



工商管理硕士（MBA）系列规划教材



# 统 计 学

龚江辉 编著

▶ Statistics



北京师范大学出版集团  
BEIJING NORMAL UNIVERSITY PUBLISHING GROUP  
北京师范大学出版社



工商管理硕士（MBA）系列规划教材



文森大学堂

[www.jswsedu.com](http://www.jswsedu.com)

立体化教材

# 统计学

龚江辉 编著

Statistics



北京师范大学出版集团

BEIJING NORMAL UNIVERSITY PUBLISHING GROUP

北京师范大学出版社

---

### 图书在版编目(CIP)数据

统计学/龚江辉编著. —北京:北京师范大学出版社, 2018. 7  
(工商管理硕士(MBA)系列规划教材)  
ISBN 978-7-303-23660-2

I. ①统… II. ①龚… III. ①统计学—研究生—教材  
IV. ①C8

中国版本图书馆 CIP 数据核字(2018)第 084174 号

---

营销中心电话 010-62978190 62979006  
北师大出版社科技与经管分社 www.jswsbook.com  
电子信箱 jswsbook@163.com

---

TONGJIXUE

出版发行: 北京师范大学出版社 www.bnup.com  
北京市海淀区新街口外大街 19 号  
邮政编码: 100875

印 刷: 北京玺诚印务有限公司  
经 销: 全国新华书店  
开 本: 730 mm×980 mm 1/16  
印 张: 18.5  
字 数: 306 千字  
版 次: 2018 年 7 月第 1 版  
印 次: 2018 年 7 月第 1 次印刷  
定 价: 48.80 元

---

策划编辑: 张自然 责任编辑: 张自然  
美术编辑: 刘 超 装帧设计: 刘 超  
责任校对: 赵非非 黄 华 责任印制: 赵非非

---

### 版权所有 侵权必究

反盗版、侵权举报电话: 010-62978190  
北京读者服务部电话: 010-62979006-8021  
外埠邮购电话: 010-62978190  
本书如有印装质量问题, 请与印制管理部联系调换。  
印制管理部电话: 010-62979006-8006

## 前

## 言

本教材适合于普通高校经济学或管理学专业本科生和硕士生使用,也可用于其他文科类专业的本科及硕士统计学教学。

作为一本主要面向文科背景本科生使用的教材,本书编写的指导思想包括两个方面:降低数学门槛,提高实用性。

与许多高校目前广泛使用的统计学教材相比,本教材极大地简化了有关概率论与数理统计的内容,改变了统计学教材中一半篇幅充斥数学公式的状况。这有助于降低文科背景学生对于统计学课程的畏惧心理,使其能够带着轻松的心态来学习这门课程。

在高校经管专业的培养方案中,概率论与数理统计是作为统计学原理课程的先修课开设的,这就决定了“统计学”与“概率论与数理统计”这两门课程应当各有侧重。概率论与数理统计主要是为统计学提供数学基础,而统计学则主要是面向于经管专业的工作和科研实践需要。因为学生在学习统计学课程之前已经修读过概率论与数理统计课程,统计学教材完全没有必要重复介绍有关概率论的内容,涉及推断统计的部分,也应当局限于对方法的介绍,而不必纠缠于原理的证明。

本教材的内容和编排体系经历过20年课堂教学的实践检验,证明能够让学生较好地掌握包括推断统计在内的统计学内容,即使是没有修读过“概率论与数理统计”先修课的文科背景学生,按照本教材中的介绍也能够较为容易地理解参数估计与假设检验的意义和方法。

为了照顾学有余力的学生,本教材对于部分涉及数学原理

的内容采取了穿插“知识点”的做法，将一些严谨的数学定义放在单独的知识点中，数学基础较好的学生可以参考这部分内容，数学基础不足的学生即使略过这些内容，也不会影响到对全书的学习。

在实用性方面，本教材在介绍各种统计方法时均强调其在实践中的应用方式，并提供了大量案例予以说明，其中一些比较详细的案例包括第2章中“在一个城市调查1000个样本的抽样设计”，第3章中“各省区城市设施水平综合评价”“歌手大奖赛分值的转换”，第6章中“市场调查中确定样本量的方法”等。

