



工业和信息化高等院校
“十三五”规划教材立项项目



21 世纪高等学校
经济管理类规划教材
高校系列

S STATISTICS

统计学原理

(附微课视频 第3版)

+ 刘泽 刘建冰 主编
+ 王苹香 张伟 严瑜 副主编

精选案例，贴近现实生活
突出实务，强化练习操作
力求创新，增加视频微课

ECONOMICS
AND
MANAGEMENT



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS



工业和信息化高等院校
“十三五”规划教材立项项目



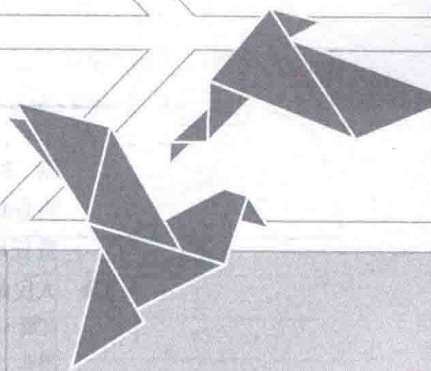
21 世纪高等学校
经济管理类规划教材
高校系列

STATISTICS

统计学原理

(附微课视频 第3版)

+ 刘泽 刘建冰 主编
+ 王苹香 张伟 严瑜 副主编



ECONOMICS

MANAGEMENT

人民邮电出版社
北京

图书在版编目 (CIP) 数据

统计学原理：附微课视频 / 刘泽, 刘建冰主编. —
3版. — 北京：人民邮电出版社, 2018.7
21世纪高等学校经济管理类规划教材. 高校系列
ISBN 978-7-115-48512-0

I. ①统… II. ①刘… ②刘… III. ①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108641号

内 容 提 要

本书系统讲述了统计学中数据收集、数据整理和数据分析的基本方法, 突出了 Excel 在数据整理和数据分析中的应用。全书共 10 章, 分别是: 总论、数据收集、数据整理与显示、数据特征的描述、抽样估计、假设检验、方差分析、统计指数、相关与回归分析及时间序列分析与预测。为了提高学生的学习兴趣, 本书每章开篇都设有一个与现实社会经济生活非常贴近的案例资料; 为了提高学生的实践能力, 本书有针对性地给出了 Excel 的操作演示; 为了让学生能够及时地检查自己的学习效果, 本书按“节”给出了相应的练习题; 为了增加学生的学习兴趣, 本书添加了视频资料, 读者扫描书中的二维码即可观看。

本书可以作为本科经管类各专业的专业基础课教材, 也可以作为成人培训教材或自学的参考资料。

◆ 主 编 刘 泽 刘建冰
副 主 编 王苹香 张 伟 严 瑜
责任编辑 李育民
责任印制 焦志炜

◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
固安县铭成印刷有限公司印刷

◆ 开本: 787×1092 1/16
印张: 15
字数: 433 千字

2018 年 7 月第 3 版
2018 年 7 月河北第 1 次印刷

定价: 45.00 元

读者服务热线: (010)81055256 印装质量热线: (010)81055316
反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147 号

前言



统计学是一门关于数据搜集、整理与分析的方法论科学，是普通高校经济类、管理类各专业的专业基础课。能够利用统计方法进行量化分析是大数据时代高校经济类、管理类各专业学生的一项必备基本技能。在本书编写过程中，编者尽量采用通俗易懂的语言来阐述统计学的基本概念、基本原理和常用方法，减少公式推导，强化统计思想的理解和统计方法的应用。本书通过案例、图示、表格、视频、软件操作等形式帮助学生加强对教学内容的理解和掌握。本书主要有以下4个特点。一是结合现实案例。每章都有一个与教学内容相关的、贴近我国社会经济生活现实的案例。二是在各章内容中添加了微课，使教学过程更加生动有趣。三是突出了 Excel 操作。本书以提高应用能力为目标，以常用的 Excel 软件为实现统计计算和分析的操作软件，可以帮助学生提高快速、准确处理数据的能力。四是编写体例新颖，结构紧凑。体例上，开篇有知识目标、能力目标和案例引入，各章通过案例引发学生的思考；接下来是知识介绍，并按“节”而不是按“章”给出相应的 Excel 操作和练习题，便于学生及时和更有针对性地学习；之后结合所学知识解析案例，引导学生学以致用；最后通过相关知识图示对全章内容进行扼要总结。

