用,我们通过以下几个方面进行阐述。

1. 个人理财中的统计

买卖股票已经逐渐成为人们生活中投资理财的方式之一,人们都希望能低价买,高价卖,获取差价收益;或能适时地买入绩优稳健成长的上市公司股票,将来获得丰厚的回报。

股价与宏观经济形势、股市的运行状态、公司的质地等因素有关。判断宏观经济形势涉及诸多因素,若进入国家统计局网站,可以查询到不同区域社会经济运行与发展情况的统计数据与一些经济走势统计图。这些数据传递的信息能帮助我们了解区域的社会经济运行与发展情况。具体来说,2009年8月开始,国家统计局每月11日左右集中发布我国经济运行的月度统计数据。2024年2月份,我国居民消费价格指数(CPI)环比上涨1.0%,同比上涨0.7%,工业生产者出厂价格指数(PPI)同比下降2.7%,……。CPI、PPI的高低与变化走势,直接影响国家的宏观经济调控措施的出台与力度,影响投资的方向与取舍。2024年9月13日,上证指数收于2704.09点,从2021年9月13日的3723.84点下跌了1019.75点,跌幅为27.38%;……。2007年10月17日上证指数收于6124.04点,比2005年6月6日的998.23点上涨了5125.81点,涨幅达413.49%,……。这些风险的变化趋势数据有助于我们了解股市系统性风险的变化,判断买卖股票的时机。贵州茅台是品牌国酒,于2001年8月27日上市,可以查到其1998年以来历年的净利润增长率、净资产收益率、净资产增长率等财务指标与其股票市场上的股价、市盈率,并将其与行业平均水平比较,有助于我们判断贵州茅台股票的相对投资价值。

2. 会计中的统计

企业会计进行产品成本核算时,为了确定单位产品成本中的材料成本,除不能互相替换的存货项目及为特定项目生产和存放的存货需要运用个别辨认法外,常常采用统计平均法(加权平均法、移动平均法)核算材料的加权平均单位成本,将其与用于该产品的材料量相乘,除以产品的产量,计算得出产品的单位材料成本。

会计师为能估计某一特定类型产品与产量相联系的生产成本,常常搜集某一特定制造业的产量与总成本的若干数据组,求出关于产量与成本的估计的回归方程,做诸如产量已知情形下的总成本预测和生产每件产品的可变成本预测,分析总成本变动中产量变动影响的程度等。

3. 审计中的统计

会计师事务所对客户进行审计时要用统计抽样方法。例如,注册会计师确定列示在客户资产负债表上的应收账款金额是否真实可靠、内容完整。客户应收账款的业务量通常很大,验证每一笔应收账款需要一定的时间和费用,所以审计人员不逐笔验证客户的应收账款业务,只从账户中随机抽取一个子集作为样本,在查看样本账户的准确性后,得出有关列示在客户资产负债表上的应收账款金额是否可以接受的结论。

4. 财务管理中的统计

在公司的日常运营中,现金流量管理是最重要的经营活动之一。是否能够保证公司拥有足够的现金收入,以满足目前和未来的偿债义务,决定着公司的财务风险状况。现金流量管理的一个关键因素是对应收账款的分析和控制。通过度量未付款发票的平均期限和资金数额,管理人员能够预测可用现金并监控应收账款状态的变化。公司设置了如下目标:未付款发票的平均期限不应超过45天;超过60天的未付款发票的资金数额不应超过所有应收账款总额的5%。

在最近对应收账款的总结中,统计未付款发票期限的结果如下。

平均数: 40天。

中位数: 35天。

众数: 30天。

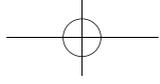
超过60天的占3%。

平均数、中位数、众数这些数据(统计指标)的含义与计算参见第4章的内容。

这些统计指标表明,未付款发票的平均数(即平均期限)是40天,没有超过45天;而中位数表明有一半的发票已经超过35天没有付款;发票期限最高的频数——众数为30天,表明一张未付款发票的最普通时间长度是30天;应收账款总额中只有3%超过60天,低于5%的要求。基于这些统计信息,管理人员可以感到满意,因为应收账款和收入现金流都处于控制之中。

5. 储存管理中的统计

在企业的成千上万种库存物资中,少数几种库存量占大部分,并占用了大部分资金。有效的仓储管理,需要对库存物品的平均资金占用额进行分析,以了解哪些物品占用资金



多,以便实行重点管理。具体做法如下。

首先,收集每种库存物品的平均库存量、每种物品的单价等数据;然后,对收集来的数据资料进行整理,分别将平均库存乘以单价,计算各种物品的平均资金占用额,并按平均资金占用额的大小进行排序,观察累计品目百分数和平均资金占用额累计百分数。将累计品目百分数为5%~15%,而平均资金占用额累计百分数为60%~80%的物品,确定为A类;将累计品目百分数为20%~30%,而平均资金占用额累计百分数也为20%~30%的物品,确定为B类;其余为C类,编制ABC分析表。当然,将ABC分析的结果绘成以累计品目百分数为横坐标,以累计资金占用额百分数为纵坐标的ABC分析图,会更直观明了。

ABC分析法理顺了品目繁多的库存物品,明确了重点,这样就可以权衡管理力量与经济效果,对三类库存物品进行有区别的管理,达到以下目的:①压缩总库存量;②解放被占压的资金;③使库存结构合理化;④节约管理力量。