统计更多的是一种认识世界的工具，在统计实践中，研究人员屡屡打破陈规，不拘泥于某种定式，创造性地应用统计方法。例如，从严格的意义上说，Likert量表的测量结果属于顺序尺度数据，不能进行加减乘除的运算，但大多数研究人员都会使用Likert量表数据进行因子分析、回归分析等面向间隔尺度或比率尺度数据的统计计算。本教材的编写考虑到了这种情况，在介绍各种统计方法与概念时，对于实践中的变通手段也进行了讨论，以帮助学生更好地掌握实践技能。

本教材还有一个特点就是对抽样调查技术和方法给予了较大的篇幅，涉及有关各类抽样方式的介绍、非抽样误差等内容。介绍这些知识的原因在于抽样调查是实践中获得统计数据的主要手段，通过对抽样调查知识的学习，学生不仅能够掌握数据收集的方法，而且能够对数据的随机性以及调查数据质量等问题有更为深刻的体会。

随着计算机的普及，传统的依赖统计表查阅统计分布函数值的做法已经被抛弃，本教材没有提供经典的统计函数表，而是在讨论各种分布时分别介绍了如何利用Excel软件和R语言计算统计函数值的方法。

同时，本教材配置了各章课件及书中演算所需的各章的数据库，还有部分思考与练习题答案。注册登录后通过扫描书中各章二维码阅读课件；通过扫描第一章二维码，下载解压缩Excel表格即可使用数据库；扫描部分思考与练习题参考答案处的二维码即可看到答案PDF版。

本教材编写过程中参考了许多前辈和同行的相关教材和著作，从中汲取了大量宝贵的思想，恕不能一一致谢。由于作者才疏学浅，本教材难免会有错漏，敬请广大使用者给予批评。

龚江辉  
于北京师范大学  
2018.2.13



北京师范大学出版集团

BEIJING NORMAL UNIVERSITY PUBLISHING GROUP

北京师范大学出版社科技与经管分社

地址：北京市海淀区信息路甲 28 号科实大厦 C 座 12B

电话：010-62979096/8896 传真：010-62978190

网址：jswsbook.com 邮箱：jswsbook@163.com



官方微信公众号



官方微博

## 教师样书申请表

尊敬的老师，您好！

请您在我社网站的最新教材目录中选择与您教学相关的样书(每位教师每学期限选 1-2 种)，并以清晰的字迹真实、完整填写下列栏目后经所在院(系)的主要负责人签字或盖章。符合上述要求的表格将作为我社向您提供免费教材样书的依据。本表复制有效，可传真或函寄，亦可发 E-mail。

姓名：\_\_\_\_\_ 性别：\_\_\_\_\_ 年龄：\_\_\_\_\_ 职务：\_\_\_\_\_ 职称：\_\_\_\_\_

院校名称：\_\_\_\_\_ 大学(学院) \_\_\_\_\_ 学院(系) \_\_\_\_\_ 教研室 \_\_\_\_\_

通信地址：\_\_\_\_\_

邮编：\_\_\_\_\_ 座机：\_\_\_\_\_ 手机：\_\_\_\_\_

E-mail：\_\_\_\_\_ 微信：\_\_\_\_\_ QQ：\_\_\_\_\_

教授课程 \_\_\_\_\_ 学生层次 \_\_\_\_\_ 学生人数/年 \_\_\_\_\_ 用书时间 \_\_\_\_\_

\_\_\_\_\_ ☐ 研究生 ☐ 本科 ☐ 高职 \_\_\_\_\_ ☐ 春季 ☐ 秋季

现使用教材 \_\_\_\_\_ 版本 \_\_\_\_\_ 换教材意向 \_\_\_\_\_

\_\_\_\_\_ 出版社 ☐ 有 ☐ 无

换教材原因 \_\_\_\_\_

课程 \_\_\_\_\_

原因 \_\_\_\_\_

曾编教材情况

| 书 名 | 出 版 社 | 主编/副主编/参编 | 出版时间 |
|-----|-------|-----------|------|
|     |       |           |      |

教材编写意向： ☐ 近期有编写意向 ☐ 目前暂无意向

希望编写教材名称：\_\_\_\_\_

申请样书

| 书 名 | 书号 (ISBN) | 作 者 | 定 价 |
|-----|-----------|-----|-----|
|     |           |     |     |
|     |           |     |     |