本书共分为10章，分别是：总论、数据收集、数据整理与显示、数据特征的描述、抽样估计、假设检验、方差分析、统计指数、相关与回归分析及时间序列分析与预测。本书参考学时为54学时，各章的参考学时见学时分配表。

学时分配表

章名	学时
第一章 总论	4
第二章 数据收集	4
第三章 数据整理与显示	6
第四章 数据特征的描述	8
第五章 抽样估计	6
第六章 假设检验	4
第七章 方差分析	4
第八章 统计指数	6
第九章 相关与回归分析	6
第十章 时间序列分析与预测	6
合计	54

本书由刘泽、刘建冰担任主编，王苹香、张伟、严瑜担任副主编。刘泽负责全书框架的设计及统稿，并编写了第一章、第八章、第九章、第十章；刘建冰编写了第二章、第五章；王苹香编写了第三章；张伟编写了第六章、第七章；严瑜编写了第四章。

在本书编写过程中，作者借鉴和参考了不少同仁编写的教材，在此一并表示感谢！

编 者

2018年6月

目 录 Contents

第一章 总论 / 1

案例引入 / 1

第一节 统计的内涵 / 1

- 一、统计学 / 2
- 二、统计工作 / 2
- 三、统计数据 / 3

第二节 统计的职能 / 4

第三节 统计学研究方法 / 5

第四节 统计学中的基本概念 / 5

- 一、总体、样本与个体 / 5
- 二、参数与统计量 / 6
- 三、变量 / 6
- 四、练习题 / 7

第五节 统计应用软件简介 / 8

案例解析 / 9

相关知识图示 / 10

第二章 数据收集 / 11

案例引入 / 11

第一节 统计数据的来源 / 12

- 一、统计数据的间接来源 / 12
- 二、统计数据的直接来源 / 13
- 三、练习题 / 18

第二节 统计调查方案 / 19

- 一、统计调查方案的内容 / 19
- 二、案例——我国住户收支与生活状况调查方案 / 22

第三节 问卷设计 / 24

- 一、问卷的含义 / 24
- 二、问卷的基本结构 / 24
- 三、问卷设计的步骤 / 25
- 四、问题的设计 / 25
- 五、问题答案的设计 / 26
- 六、案例——购物中心消费者意见调查问卷 / 26
- 七、练习题 / 29

案例解析 / 29

相关知识图示 / 30

第三章 数据整理与显示 / 31

案例引入 / 31

第一节 数据整理的程序 / 32

第二节 统计分组 / 33

- 一、统计分组的概念 / 33
- 二、品质数据分组 / 33
- 三、数值型数据分组 / 34
- 四、统计分组的形式 / 36

第三节 频数分布 / 37

- 一、频数分布的概念 / 37
- 二、频数分布的编制 / 37
- 三、频数分布的常见形式 / 39
- 四、Excel操作 / 40
- 五、练习题 / 44

第四节 统计表和统计图 / 45

- 一、统计表 / 45
- 二、统计图 / 46
- 三、Excel操作 / 50
- 四、练习题 / 52

案例解析 / 54

相关知识图示 / 54

第四章 数据特征的描述 / 55

案例引入 / 55

第一节 总量与相对量的测度 / 56

- 一、总量指标 / 56
- 二、相对指标 / 57
- 三、练习题 / 61

第二节 集中趋势的测度 / 62

- 一、集中趋势的含义 / 62
- 二、集中趋势的度量方法 / 62
- 三、Excel操作 / 70
- 四、练习题 / 73

第三节 离散程度的测度 / 75

- 一、离散程度的含义 / 75
- 二、离散程度的度量方法 / 75
- 三、数据的标准化 / 80
- 四、Excel操作 / 81
- 五、练习题 / 83

案例解析 / 84

相关知识图示 / 85

第五章 抽样估计 / 86

案例引入 / 86

第一节 抽样与抽样分布 / 87

一、抽样估计 / 87

二、抽样方法 / 87

三、抽样估计的基本概念 / 91

四、样本统计量的抽样分布 / 92

五、Excel操作 / 96

六、练习题 / 96

第二节 参数估计的方法 / 97

一、点估计 / 97

二、区间估计 / 98

三、Excel操作 / 102

四、练习题 / 106

第三节 必要样本容量的确定 / 108

一、影响必要样本容量的因素 / 108

二、简单随机抽样条件下样本容量的确定 / 108

三、练习题 / 109

案例解析 / 110

相关知识图示 / 112

第六章 假设检验 / 113

案例引入 / 113

第一节 假设检验的基本问题 / 114

- 一、假设检验的基本思想 / 114
- 二、假设检验的步骤 / 114
- 三、假设检验的两类错误 / 118
- 四、假设检验与参数估计的关系 / 119
- 五、练习题 / 120