6. 生产管理中的统计

控制图在质量管理中有着极为广泛的应用。人们利用控制图,科学地区分正常波动和异常波动,及时调整消除异常波动,实现对工序过程的质量波动性进行控制,使生产过程处于受控状态。例如,假定有一台250克的软饮料灌装机,定期从产品中选择一些听装饮料作为样本,计算出样本灌装量的平均值。这一平均值(\bar{x})标在一张 \bar{x} 控制图上。当该数值位于控制上限以上时,则表明产品灌装过量;当该数值位于控制下限以下时,则表明产品灌装不足;当 \bar{x} 的值位于控制图的控制上限和控制下限之间时,表示处于“控制”状态,则允许连续生产。这样, \bar{x} 控制图就能帮助确定何时必须调整和修正生产过程。

7. 市场营销中的统计

市场营销要求企业以市场需求为导向,需要实施有效的市场调研,通过统计调查,能深入研究消费心理、消费习惯,以及消费行为的特征、变化和趋向;通过深入研究其他商贸企业促销措施成败概率和实际效果,通过分析各类商品销售额和促销活动的统计资料,能更好地理解促销活动和销售额之间的关系。这样的分析对制定各种产品未来的市场营销战略大有裨益。

8. 国际贸易中的统计

无论是国际贸易的一些基本概念、理论分析还是政策分析,都大量运用统计学知识。例如,借助统计指数计算体现交换福利的贸易条件,而在计算价格指数时主要采用派氏指数法,相关的数据收集多采用海关的编码分类抽取一定的样本进行统计;在计算贸易依存度指标时,需要结合多种统计方法收集和整理大量相关数据,既涉及进出口贸易额数据,又涉及生产总值等数据;比较优势理论通过计算机会成本这一指标进行数据论证,要素禀赋学说通过要素密集度和要素丰裕度两个相对指标的计算来确定贸易伙伴国之间的比较优势,产业内贸易理论借助产业内贸易指数验证一国国际分工及对外贸易的深度,并且借助规模报酬和规模经济等指标说明产业内贸易的必要性;在国际贸易政策分析中通常结合进出口额及关税税率等数据,利用加权算术平均数法进行核算,分析各国关税水平。

9. 金融中的统计

根据《金融统计管理规定》，中国人民银行总行定期公布全国性金融统计资料，即月后20日内通过新闻媒体和中国人民银行网站，向全社会公布月度金融机构货币供应量、信贷收支及资产负债主要指标等金融统计资料。依据金融、经济统计资料，通过一定的计量方法，我们可以分析信贷规模和货币供应量与经济活动之间的关系，并解释其对实体经济的作用，帮助理解我国货币政策传导机制，以更好地把握货币政策调控目标的货币供应量和信贷规模。

假如某公司准备投资5000万元的金融资产(股票、外汇等)，其1个月后损失超过50万元的可能性有多大？能有90%的把握保证损失不会超过多少？实际上，这是风险预测：判定在持有期末，因资产价格的变化造成的损失超过限定额度的概率，并且以给定的置信度确定持有期末可能损失的最大额度。如果要限定1个月后损失超过50万元的可能性不大于1%，那么初始投资额最多应为多少？这是一个投资决策问题：限定能承受的损失额度，并使得在持有期末的可能损失不超过限定额度的概率低于某个非常低的水平，然后在这样的要求下决定初始投资额度。这显然是运用统计方法解决金融中的问题。

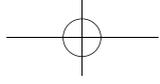
10. 经济研究中的统计

由于现代经济发展日益错综复杂，经济学研究又不断深化，对分析精确性的要求越来越高，因此，现代经济学在研究方法上大量运用统计学、计量经济学的方法，进行经济数量关系的分析。比如，人们经常要求经济学家对未来的经济或其某一方面的发展做出预测。在预测时需要用到各种统计数据，还要有适当的统计方法。例如，在预测通货膨胀率时，经济学家利用诸如生产者价格指数、失业率、制造业开工率等指标的统计数据，把这些统计指标输入预测通货膨胀率的计算预测模型中，进行运算、分析与解释。

统计方法在经济管理中应用的例子不胜枚举。总之在宏观经济管理中，政府会进行大量的统计调查和统计分析预测：GDP增长及其走势；产业结构状况及其变动趋势；每月的居民消费价格指数及由此可反映出的通货膨胀情况；房屋动工数量及货币回笼速度等统计指标。政府依据统计数据制定宏观经济政策，并利用统计数据评价与检查施政的绩效。在工商企业管理中，管理人员经常需要在未来条件不确定的情况下，做出牵涉资金数额巨大甚至是有关企业生死存亡的重大经营决策，而在日常的经营管理中，更时时面临着各种各样的决策问题。利用统计决策模型，则可以降低决策风险，做出明智的选择。

1.4 统计学的基本范畴

统计学的基本范畴是阐明各个统计方法所蕴含的统计思想的核心概念，通常称其为统计学的基本概念。统计学的基本概念是统计认识的基本工具，深刻理解统计学的基本概念，是学好应用统计学这门课程的基础。



1.4.1 统计学的基本概念

构成统计学基本范畴的基本概念,即统计学的基本概念,包括统计总体、样本和总体单位、统计标志与变量、统计指标与指标体系等。

1. 统计总体、样本和总体单位

统计总体是指统计所要研究对象的全体,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,简称总体或母体;构成总体的每一个个别事物称总体单位,也称为个体。总体中总体单位的数量称为总体容量。