# 目 录

|                             |      |
|-----------------------------|------|
| <b>第 1 章 绪 论</b> .....      | (1)  |
| 【学习目标】 .....                | (1)  |
| 【关键术语】 .....                | (1)  |
| 【本章要点】 .....                | (1)  |
| 1.1 统计学简介 .....             | (1)  |
| 【案例 1-1】如何统计一个水库里的鱼 .....   | (2)  |
| 1.2 统计学的学科构成 .....          | (3)  |
| 1.2.1 学术角度的学科构成 .....       | (3)  |
| 1.2.2 统计学习角度的学科构成 .....     | (5)  |
| 1.3 统计学科简史 .....            | (7)  |
| 1.3.1 统计学的起源 .....          | (7)  |
| 1.3.2 现代统计学的形成 .....        | (8)  |
| 1.4 统计学的课程学习方法 .....        | (8)  |
| 1.4.1 形成运用统计思想进行分析的能力 ..... | (9)  |
| 1.4.2 掌握基本的统计方法 .....       | (9)  |
| 1.4.3 培养进行统计计算的能力 .....     | (10) |
| 【思考与练习】 .....               | (13) |
| <b>第 2 章 统计调查</b> .....     | (14) |
| 【学习目标】 .....                | (14) |
| 【关键术语】 .....                | (14) |

|                                     |      |
|-------------------------------------|------|
| <b>【本章要点】</b>                       | (14) |
| 2.1 统计指标                            | (15) |
| 2.1.1 总体与个体                         | (15) |
| <b>【案例 2-1】</b> 农民工总体的定义            | (15) |
| 2.1.2 标志与指标                         | (16) |
| 2.1.3 统计尺度                          | (17) |
| <b>【案例 2-2】</b> 李克特量表               | (17) |
| <b>【案例 2-3】</b> 比率尺度与间隔尺度的对比        | (19) |
| 2.1.4 指标的六要素                        | (20) |
| 2.2 统计调查的分类                         | (22) |
| 2.2.1 按选择样本的范围分类                    | (22) |
| <b>【案例 2-4】</b> 典型调查案例：毛泽东《寻乌调查》    | (24) |
| 2.2.2 按获取信息的方法分类                    | (25) |
| 2.3 抽样调查                            | (27) |
| 2.3.1 概率抽样                          | (27) |
| <b>【案例 2-5】</b> 对比分层抽样与简单随机抽样       | (31) |
| <b>【案例 2-6】</b> 群内差异大、群间差异小的抽样情境    | (36) |
| 2.3.2 概率抽样案例                        | (37) |
| 2.3.3 非概率抽样                         | (44) |
| <b>【案例 2-7】</b> 滚雪球式抽样案例            | (45) |
| 2.4 调查误差                            | (46) |
| 2.4.1 抽样误差                          | (46) |
| 2.4.2 非抽样误差                         | (46) |
| <b>【案例 2-8】</b> Rugg 进行的关于问卷措辞的对比实验 | (50) |
| 2.4.3 敏感性问题调查                       | (51) |
| <b>【思考与练习】</b>                      | (53) |
| <b>第 3 章 描述统计</b>                   | (55) |
| <b>【学习目标】</b>                       | (55) |
| <b>【关键术语】</b>                       | (55) |
| <b>【本章要点】</b>                       | (55) |
| 3.1 数据预处理                           | (56) |

|       |                                   |      |
|-------|-----------------------------------|------|
| 3.1.1 | 缺失值处理 .....                       | (56) |
| 3.1.2 | 数据分组 .....                        | (59) |
|       | 【案例 3-1】组中值的确定 .....              | (62) |
| 3.1.3 | 数据分布 .....                        | (63) |
|       | 【案例 3-2】在一副扑克牌中抽取 10 张形成的分布 ..... | (64) |
|       | 【案例 3-3】多峰分布 .....                | (69) |
| 3.2   | 相对指标 .....                        | (70) |
| 3.2.1 | 总量指标和相对指标 .....                   | (70) |
| 3.2.2 | 描述总量指标和相对指标的一些常用术语 .....          | (71) |
| 3.3   | 平均指标 .....                        | (72) |
| 3.3.1 | 算术平均数 .....                       | (72) |
|       | 【案例 3-4】简单算术平均数的计算 .....          | (73) |
|       | 【案例 3-5】加权算术平均数的计算 .....          | (73) |
| 3.3.2 | 调和平均数 .....                       | (75) |
|       | 【案例 3-6】调和平均数的计算 .....            | (75) |
| 3.3.3 | 几何平均数的计算 .....                    | (76) |
|       | 【案例 3-7】几何平均数 .....               | (76) |
|       | 【案例 3-8】算术平均数并不具有天然的优先权 .....     | (78) |
| 3.3.4 | 众数 .....                          | (79) |
|       | 【案例 3-9】分组数据的众数 .....             | (79) |
| 3.3.5 | 中位数 .....                         | (80) |
|       | 【案例 3-10】分组数据的中位数 .....           | (81) |
| 3.3.6 | 分位数 .....                         | (81) |
|       | 【案例 3-11】分位数的计算 .....             | (81) |
| 3.3.7 | 截尾均值 .....                        | (82) |
|       | 【案例 3-12】确定极端值的边界 .....           | (83) |
| 3.4   | 离散程度指标 .....                      | (83) |
|       | 【案例 3-13】离散程度的对比 .....            | (83) |
| 3.4.1 | 极差 .....                          | (83) |
| 3.4.2 | 内距 .....                          | (84) |
| 3.4.3 | 平均绝对差 .....                       | (84) |
|       | 【案例 3-14】平均绝对差的计算 .....           | (84) |