第二节 单总体参数的假设检验 / 121

- 一、总体均值的假设检验 / 121
- 二、总体比例的假设检验 / 124
- 三、Excel操作 / 125
- 四、练习题 / 126

第三节 两总体参数的假设检验 / 127

- 一、两独立样本均值的抽样分布 / 127
- 二、两独立总体均值之差的假设检验 / 128
- 三、练习题 / 129

案例解析 / 129

相关知识图示 / 130

第七章 方差分析 / 131

案例引入 / 131

第一节 方差分析的基本问题 / 132

- 一、方差分析的产生背景与基本概念 / 132
- 二、方差分析的基本原理 / 133

三、练习题 / 134

第二节 单因素方差分析 / 134

- 一、单因素方差分析的数据结构 / 135
- 二、单因素方差分析的步骤 / 135
- 三、单因素方差分析中的多重比较 / 139
- 四、Excel操作 / 140
- 五、练习题 / 141

第三节 双因素方差分析 / 143

- 一、双因素方差分析应用的背景及类型 / 143
- 二、无交互作用的双因素方差分析 / 144
- 三、Excel操作 / 147
- 四、练习题 / 149

案例解析 / 149

相关知识图示 / 151

第八章 统计指数 / 152

案例引入 / 152

第一节 统计指数的含义、作用和分类 / 153

- 一、统计指数的含义 / 153
- 二、统计指数的作用 / 154
- 三、统计指数的分类 / 154
- 四、练习题 / 155

第二节 综合指数 / 155

- 一、综合指数的编制原理 / 155
- 二、综合指数的计算形式 / 156
- 三、Excel操作 / 158
- 四、练习题 / 159

第三节 平均指数 / 160

- 一、平均指数的编制原理 / 160
- 二、平均指数的计算形式 / 160
- 三、综合指数和平均指数的区别 / 163
- 四、Excel操作 / 163
- 五、练习题 / 164

第四节 指数因素分析法 / 165

- 一、指数体系 / 165
- 二、因素分析 / 165
- 三、总量指标两因素分析 / 166
- 四、总量指标多因素分析 / 167
- 五、平均指标两因素分析 / 168
- 六、练习题 / 168

案例解析 / 170

相关知识图示 / 172

第九章 相关与回归分析 / 173

案例引入 / 173

第一节 相关分析 / 174

- 一、相关关系的含义 / 174
- 二、相关关系的类型 / 174
- 三、相关关系的判断方法 / 176
- 四、Excel操作 / 177
- 五、练习题 / 178

第二节 一元线性回归分析 / 180

- 一、回归分析的含义 / 180
- 二、一元线性回归方程 / 180
- 三、参数的最小二乘估计 / 180
- 四、拟合优度与显著性检验 / 181
- 五、一元线性回归预测 / 184
- 六、Excel操作 / 185
- 七、练习题 / 186

第三节 多元线性回归分析 / 188

- 一、多元线性回归模型 / 188
- 二、多元线性回归方程的拟合优度检验 / 188
- 三、多元线性回归方程的显著性检验 / 189
- 四、练习题 / 189

案例解析 / 191

相关知识图示 / 192

第十章 时间序列分析与预测 / 193

案例引入 / 193

第一节 时间序列的描述 / 194

- 一、时间序列的含义 / 194
- 二、时间序列的影响因素 / 195
- 三、时间序列的图形描述 / 195
- 四、时间序列的指标描述 / 196
- 五、Excel操作 / 202
- 六、练习题 / 203

第二节 时间序列外推预测 / 204

- 一、时间序列预测方法的选择 / 204
- 二、时间序列预测的程序 / 205
- 三、移动平均预测法 / 206
- 四、指数平滑预测法 / 208
- 五、线性趋势外推法 / 209
- 六、Excel操作 / 211
- 七、练习题 / 213

第三节 季节变动预测 / 215

一、季节指数的计算	/ 215
二、季节指数预测	/ 218
三、Excel操作	/ 218
四、练习题	/ 219
案例解析	/ 220
相关知识图示	/ 221