最常见的总体是由自然物体组成的总体。例如,要研究全国的人口状况,则全国人口就是总体,每一个人就是总体单位。又如,要研究一批产品的质量状况,则该批产品就是总体,每件产品是总体单位。可见,总体单位与总体的关系是个别与整体的关系,它们是相互依存的。

如果总体中只包含有限可数个单位数,则称该总体为有限总体;如果总体单位数是无限的,即总体容量为无穷大,则称该总体为无限总体。例如,全国的人口、某种产品产量等都是有限总体;而宇宙中的星球、海洋中的鱼则可看作无限总体。

统计总体和总体单位的确定取决于统计研究的目的。研究目的不同,统计总体往往也不同。例如,研究目的是了解全国工业企业的状况,则总体是全国所有工业企业组成的集合,每一个工业企业是一个总体单位;若研究目的是了解某企业职工的工资收入状况,则该企业的全部职工构成总体,每个职工就是总体单位。

样本是从总体中抽取一部分个体所组成的集合。样本中所包含的个体数,称为样本容量或样本单位数。

样本是与总体相对应的概念。总体是在一个特定研究中所有个体组成的集合,样本是总体的一个子集,是总体的代表和缩影。总体是要研究的对象,而样本则是所要观察的对象,对样本进行观察的目的是要对总体的数量特征做出估计与判断。

总体与样本也取决于统计研究的目的。例如,一个国家进行人口普查,以了解本国某一时点的人口状况,则一个国家该时点的所有人口是一个总体。但如果要从历史上动态考察该国人口变化规律,则若干时点上的人口总体就成为样本,动态上的人口总体才是与之相对应的总体。

2. 统计标志与变量

统计活动的研究对象虽然是统计总体,但是,必须从组成该总体的各个个体认识入手,如考察各个个体与总体研究目的相联系的某些特征,考察这些特征在总体中各个个体间的分布情况,以及蕴含在一个个带有一定偶然性的个体结果中的偶然现象之必然规律。例如,人口普查是为了考察全体国民的性别、民族、年龄、文化程度、职业等特征上的分布状况及其变化规律,为此,我们必须从对每个国民的性别、民族、年龄、文化程度、职业等具体表现的登记入手;又如,对工业企业总体进行调查,是为了考察不同行业、资产规模、职工人数、产值、利润等特征上企业数量的分布情况及其演变规律,所以必须从对每个企业所属行业、资产规模、职工人数、产值、利润等具体表现的登记入手。因此,要

研究统计总体的数量特征，需要从对统计总体各单位的属性或特征，以及各个总体单位在某些属性或特征上的具体表现的观察开始。

总体各单位所共同具有的属性或特征称为统计标志，简称标志，它是说明总体单位属性或特征的名称。而标志在各总体单位的具体表现则称为标志表现。例如，对于人口总体，性别、民族、年龄、文化程度、职业等都是标志，而张三是男性，汉族，25岁，大学毕业，现从事会计工作则分别是上述各标志的一个具体表现。又如，对于某区域工业企业总体，企业所属的行业、资产规模、职工人数、总产值、利润都是标志，而强盛电视机厂属于电子行业，拥有资产总额2亿元，职工3000人，营业收入4亿元，净利润4000万元则分别是上述各标志的标志表现。

由于总体是由同类事物的全体构成的，所以在一个总体中必然有些标志在各个个体上的标志表现都相同，这样的标志称为不变标志。例如，在全国人口总体中，国籍这一标志就是不变标志，这是由于每个人的国籍相同这一共性(即同质性)特征把每个人集合在一起构成了全国人口这一总体。显然，一个总体至少有一个不变标志。

总体中必然有些标志在各个总体单位上的标志表现是不同的，这样的标志称为可变标志，它们正是统计活动所要考察的对象。例如，人口总体中，性别、年龄、民族、文化程度、职业都是可变标志，因为每个人在这些标志上的标志表现会有不同(即差异性)。一个总体至少有一个可变标志。同时，构成总体的总体单位必须是大量的，才能规避由于各种偶然因素的干扰产生差异性而得出错误的结论。大量性、同质性和差异性总体必须同时具备的3个基本特性。

总体中的可变标志是统计活动的基础和依据。为了能够使用数学方法，每个可变标志都可以视为变量，用 x 或 y 表示，而各个总体单位在可变标志上的标志表现则可看成该变量的各个取值，即变量值。实际上，统计活动研究总体是研究总体中的某个变量，该变量的全部取值是统计活动的考察对象，因而，变量的全部取值也可以构成总体。

3. 统计指标与指标体系

统计指标是统计活动按照一定的统计方法，对总体单位数、总体各单位的标志表现进行记录、核算、汇总、综合而形成的，用于反映统计总体某一综合数量特征的科学范畴。它一般包括指标概念和指标数值两个方面，此外，还包括计算方法、计算单位、时间规定性和空间规定性，因此，一个完整的统计指标由以上6个要素组成。需要说明的是，在讨论统计理论时和在统计设计阶段，统计指标是指说明总体数量特征的名称。当然，根据样本中各单位的标志表现计算的是样本指标，样本指标是用来推断总体指标的。

统计指标具有数量性、综合性和具体性的特点。也就是说，它是许多个体现象具体数量综合的结果，可以说明一定的、具体的总体现象的数量特征。

统计指标是统计学中最重要的基本概念之一。统计正是用统计指标对总体的现状、发展变化、内部结构及其与外部的数量关系进行计量描述及分析研究的。统计指标在经济分析和管理决策中占有中心地位。