|                                                  |              |
|--------------------------------------------------|--------------|
| 3.4.4 方差和标准差 .....                               | (85)         |
| 【案例 3-15】方差的计算 .....                             | (85)         |
| 3.4.5 离散系数 .....                                 | (85)         |
| 【案例 3-16】离散系数的计算 .....                           | (86)         |
| 3.4.6 异众比率 .....                                 | (86)         |
| 【案例 3-17】异众比率的计算 .....                           | (86)         |
| 3.5 偏态和峰态 .....                                  | (87)         |
| 3.5.1 偏态及偏态系数 .....                              | (87)         |
| 3.5.2 峰态及峰态系数 .....                              | (88)         |
| 3.6 数据变换 .....                                   | (89)         |
| 3.6.1 线性变换 .....                                 | (90)         |
| 【案例 3-18】中心化变换 .....                             | (91)         |
| 【案例 3-19】数据变换的应用：各省区城市设施水平综合评价 .....             | (92)         |
| 3.6.2 非线性变换 .....                                | (96)         |
| 【案例 3-20】歌手大奖赛分值的转换 .....                        | (98)         |
| 3.7 箱线图 .....                                    | (101)        |
| 3.7.1 次序统计量的使用 .....                             | (101)        |
| 【案例 3-21】利用中国 2013 年 15 个省区市制造业人均工资计算次序统计量 ..... | (104)        |
| 3.7.2 箱线图 .....                                  | (105)        |
| 【思考与练习】 .....                                    | (109)        |
| <b>第 4 章 统计指数 .....</b>                          | <b>(112)</b> |
| 【学习目标】 .....                                     | (112)        |
| 【关键术语】 .....                                     | (112)        |
| 【本章要点】 .....                                     | (112)        |
| 4.1 统计指数概念 .....                                 | (112)        |
| 4.1.1 个体指数的计算 .....                              | (113)        |
| 4.1.2 总指数的计算 .....                               | (113)        |
| 4.2 综合指数的计算 .....                                | (115)        |
| 4.2.1 综合指数模型的导出 .....                            | (115)        |
| 【案例 4-1】拉氏指数与帕氏指数的计算 .....                       | (116)        |

|                                    |              |
|------------------------------------|--------------|
| 4.2.2 综合指数模型的推广 .....              | (117)        |
| 【案例 4-2】分析一个城市的机动车保有量增长了多大比例 ..... | (118)        |
| 4.3 平均指数的计算 .....                  | (119)        |
| 4.3.1 平均指数模型 .....                 | (119)        |
| 4.3.2 平均指数模型的应用 .....              | (120)        |
| 【案例 4-3】商品零售价格指数的计算模型 .....        | (120)        |
| 4.4 指数因素分析 .....                   | (123)        |
| 4.4.1 个体指数因素分析 .....               | (123)        |
| 【案例 4-4】个体指数因素分析 .....             | (124)        |
| 4.4.2 双因素综合指数因素分析 .....            | (124)        |
| 【案例 4-5】双因素综合指数因素分析 .....          | (124)        |
| 4.4.3 多因素综合指数因素分析 .....            | (126)        |
| 【案例 4-6】多因素综合指数因素分析 .....          | (126)        |
| 4.5 总平均数指数因素分析 .....               | (127)        |
| 4.5.1 概念 .....                     | (127)        |
| 【案例 4-7】某城市平均房价指数 .....            | (127)        |
| 4.5.2 影响总平均数指数的因素 .....            | (128)        |
| 4.5.3 总平均数指数的分解 .....              | (129)        |
| 4.6 指数修正 .....                     | (129)        |
| 4.6.1 指数偏误与费雪理想公式 .....            | (129)        |
| 4.6.2 购买力平价法 .....                 | (130)        |
| 【思考与练习】 .....                      | (132)        |
| <b>第 5 章 时间序列分析</b> .....          | <b>(135)</b> |
| 【学习目标】 .....                       | (135)        |
| 【关键术语】 .....                       | (135)        |
| 【本章要点】 .....                       | (135)        |
| 5.1 时间序列的概念与构成 .....               | (135)        |
| 5.1.1 时间序列的概念 .....                | (135)        |
| 【案例 5-1】把国民总收入视为时间的函数 .....        | (136)        |
| 5.1.2 时间序列的构成 .....                | (137)        |
| 【案例 5-2】某景区接待游客人数的季节变动 .....       | (139)        |