附录 / 222

附表1 正态分布分位数表	/ 222
--------------	-------

附表2 t 分布表	/ 223
附表3 F 分布表	/ 224
附表4 随机数字表	/ 230
附表5 Excel中的统计函数	/ 230
附表6 Excel中的数学函数	/ 231

参考文献 / 232



知识目标

1. 统计的内涵
2. 统计的职能
3. 统计学的基本概念：总体、样本、个体、参数、统计量、变量



能力目标

1. 能正确理解统计的内涵
2. 能正确理解统计学在认识世界和管理决策中的作用
3. 能根据特定研究对象理解统计学中的基本概念



案例引入

统计助力 Tesco 实现公司价值

现在几乎每家超市、购物中心都会推出会员卡。顾客办理会员卡通常是由于受到会员折扣价的吸引，或者希望能够获得超市、购物中心定期的返利；商家则希望通过会员卡赢得顾客的忠诚。会员卡为商家累积了大量的数据，包括顾客的姓名、联系方式等会员信息和每一次的交易数据。商家如果能够对这些数据进行合理的分析和利用，这些数据将成为商家最重要的资产。Tesco（特易购）是英国一家大型超级连锁零售商，它在1995年启动了一个大规模的会员卡项目——Clubcard。Clubcard不是英国连锁零售业的第一个会员卡项目，但它是英国最成功的会员卡项目之一。为了达到以数据为驱动进行市场营销的目标，Tesco在Clubcard项目启动之前就确定了一个专长于数据分析的合作伙伴，该合作伙伴是善于从业务角度出发利用数据的Dunnhumby公司。Tesco和Dunnhumby之间十余年的合作为双方带来了巨大的价值。Tesco宣称，分析和利用Clubcard的数据使公司每年因不囤积销量不好的产品而节省3.5亿英镑。Dunnhumby由当初合作时的一小公司成长为目前世界知名的市场调查公司。当年在启动Clubcard项目时，Tesco的市场份额就超越了它的最大竞争对手Sainsbury，目前Tesco已经成为英国最大、世界第三大的食品零售商，排在沃尔玛和家乐福之后。2004年7月，特易购进入中国市场。

请思考以下问题。

1. Dunnhumby是怎样借助统计工具和先进的计算机技术对会员卡信息和交易数据进行分析，并帮助Tesco获得成功的呢？
2. 统计的应用领域极为广泛，统计方法也极为丰富，面对纷繁复杂的现实数据，您是否想探个究竟，一显身手？

第一节

统计的内涵

“统计”一词有三种含义：统计学、统计工作和统计资料。统计学是一系列统计方法的理论阐述；统计工作是统计的具体实践活动；统计资料是统计工作各阶段的成果，也可称为统计数据。

一、统计学

统计学是收集、整理、分析和解释统计数据的科学，是一门方法论性质的学科。其目的是探索研究对象的数量特征和数据的内在规律。

统计学的这一定义包含了以下3个要点。

(1) 统计学是研究“数据”的科学，离开了数据，统计学也就失去了它存在的意义，因此有人称统计学是“数据的科学”。

(2) 统计学是方法论科学，是研究数据的“工具”，因此它适用于所有有数据存在的学科领域。

(3) 统计学研究的不是抽象的数据，而是“有载体的统计数据”，因此利用统计方法得到的任何数据特征和数据规律都与某一研究对象紧密相连。

统计学的研究对象是现象总体的数量方面。统计学通过对总体中大量个体的差异化数据进行收集、整理和分析，获得研究对象总体的特征和规律。

英文“Statistics”一词包含了两层含义：一是作为一门学科的统计学，二是统计数据或统计资料。这也说明了统计学与统计数据之间密不可分的关系。

二、统计工作

一般将统计工作概括地分为4个阶段，即统计设计、数据收集、数据整理和数据分析，如图1-1所示。

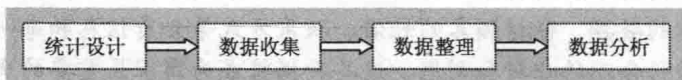


图 1-1 统计工作过程

1. 统计设计

统计设计是根据统计研究的目的和研究对象的特点，对统计工作的各个方面及各个环节所做的通盘考虑与安排。其基本任务是制定各种统计工作方案，主要内容应包括统计指标体系、统计调查方案、统计汇总或整理方案以及统计分析方案等诸多方面。统计设计是统计工作实施的基本依据，是使统计工作协调、有序、顺利进行的必要条件。