统计指标按所反映总体的内容和数量性质的不同，分为数量指标和质量指标。数量指标是说明总体现象的规模大小和数量多少的指标，一般用绝对数来表示，如企业总数、利

润额等。质量指标是说明总体内部、总体之间数量关系或总体单位水平的指标,是数量指标的派生指标,一般用相对数和平均数形式表示,如平均工资、价格水平、人口密度等。

统计指标按其计量单位的不同,可分为实物指标和价值指标。实物指标是根据事物的自然属性,采用自然物理单位计量的指标。实物指标的最大特征是具体明了,可直观地反映事物发展的规模和水平,是计算其他指标的基础。但计量不同事物现象的实物指标不能直接相加,缺乏综合概括能力。价值指标是以货币单位计算的统计指标,又称货币指标。其最大特点是综合性和概括能力强,但比较抽象,同时受价格水平的制约。此外,还有以劳动时间(即定额工时)为单位的劳动量指标。由于各企业的定额水平不同,该指标往往仅限于企业内部使用,不适合用于进行企业间的汇总。

一个统计总体往往有许多数量特征,可以从多方面、多角度、多层次来描述。但一个统计指标只反映其中一方面的一个特征,要全面反映总体各方面的特征,就必须构造和使用一系列相互联系的指标。这种反映同一总体现象的一系列相互联系的统计指标所形成的体系,称为统计指标体系。

例如,工业企业生产经营活动的过程既是供、产、销的过程,也是产品的价值形成与实现的过程,同时又是企业的资金周转过程,如图1-1所示。全面反映一个工业企业生产经营活动的过程及结果,就需要由工业企业的总产值(C_1+C_2+V+M)、增加值(C_2+V+M)、销售额、原材料消耗及库存、利润总额、工资总额、劳动生产率、资金利润率等一系列统计指标组成的指标体系来完成。

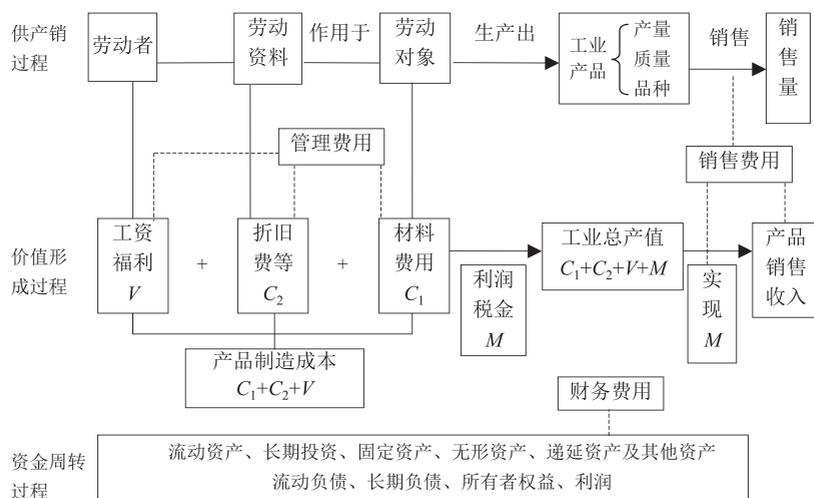


图 1-1 工业企业生产经营活动的过程及结果

1.4.2 统计思想的概念框架

构成统计学基本范畴的统计总体、样本和总体单位、统计标志与变量、统计指标与指标体系等基本概念能清晰地勾勒出统计的基本思想,如图1-2所示。

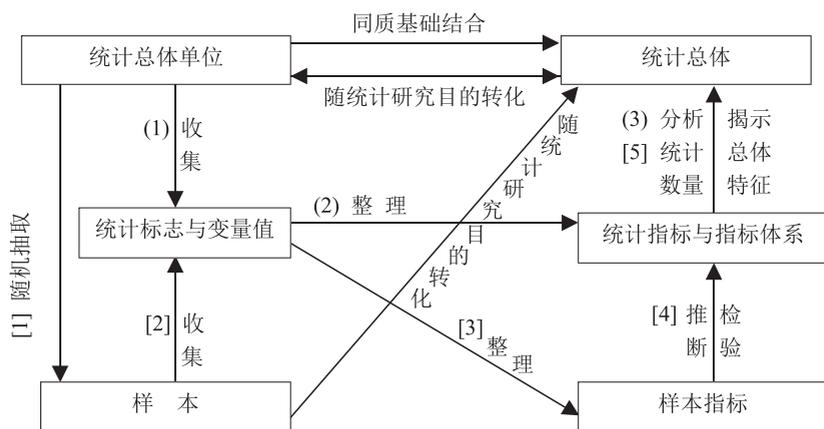


图 1-2 统计基本概念的关系

注：(1)→(2)→(3)勾勒的是全面调查下数据的收集、整理和分析路径；[1]→[2]→[3]→[4]→[5]勾勒的是抽样调查下数据的收集、整理和分析路径。

图1-2清晰地表明：统计总体是统计所要研究的对象，统计活动的目的是认识统计总体的数量特征。统计总体由总体单位构成，两者互相依存；统计标志说明总体单位的某种共同的属性或特征；统计指标说明总体的某一数量特征；统计指标体系则是多方面、多角度、多层次地描述统计总体的多重数量特征。统计指标由统计标志表现(变量值)整理所得，统计指标体系由反映同一总体现象的一系列相互联系的统计指标构成。