|                                   |              |
|-----------------------------------|--------------|
| 5.1.3 时间序列的表现形式 .....             | (141)        |
| 【案例 5-3】加法模型的应用 .....             | (142)        |
| 【案例 5-4】乘法模型的应用 .....             | (142)        |
| 5.2 趋势变动的测定 .....                 | (143)        |
| 5.2.1 修匀方法与拟合方法 .....             | (143)        |
| 【案例 5-5】油料作物产量 .....              | (143)        |
| 5.2.2 移动平均法 .....                 | (145)        |
| 5.2.3 曲线回归法 .....                 | (149)        |
| 5.3 季节变动的测定 .....                 | (154)        |
| 5.3.1 季节指数的概念 .....               | (154)        |
| 【案例 5-6】季节变动情况 .....              | (154)        |
| 5.3.2 按月(季)平均法 .....              | (154)        |
| 5.3.3 趋势剔除法 .....                 | (155)        |
| 5.4 循环变动测定的思路 .....               | (160)        |
| 【思考与练习】 .....                     | (160)        |
| <b>第 6 章 抽样分布与参数估计 .....</b>      | <b>(163)</b> |
| 【学习目标】 .....                      | (163)        |
| 【关键术语】 .....                      | (163)        |
| 【本章要点】 .....                      | (163)        |
| 6.1 抽样分布 .....                    | (163)        |
| 6.1.1 参数与统计量 .....                | (164)        |
| 6.1.2 抽样分布的概念 .....               | (164)        |
| 【案例 6-1】观察抽样分布的形成 .....           | (165)        |
| 6.1.3 简单随机抽样条件下样本平均数的抽样分布特征 ..... | (166)        |
| 【案例 6-2】如何在扑克牌游戏中找到获胜的概率区间 .....  | (170)        |
| 6.1.4 正态分布的标准化变换 .....            | (171)        |
| 【案例 6-3】估计样本中儿童平均体重的区间 .....      | (172)        |
| 6.2 简单随机抽样条件下总体平均数的估计 .....       | (173)        |
| 6.2.1 总体平均数的估计方法 .....            | (173)        |
| 【案例 6-4】通过样本估计一群儿童的平均体重 .....     | (173)        |
| 6.2.2 样本标准差的运用 .....              | (174)        |

|                                      |              |
|--------------------------------------|--------------|
| 【案例 6-5】估计一群大学生的平均体重 .....           | (175)        |
| 6.2.3 无放回抽样条件下的简单随机抽样误差 .....        | (176)        |
| 6.2.4 影响抽样误差的因素 .....                | (177)        |
| 6.3 样本量的计算 .....                     | (179)        |
| 6.3.1 简单随机抽样条件下样本量的计算 .....          | (179)        |
| 【案例 6-6】居民调查中的样本量计算 .....            | (180)        |
| 6.3.2 总体标准差未知情况下的处理方法 .....          | (180)        |
| 【案例 6-7】估计学校中拥有笔记本电脑的学生所占比重 .....    | (182)        |
| 6.3.3 市场调查中确定样本量的方法 .....            | (182)        |
| 【案例 6-8】某地方政府对乡镇工作群众满意度的评测调查设计 ..... | (183)        |
| 【案例 6-9】收视率调查的样本量计算 .....            | (184)        |
| 6.4 一般参数估计问题 .....                   | (185)        |
| 6.4.1 参数估计的概念 .....                  | (185)        |
| 6.4.2 参数估计的一般方法 .....                | (189)        |
| 6.4.3 常用的抽样分布 .....                  | (190)        |
| 6.4.4 利用常见抽样分布进行参数估计 .....           | (196)        |
| 【案例 6-10】方差未知条件下的单总体均值估计 .....       | (198)        |
| 【案例 6-11】单总体方差估计 .....               | (199)        |
| 【案例 6-12】方差已知条件下两个总体均值之差的估计 .....    | (200)        |
| 【案例 6-13】两个总体方差比的估计 .....            | (201)        |
| 【案例 6-14】单总体的比例估计 .....              | (202)        |
| 【案例 6-15】样本比例为 0 时的区间估计 .....        | (203)        |
| 【思考与练习】 .....                        | (203)        |
| <b>第 7 章 假设检验 .....</b>              | <b>(205)</b> |
| 【学习目标】 .....                         | (205)        |
| 【关键术语】 .....                         | (205)        |
| 【本章要点】 .....                         | (205)        |
| 7.1 检验原理 .....                       | (205)        |
| 7.1.1 假设检验的概念 .....                  | (205)        |
| 7.1.2 假设检验的步骤 .....                  | (207)        |
| 【案例 7-1】章鱼保罗是神迹吗 .....               | (207)        |