2. 数据收集

数据收集是根据统计设计的要求，采用科学的数据收集方法，收集总体中全部或部分个体数据资料的工作过程。数据收集是统计工作的基础环节，是认识事物的起点。原始数据质量的高低，直接影响到分析结论是否可靠。因此，数据收集阶段应尽可能地降低统计调查误差。

3. 数据整理

数据整理是对收集来的统计数据进行科学的加工整理，使之系统化、条理化，把大量反映个体特征的零散资料转化为反映总体综合数量特征的统计资料的过程。数据整理是数据收集和数据分析之间承上启下的中间环节，其主要任务是对收集的数据进行分组、归类，并用合适的统计表或统计图展示整理的结果。

4. 数据分析

数据分析是运用统计方法及与分析对象有关的知识，定量与定性地对研究对象总体进行分析，以揭示研究对象数量特征和数量规律的统计工作阶段。数据分析是统计工作的最后一环，也是关键的一步。统计分析报告是数据分析的结果，其质量既取决于数据本身的质量，又取决于统计分析人

员对统计分析利用的程度。

一般来说，统计工作的 4 个阶段是依次进行的，任何一个阶段的工作失误都会影响到下一阶段的工作质量。在某些情况下，为保证统计工作整体取得好的效果，各阶段也会交叉进行。例如，统计人员在数据收集和整理阶段可以进行一些必要的分析，或者对原设计方案进行适当的改进；在数据分析阶段，如果现有资料不能满足分析的需要，可以做一些必要的补充调查、数据整理等工作。

三、统计数据

统计数据是对现象进行计量和分析的结果，是研究对象的特征表现。对个体而言，其特征既可以用数字表示，又可以用文字表示；对总体而言，其特征都是用数字表示的。

一个完整的统计数据应包括 3 个最基本的构成要素：数据名称、数值和数据的计量单位。在具体表述一个统计数据时，很多时候只有这 3 个要素是不够的，还应包含对数据其他方面的内涵规定，如数据所属的时间和空间范围等。

例如，“工业机器人 95 351 台”传递的信息是不完整的，如果改为“2017 年前三季度，我国工业机器人有 95 351 台”，数据的内外边界就很清楚了。

统计数据从不同的角度有不同的分类，通常可以从以下 3 个角度进行分类，如图 1-2 所示。

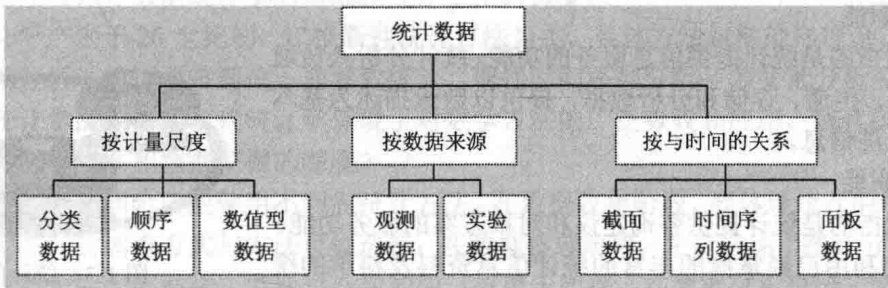


图 1-2 统计数据的分类

1. 按计量尺度分类

按计量尺度的不同，统计数据分为分类数据、顺序数据和数值型数据。

(1) 分类数据用文字来表述，表明事物的不同属性或类别。在进行统计处理时，我们用数字代码来表示分类数据的各个类别。例如，用“1”代表“男性”，用“2”代表“女性”，这里的“1”和“2”只表达了现象分属于两个类别这样一层含义。企业所属的不同行业、企业的不同所有制类型、不同职业等都属于分类数据。

(2) 顺序数据用文字来表述，表明有顺序的不同类别。在进行统计处理时，我们同样可以用数字代码来表示。例如，用“1”“2”“3”“4”“5”分别代表考试成绩“不及格”“及格”“中”“良”“优”，这里的“1”“2”“3”“4”“5”不仅表明现象分属于不同的类别，而且表明这些类别的顺序。不同的教育程度、产品的不同等级等均属于顺序数据。

(3) 数值型数据用数字来表现各种现象的数量特征，是使用自然或度量衡单位对事物进行计量的结果。一个企业的职工人数、产值、销售收入、市场占有率等都是用具体的数字来表示的。数值型数据不仅能表明现象数量上的不同和大小顺序，数据之间还能进行数量运算。大部分统计分析方法适用于数值型数据，统计处理的大多是数值型数据。