统计可以收集总体的全部单位的数据来认识总体的数量特征，也可以通过从总体中随机抽取一部分个体进行观察，达到认识统计总体数量特征的目的。

1.4.3 统计数据的研究过程

一项统计活动是依据该项统计活动的目的和任务，在对统计活动的内容和程序做出通盘考虑和安排的基础上，从收集统计数据(原始数据或次级数据)开始，进而对其进行整理加工，经过分析，得出结论。显然，对统计数据的研究过程可概括为：统计设计；统计数据收集；统计数据整理；统计数据分析与解释；统计资料的提供。从这5个基本环节实现统计信息、统计咨询和统计监督3种职能，如图1-3所示。

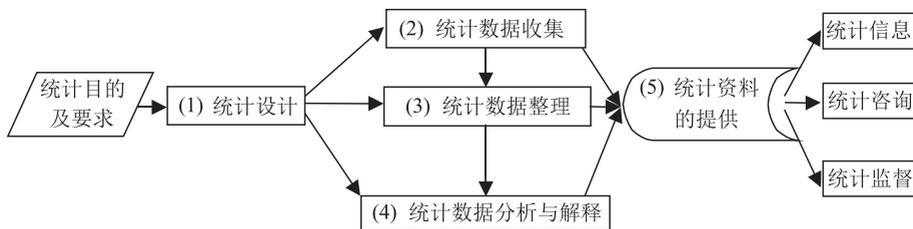
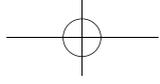


图 1-3 统计活动过程

1. 统计设计

统计设计是统计工作的第一个工作阶段，是根据统计研究对象的性质和研究的目的，对统计工作的内容和程序进行通盘考虑和安排。



统计设计的结果通常表现为各种设计方案,如统计指标和统计指标体系、分类目录、统计报表制度、调查方案、汇总整理方案等诸多方面的内容。

从认识顺序来讲,统计设计是统计活动的第一阶段,但在实际工作中,有时并不被认为是统计活动的开始,而表现为统计活动连续活动中的一个环节,即统计设计的修正与统计活动过程中对设计方案的修正和改进。

例如,一个单项的统计全阶段设计应包括以下几个方面的内容。

- (1) 明确规定统计的目的。
- (2) 确定统计研究的对象与单位。
- (3) 进行统计指标和统计指标体系的设计。
- (4) 进行与统计指标体系相联系的统计分类和统计分组设计。
- (5) 收集统计数据方法的设计。
- (6) 统计数据汇总整理的设计。
- (7) 确定为达到统计研究目的所需要的分析。
- (8) 统计力量的组织与安排。
- (9) 统计活动各个部门的协调。

2. 统计数据收集

统计数据收集是按照统计设计的要求,有针对性地获取所需要的数据的环节,是统计数据整理与分析的基础。数据收集是否准确、及时、完整,直接影响统计分析质量的高低。如果收集的数据失真,则会导致错误的分析结论。

3. 统计数据整理

统计数据整理是根据统计研究的目的和要求,对收集的数据进行审核、分类、加工、汇总和显示,使零散的资料条理化、系统化、综合化,成为能反映总体特征的统计数据的环节,也称为统计整理环节。整理结果表现为统计表、统计图或统计指标。

4. 统计数据分析与解释

统计数据分析是对加工整理好的统计资料,围绕统计研究所确定的任务,运用各种统计分析方法进行研究,得出有用的定量结论的环节,也称为统计分析环节。数据分析的实质旨在揭示研究对象的基本特征和发展规律性,并做出科学的判断和结论。统计分析是理性认识阶段,是统计发挥作用的关键环节。数据解释是对整理和分析的数据或有关数量结果进行说明,即说明为什么会得出这些数据,这些数据的含义分别是什么,从中能得出哪些具有规律性的结论,需要进一步探讨哪些问题等。数据解释是对数据分析的深化。

5. 统计资料的提供

提供统计资料是统计信息社会化的重要步骤,在全面分析、系统搜集整理各种统计信息的基础上,建立数据库和信息库,以灵活多样的方式向社会提供资料。

总之,统计数据的研究过程是通过这5个环节,从定性到定量,再到更高的定性,最后达到对社会经济和自然现象本质与规律性的认识。与此同时,统计随之实现统计信息、统计咨询与统计监督三大统计职能。

1.5 描述统计与推断统计

统计数据研究经过5个环节,各个环节需要运用不同的统计方法。综合来看,统计研究过程的起点是反映总体单位的统计数据,终点是揭示客观现象总体数量特征、数量关系,探索其数量规律性。如果要通过收集总体的全部单位的数据来认识总体的数量特征,就需要将收集的数据进行加工整理,通过表格、图形或数值形式对其进行恰当、科学的描述,其被称为描述统计;如果获得的只是样本数据,那么要揭示总体数量特征、数量关系,探索其数量规律性,就要运用概率论的理论,并根据样本信息对总体进行科学推断,其被称为推断统计。因此,贯穿于统计数据研究过程的统计研究方法主要有描述统计与推断统计两大类,它们是将数据转为有意义的、易于理解的统计信息的方法。

1.5.1 描述统计

描述统计是研究如何取得反映客观现象的数据,并将收集的数据进行加工处理,通过表格、图形或数值形式显示,进而通过综合、概括与分析得出反映客观现象总体的数量特征、数量关系与数量规律的统计方法。