分类数据和顺序数据也称为品质数据或定性数据，数值型数据也称为数量数据或定量数据。

2. 按数据来源分类

按来源不同，统计数据可分为观测数据和实验数据。

(1) 观测数据是指通过直接调查或观察收集的数据。社会经济领域的统计数据基本上是观测数据。

(2) 实验数据是通过对实验对象、实验环境及实验过程进行有效控制而获得的统计数据。这些数据主要用于考察变量之间的因果关系。

3. 按与时间的关系分类

按与时间的关系不同，统计数据可分为截面数据、时间序列数据和面板数据。

(1) 截面数据是指某一总体中的不同个体在相同或近似相同时间上的数据表现。

(2) 时间序列数据是指将不同时间上的某一指标值列出所形成的按时间顺序排列的数据序列。

(3) 面板数据是指总体中的不同个体在不同时间上的数据表现，是时间序列数据与截面数据的结合。

第二节

统计的职能

统计的职能可以概括为信息职能、咨询职能和监督职能，如图 1-3 所示。

1. 信息职能

信息职能指的是统计提供信息服务的功能。统计分析人员通过采集、处理、传递、存储和分析数据，提供以数量描述为基本特征的社会经济信息。

2. 咨询职能

咨询职能指的是统计提供咨询建议和对策方案的服务功能。统计分析人员利用已经掌握的丰富的统计信息资料及科学的统计分析方法、先进的技术手段，开展综合分析和专题分析，为科学决策和科学管理提供可选择的咨询建议和对策方案。

3. 监督职能

监督职能指的是统计揭示事物运行中的偏差，促使事物不偏离正常轨道运行的功能。统计以定量检查、监测、预警指标体系等为手段，揭示决策及其执行过程中的偏差，使决策及其执行过程按客观规律的要求进行。

信息职能是统计最基本的职能，是保证咨询职能和监督职能得以有效发挥的前提。反过来，统计咨询职能和监督职能的强化会促进统计信息职能的强化。统计的 3 种职能相辅相成，相互作用，构成了一个有机整体，故称为统计的整体功能。

统计方法已应用到人类生活的各个领域。生活在 21 世纪的人们，将与数据朝夕相处——收集相关数据，从数据中提取信息，进行预测与决策，将是个人、企业、政府经常面对的事情。在科研领域，统计方法已应用到几乎所有研究领域。各学科广泛地应用统计方法，使得其拥有各自的统计术语，如生物统计学、医学统计学、卫生统计学、商务统计学、经济统计学、统计物理学、人口统计学、心理统计学、教育统计学、社会统计学、体育统计学等。

统计在工商业领域扮演着重要的角色。一个工商管理人员会经常面临大量企业经济管理方面的数据。这些数据只有被分析、提炼，才能成为管理人员进行管理和决策的依据。

图 1-4 简要表述了当我们遇到一个具体问题时，借助于统计工具解决问题的思维过程。



图 1-3 统计的职能



图 1-4 应用统计的思维过程

我们在讨论统计所具有的职能、统计应用的广泛性及统计对决策的重要性时，还应看到统计的局限性。统计可以帮助人们分析数据，并通过分析得出某种结论，但对统计结论的进一步解释则需要用到相关领域的专业知识。



微课：统计的源流



微课：统计与生活

第三节

统计学研究方法

统计学于 17 世纪产生，发展到今天，其应用领域越来越广泛，其方法越来越丰富。统计学的方法可以简单地划分为两大类：描述性统计方法和推断性统计方法。

传统统计学以描述性统计方法为主，主要包括数据的收集、整理、图表显示和数据的综合测度等内容。描述性统计方法是统计学的基础和统计工作的初步。

现代统计学产生于 20 世纪初，以推断性统计方法为主，具体方法包括抽样理论、参数估计、假设检验、方差分析、统计决策理论、非参数统计、现代时间序列分析、多变量分析等。推断性统计方法是现代统计学的核心。现代统计学体现了对数学方法的广泛吸收和应用，现代统计方法的应用提高了数据分析的效率和数据挖掘的深度。