我们再回头看本章的开篇案例,课程主讲教师所做的课程成绩的统计活动,就是用描述统计方法对学生的应用统计学课程学习成绩的数量特征给予综述。在此案例中,主讲教师运用观测法,记录每位学生的各种学习成绩——收集数据;计算评定学生的学期总成绩,然后分组、汇总,得出各组成绩的学生人数,以表1-1(实际工作中,直接填入课程考试质量分析表)与图1-4(实际工作中,考虑到成绩分布比较简单与课程考试质量分析表的空间有限,省略了分布图)反映学生的成绩分布,用卷面平均成绩70分,不及格的学生占17.53%等统计指标揭示学生成绩的总体数量特征——统计数据加工整理、显示;撰写考试成绩结果分析与意见——统计数据分析与解释。

表1-1 经贸学院××级学生应用统计学的成绩分布表

考试卷面成绩			学期总评成绩		
分数/分	人数/人	百分比/%	分数/分	人数/人	百分比/%
100~90	23	5.45	100~90	23	5.45
89~80	96	22.75	89~80	98	23.22
79~70	116	27.49	79~70	115	27.25
69~60	113	26.78	69~60	158	37.44
59~40	61	14.45	59~40	17	4.03
39~0	13	3.08	39~0	11	2.61
合计	422	100.00	合计	422	100.00

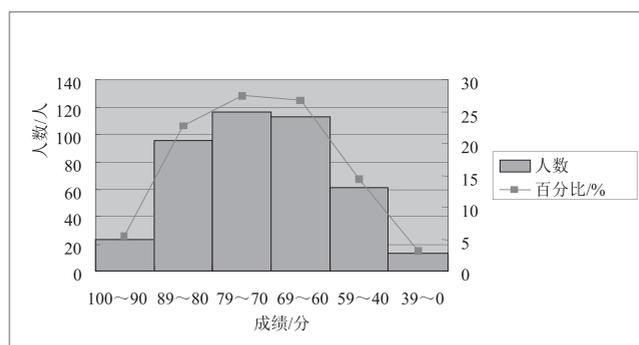


图 1-4 应用统计学卷面成绩分布

虽然本章的开篇案例是个非常简单的统计描述案例,但它让我们切实地看到了描述统计的过程与主要内容。首先要通过一定的行之有效的调查方法,收集我们关注的标志在各个总体单位中的表现,然后对这些数据进行汇总、归纳和计算,将原始资料整理成有条理的能够说明被研究的统计总体数量特征的统计指标,并运用相应的统计表、统计图将这些结果表现出来,进而进行分析与解释。

在描述统计过程中,收集统计数据可以通过查阅与询问的方式获得人们已经加工的资料,当我们需要的统计数据不能从已加工资料中获取时,可以运用大量观测(调查)法、统计实验法^[1]等统计研究的方法获得;在统计数据整理阶段有统计审核法、统计分组法、统计汇总法等方法,而数据整理结果的显示有统计表、统计图或统计指标等形式;统计数据的分析又有综合指标法、统计模型法等基本方法。描述统计的过程与主要内容如图1-5所示。

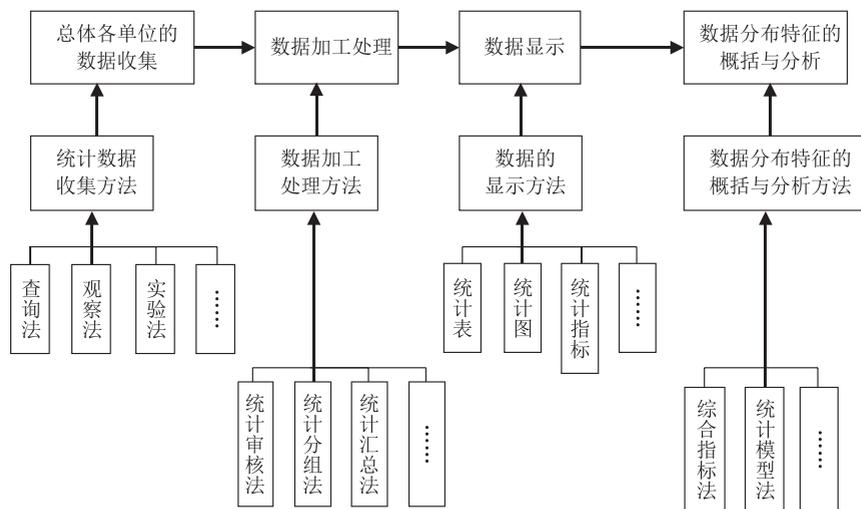


图 1-5 描述统计的过程与主要内容

描述统计方法的使用非常广泛,绝大多数的报纸、杂志、公司报告和其他出版物上的统计信息以易于读者理解的方式披露,将数据以表格、图形或数值的形式表现出来,都是

[1] 凡是在获得数据的过程中,不对调查对象数据产生的条件施加任何控制而获得数据的方法,称为观测(调查)法。凡是在获得数据的过程中,对数据产生的条件实施了控制而获得数据的方法,称为实验法。社会经济现象的统计数据收集都只能运用观察法。

在运用描述统计的方法。从现实情况来看，人们开始关注描述统计的拓展与表述，对描述统计方法的兴趣也在不断提高。

1.5.2 推断统计

推断统计主要研究如何根据样本数据来推断总体的数量特征。它的内容主要包括参数估计的方法、假设检验的方法、方差分析的方法、相关与回归分析的方法等。

在很多情形下，出于对时间、费用和其他因素的考虑，只能对样本进行调查，利用样本数据来估计总体的数量特征。比如，新宇高科电器有限责任公司是生产节能灯泡的电器公司，为了提高A种灯泡的使用寿命，产品设计部开发出一种新型灯丝的灯泡。为了评估新灯丝的优点，用生产出的80只新灯丝灯泡组成样本，并进行测试，记录每只灯泡在灯丝被烧断之前的使用小时数。记录的样本测试结果列示于表1-2。

表1-2 新宇高科电器有限责任公司80只新灯丝灯泡的使用寿命

单位：小时

1036	988	1125	995	1088	1065	1023	1075	1002	994
1047	968	1183	1058	1142	1098	945	1126	1036	987
1046	976	1087	984	1224	998	1032	1153	1103	958
1153	994	1039	1006	1214	1076	986	1078	1048	1126
1216	1122	1096	1035	1004	1053	1004	1094	1080	994
964	975	1185	1021	1007	948	1024	1136	1083	1120
1113	997	1005	1088	997	1034	985	1093	1004	1082
1047	984	1136	989	1073	1102	976	1097	1005	1152

新宇高科电器有限责任公司的产品设计部门希望利用样本数据对用新灯丝生产的所有灯泡总体的使用寿命进行推断。产品设计部门将进行测试记录的80只灯泡的使用寿命相加再除以80，计算出80只灯泡的平均使用寿命为1055小时，他们可以用这一样本结果估计灯泡总体的可能平均使用寿命或其可能平均使用寿命的区间。灯泡总体的平均使用寿命推断统计过程如图1-6所示。

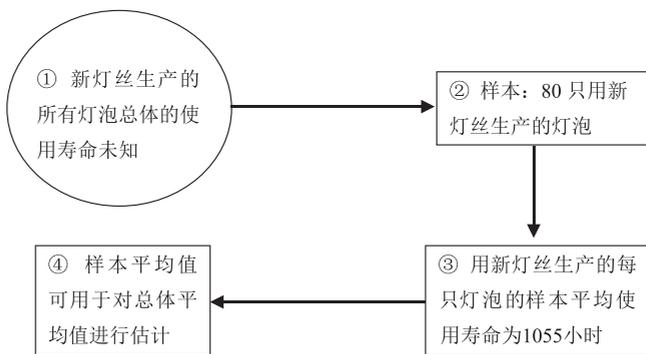


图 1-6 统计推断过程

统计利用样本对总体的特征进行推断时，通常需要相应地注明统计质量与精度。产品设计部门用样本结果估计灯泡总体的平均使用寿命为1055小时的同时，计算得出估计精度为±15小时的概率为95%。这样就可以估计得出，有95%的把握新灯泡总体的平均使用寿命

在1040小时和1070小时的区间内。

新宇高科电器有限责任公司产品设计部利用样本数据对用新灯丝生产的所有灯泡总体的使用寿命进行推断,我们能初步领会推断统计是在对样本数据进行统计描述的基础上,对统计总体的未知数量特征做出以概率形式表述的推断。其过程可用图1-7表示。

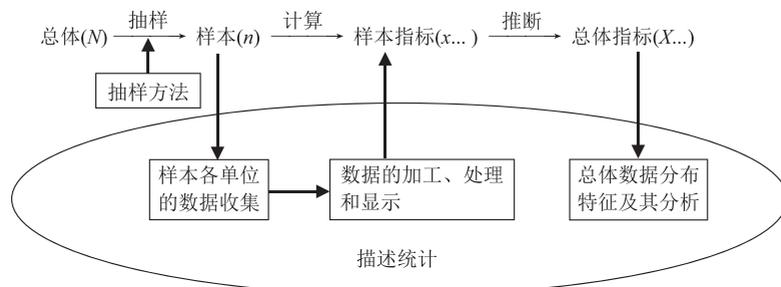


图 1-7 推断统计的过程

由于样本包含的总体信息必然不完备,用其推断总体难免有误差,为了使这种误差尽可能小,或将其控制在可以接受的范围内,统计学就需研究抽取样本的方法、推算的方法,以及推断误差的计算与控制的方法,以使推算结果的准确性尽可能满足要求,且失误尽可能小。

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,推断统计则是现代统计学的主要内容,从描述统计发展到推断统计是统计学发展成熟的重要标志。利用样本数据来估计总体的数量特征等的推断统计是统计学的一个主要贡献,它的作用越来越被人们重视。

1.6 计算机在统计中的应用

统计学的研究目标是在大量观察的基础上,揭示总体数量特征与数量关系的规律。统计数据整理与统计分析都需要对大量数据进行处理,使用手工操作对统计数据进行处理、分析不仅费时又费力,计算过程很乏味,还容易产生错误,而且手工操作对有些“海量”数据的处理是不可及的。利用统计软件进行计算机运算,极大地节省了时间,运算不易产生错误,且加快了统计信息生产流动速度,可以说,计算机技术和统计软件强大的数据处理能力使统计学的广泛应用从可能走向现实。如股票市场要求迅速从交易的数据中获得有关信息,抓住赢利的时机,计算机技术迅速的数据处理功能正好提供了这种可能性。因此,在统计学的应用中,计算机技术、统计分析方法和统计软件成为必不可少的工具。

统计软件有Excel、SPSS、EViews、MATLAB、SAS和R语言等。在众多的数据处理软件中,Excel由于能够与Windows操作系统及Office中的其他软件良好结合,而且普及面广、使用简便、功能强大、能够满足一般统计分析的需要,成为本书编纂中数据处理的重要工具。

本版除Excel软件外,还选择了R语言作为数据处理的重要工具。因为R语言作为一款专为统计分析和数据可视化而设计的开源软件,具有灵活性与扩展性、强大的图形功能、开源与社区支持等独特优势,能为应对复杂数据分析任务提供更加科学、高效的解决方案。