在统计工作的不同阶段应使用不同的统计方法。在数据收集阶段，要体现统计的“大量观察法”，具体应用时有不同的调查方式与方法可供选择；在数据整理阶段，主要应用“统计分组法”，通过分组、归类了解数据的内部结构，并通过合适的图、表展示数据；在数据分析阶段，根据分析问题的角度和深度不同，也有大量的统计分析方法可供选择。随着统计学对更多数学方法的吸收和先进信息技术的不断渗透，更有效的统计分析方法在不断产生。

本书将按照统计工作的不同阶段，分别介绍数据收集的方式与方法（第二章）、数据整理与显示方法（第三章）和数据分析方法（第四章至第十章）等。

统计方法是否有效取决于研究对象是否满足方法的适用条件或基本假设。误用统计学的方法可能会导致描述上的偏离或者推论的错误，从而可能导致决策的失误。因此，统计方法的使用者应正确使用统计方法，避免误用。

第四节

统计学中的基本概念

一、总体、样本与个体

总体是客观存在的，由性质相同的大量个体组成，是由统计研究目的决定的统计研究对象的全体。个体是组成总体的个别单位。例如，要研究某地区国有企业的生产经营状况，则该地区的全部国有企业构成总体，某个国有企业就是个体；某企业要检查某批产品的质量，则该批产品构成总体，

该批产品中的某件产品就是个体。

样本是从总体中抽取出来的一部分个体组成的整体。抽样的目的是用样本的数据特征推断总体的数据特征。例如，为了研究某产品用户的满意度，从该产品的全部用户中随机抽取 5% 构成样本，用这部分用户的满意度来估计该产品用户的满意度。总体、样本与个体的关系如图 1-5 所示。

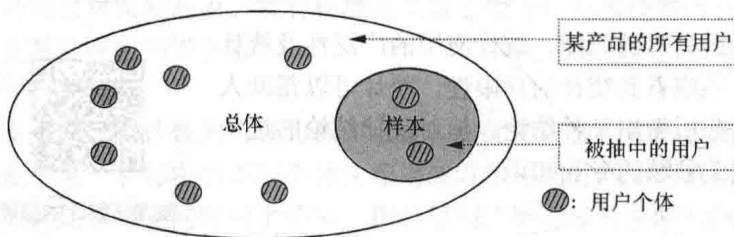


图 1-5 总体、样本和个体

二、参数与统计量

参数是描述总体综合数量特征的概括性数字度量，是对总体中所有个体某一数量特征的综合。

统计量是描述样本综合数量特征的概括性数字度量，是对样本中所有个体某一数量特征的综合。计算统计量是为了估计参数。

不管是参数还是统计量，体现的都是对个体数量特征的综合，都用数字表示。例如，随机抽取 5% 的用户进行满意度调查，5% 的用户的满意度得分是 82 分，这就是统计量；据此推算出全部用户的满意度得分为 79~85 分，这就是参数。

习惯上，也将总体的综合数量特征称为统计指标。

三、变量

变量用于描述个体特征，其具体取值称为变量值。习惯上，也将反映个体特征的变量称为标志。

1. 按变量的表现形式分类

变量按其表现形式的不同，可分为品质变量和数值型变量。

(1) 品质变量用文字表示，包括分类变量和顺序变量。分类变量是说明事物类别的一个名称，其取值是分类数据。例如，企业的“行业”属性是分类变量，其变量值表现为“IT 业”“物流业”“旅游业”等不同类别的行业。顺序变量是说明事物有序类别的一个名称，其取值是顺序数据。如“服务等级”是顺序变量，其变量值表现为“优”“良”“中”“差”4 个顺序等级。

(2) 数值型变量是说明事物数字特征的一个名称，其取值是数值型数据。例如，企业的“营业额”是数值型变量，变量值为具体的数值，如“20 万元”“30 万元”“40 万元”等。

以商场的会员卡为例，每个会员的个体信息就是变量，如某会员的性别是一个分类变量，教育程度是一个顺序变量，购买金额是一个数值型变量。

2. 按变量的取值不同分类

数值型变量根据变量的取值不同，分为离散型变量和连续型变量。

(1) 离散型变量是只能取整数值的变量，一般用来反映以自然整数计量的数量表现，变量值之间以整数位断开。例如，企业从业人数 1 000 人只能用整数来计量，在 999 和 1 000 两个整数之间不可能有小数值。

(2) 连续型变量是可以在一个或多个区间内取任意值的变量，其取值是连续不断的，不能一一列举。例如，销售收入 20 万元以货币单位计量，在顺序的两个“万元”之间可以有小数值，表示