1.6.1 计算机在描述统计中的应用

利用计算机和互联网可以进行网上调查,快速、准确地收集原始数据和次级数据,无时空和地域限制,具有传统调查无法比拟的优越性。然后利用统计软件可以迅速对统计数据进行有效性检查、筛选、排序及分组整理等工作,使统计数据系统化、条理化,符合统计分析的要求;利用统计软件中丰富的统计表和统计图的制作与编辑功能,可以形象、直观地将统计数据的分布特征和规律性显示出来,易于读者理解;利用统计软件中强大的函数功能和分析工具,可以及时、准确地计算,并提供有关统计数据的分布特征值,如均值、方差、标准差、偏度、峰度等指标。

1.6.2 计算机在推断统计中的应用

利用统计软件中的随机数和抽样工具,可以方便地实施随机抽样、等距抽样等抽样方法,快速得到研究问题所需要的样本量。然后根据样本数据,利用统计软件中的参数估计和假设检验工具或公式与函数,对总体参数快速进行点估计、区间估计、假设检验、稳健性检验等。

1.6.3 计算机在统计分析中的应用

利用统计软件中的方差分析、相关分析、回归分析和时间序列分析工具,可以轻松测定出变量的聚散状况,抑或变量间的相关关系和相关程度,测定出现象的发展水平、发展速度、长期趋势、季节变动等,并能快速得到回归方程、趋势方程和季节指数,对未来做出预测。

总之,计算机在统计中的应用,有助于提高统计活动的效率,特别是对于一些数据量较大、难以用手工完成的统计活动,计算机更能体现出它的优越性。

本书在第11章讲解了Excel的操作步骤,对大部分统计方法给出了Excel的计算过程和结果解释。

本章小结

统计学是收集、分析、表述和解释数据的科学与艺术。要有效地从事经济研究与管理,应用统计学不可不学。

运用统计学的方法研究经济、管理问题,不仅要学好统计学知识,也需要学好经济学、管理学及其背景知识,并需要具备一定的计算机技术和统计软件的运用能力,善于将统计学与经济学、宏观经济分析与微观经济分析等结合。

构成统计学基本范畴的基本概念,包括统计总体、样本和总体单位、统计标志与变量、统计指标与指标体系等。这些基本概念清晰地勾勒出统计的基本思想与统计数据的研究过程。统计总体是由总体单位构成的,两者互相依存。统计标志说明总体单位的某种共同特征。统计指标是说明总体的某一数量特征(统计指标体系可以从多方面、多角度、多层

次来描述统计总体的多重数量特征),由统计标志表现(变量值)整理所得。统计从研究总体单位入手,目的是认识统计总体的数量特征。它可以通过收集总体的全部单位的数据来认识总体的数量特征,也可以利用样本数据来估计总体的数量特征。从这些概念的关系中可发现,统计数据的研究过程包括统计设计、统计数据收集、统计数据整理、统计数据分析与解释和统计资料的提供5个环节。

描述统计研究如何取得反映客观现象的数据,并将收集的数据进行加工处理,通过表格、图形或数值形式显示,进而通过综合、概括与分析得出反映客观现象总体的数量特征、数量关系与数量规律。推断统计主要研究如何根据样本信息来推断总体的数量特征,它是在对样本数据进行统计描述的基础上,对统计总体的未知数量特征做出以概率形式表述的推断。

统计现代化是通过计算机的应用而实现的。掌握计算机和统计软件的操作是未来开展统计活动的必备技能。

练习题

一、思考题

1. “统计”一词有哪几种含义?它们之间是什么样的关系?
2. 如何认识统计的作用?
3. 总体、总体单位、标志与指标之间的相互关系如何?
4. 什么是统计指标体系?它有哪些表现形式?
5. 为什么说描述统计是统计学的基础?

二、选择题

1. 统计学研究对象的数量性包括()。

A. 数量的多少	B. 现象之间的数量关系
C. 质量互变的数量界限	D. 数量的规律
2. 统计总体的基本特征是()。

A. 同质性、数量性、变异性	B. 数量性、具体性、综合性
C. 大量性、同质性、差异性	D. 总体性、社会性、大量性
3. 现有200家公司每位职工的工资资料,如果要调查这200家公司的工资水平情况,则统计总体为()。

A. 200家公司的全部职工	B. 200家公司
C. 200家公司全部职工的工资	D. 200家公司每位职工的工资
4. 一个统计总体()。

A. 只能有一个标志	B. 可以有多个标志
C. 只能有一个指标	D. 可以有多个指标

(2) 总体是什么?

(3) 在这种情况下为什么要用样本?

(4) 基于该传媒公司的研究, 可以进行何种决策和行动?

4. 6名学生考试成绩的样本值(分)为: 72, 65, 82, 77, 90, 76。下列表述中, 哪些是正确的? 哪些受到怀疑? 为什么?

(1) 6名学生样本的平均成绩是77分。

(2) 参加考试的所有学生平均成绩是77分。

(3) 参加考试的学生成绩的估计值是77分。

(4) 参加考试的学生中大多数人的成绩在70分至85分。

(5) 学生成绩在65分至90分。

5. (1) 写出某电视机厂、某高等学校、某医院等单位中你所熟悉的统计总体和总体单位。

(2) 写出电视机、大学生、医生等总体单位的品质标志和数量标志, 以及不变标志和可变标志。