比万元更小的计量单位上的取值。

四、练习题

1. 选择题

(1) 利用统计方法认识研究对象包括的统计活动有()。

- A. 解释数据 B. 收集数据 C. 分析数据 D. 整理数据

(2) 下列属于某一有序类别的非数值型数据的是()。

- A. 支付方式(现金、支票、刷卡) B. 购物金额
C. 企业规模(大、中、小) D. 年龄

(3) 一个统计数据的构成要素有()。

- A. 名称 B. 数值 C. 时间范围 D. 计量单位

(4) 在不同时间上收集到的数据是()。

- A. 观测数据 B. 实验数据 C. 截面数据 D. 时间序列数据

(5) 下列用于描述总体特征值的是()。

- A. 统计量 B. 变量 C. 参数 D. 变量值

(6) 下列属于离散型变量的是()。

- A. 某企业职工总人数 B. 城乡居民储蓄存款余额
C. 职工的月收入 D. 全国城镇居民家庭总户数

2. 思考题

(1) 什么是统计学? 举出几个统计应用的例子。

(2) 简述统计工作的流程。

(3) 统计数据可分为哪几种? 举例说明各类数据的特点。

3. 综合应用题

(1) 判断下列数据属于分类数据、顺序数据还是数值型数据: 品牌、职业、产品的满意度(满意、较满意、一般、不太满意、不满意)、考试成绩、市场占有率、流动资金占用额、学历、购物方式、月收入。

(2) 某市的城市抽样调查队随机抽取了 1 000 户居民作为固定样本, 记录其每月的消费支出情况, 连续记录了 12 个月, 并对这 1 000 户居民 12 个月的每月消费总额及消费构成进行了汇总和分析, 以此估计该市居民的消费支出情况。要求如下。

① 指出该调查的总体、样本、个体及参数、统计量、变量。

② 某户居民每月支出额是分类变量、顺序变量还是数值型变量?

③ 某户居民每月食品支出额是离散型变量还是连续型变量?

④ 某月 1 000 户居民的支出额和 12 个月 1 000 户居民的平均支出额中, 哪个是时间序列数据? 哪个是截面数据?

⑤ 上述数据是观测数据还是实验数据?

(3) 某校组织了一次学生消费水平的调查。该校共有在校生 5 000 人, 随机调查了 800 人。要求如下。

① 指出该项调查的总体、总体单位、样本、样本单位。

② 针对该项调查, 设计调查内容, 并举例说明什么是变量、统计量、参数, 什么是离散型变量和连续型变量。

第五节

统计应用软件简介

从古人在树木上刻痕以计算家畜的数量，到今天通过计算机软件来收集、储存、分析数据，人类处理和分析数据的速度和规模已发生了革命性的变化。人类正行驶在以数据为载体的信息高速公路上。在人们的日常生活、政府行政事务管理、企业生产经营管理、科学研究等各个领域，手工计算已被计算器、计算机所取代，数据分析软件可帮助人们快速、高效地完成数据分析任务。

在统计学的学习中，众多公式、符号及数字令很多人望而生畏，感到枯燥乏味；同时，数据量较大、较为复杂的运算，也难以手工进行。近年来，借助于计算机和数据处理软件，统计教学和统计学习变得更高效了。

目前，可用于统计数据处理和分析的软件很多。办公自动化软件 Excel 就具有强大的数据处理功能，其应用非常普遍。专业的统计软件有 SPSS、SAS、Minitab、Statistica、Eviews 等。本书的适用对象是非统计学专业的学生，不涉及很复杂高深的模型与运算，因此选择具有较高普及率的 Excel 软件作为统计入门学习和掌握常用数据处理方法的工具。

Excel 具有强大的表格格式化功能、计算和函数功能、图表制作功能等，分析人员可以利用 Excel 的相应命令来制作图表（见图 1-6），进行数据的分类汇总（见图 1-7），利用统计函数或相应的数据分析工具完成数据的运算或分析等（见图 1-8~图 1-10）。Excel 具有功能强大和使用方便的特点，广泛应用于会计、财务、金融、营销、贸易、统计、行政等领域，是比较适合非统计学专业经济管理人员使用的应用统计软件。

本书中有关运算的任务的实施都借助于 Excel 来完成，很少涉及手工计算，相应的公式推导、手工计算简捷公式全部省略，读者可将注意力放在统计方法的基本原理、Excel 的操作步骤和统计结果的解读上。



图 1-6 插入图表



图 1-7 数据的分类汇总