|                                 |              |
|---------------------------------|--------------|
| 【案例 7-2】对比弃真错误与取伪错误 .....       | (208)        |
| 7.1.3 假设检验的设计原理 .....           | (210)        |
| 【案例 7-3】工厂质检人员的选择 .....         | (211)        |
| 【案例 7-4】总体均值检验 .....            | (212)        |
| 【案例 7-5】总体分布检验 .....            | (213)        |
| 【案例 7-6】相关性检验 .....             | (213)        |
| 【案例 7-7】裁判的选择 .....             | (215)        |
| 7.2 利用常用抽样分布的假设检验 .....         | (216)        |
| 7.2.1 假设检验的一般方法 .....           | (216)        |
| 【案例 7-8】备择假设的设定对检验结果的影响 .....   | (219)        |
| 7.2.2 假设检验实例 .....              | (220)        |
| 【案例 7-9】房价变动的均值检验 .....         | (220)        |
| 【案例 7-10】地方政府是否实现了对房价的承诺 .....  | (223)        |
| 【案例 7-11】总体方差的检验 .....          | (225)        |
| 【案例 7-12】是否男性比女性更喜欢某部电视剧 .....  | (226)        |
| 7.3 非参数统计 .....                 | (228)        |
| 7.3.1 Pearson 拟合优度检验 .....      | (228)        |
| 【案例 7-13】匿名手稿是否为某作家所著 .....     | (229)        |
| 7.3.2 符号检验 .....                | (230)        |
| 【案例 7-14】是否家庭中丈夫承担的家务劳动更少 ..... | (231)        |
| 7.3.3 游程检验 .....                | (232)        |
| 【案例 7-15】学生的行走顺序有规律吗 .....      | (232)        |
| 7.3.4 上下游程检验 .....              | (233)        |
| 7.4 列联表分析 .....                 | (234)        |
| 7.4.1 列联表的概念 .....              | (234)        |
| 【案例 7-16】学生的来源与口味是否相关 .....     | (234)        |
| 7.4.2 列联表分析的步骤 .....            | (235)        |
| 7.4.3 列联表分析的其他问题 .....          | (237)        |
| 【案例 7-17】不同专业的学生选择餐厅是否有差异 ..... | (238)        |
| 【思考与练习】 .....                   | (239)        |
| <b>第 8 章 相关性分析 .....</b>        | <b>(241)</b> |
| 【学习目标】 .....                    | (241)        |

|                           |       |
|---------------------------|-------|
| 【关键术语】                    | (241) |
| 【本章要点】                    | (241) |
| 8.1 相关性概述                 | (241) |
| 8.1.1 相关性的概念              | (241) |
| 8.1.2 相关的类型               | (242) |
| 8.2 相关系数                  | (243) |
| 8.2.1 Pearson 相关系数        | (243) |
| 8.2.2 Spearman 等级相关系数     | (247) |
| 8.2.3 Kendall $\tau$ 相关系数 | (249) |
| 8.3 方差分析                  | (250) |
| 8.3.1 基本概念                | (250) |
| 【案例 8-1】产品包装颜色对销售量是否有影响   | (250) |
| 8.3.2 方差分析的步骤             | (251) |
| 8.3.3 对方差分析基本假设的讨论        | (255) |
| 8.4 回归分析                  | (256) |
| 8.4.1 回归分析中的概念            | (257) |
| 8.4.2 线性回归方法              | (259) |
| 8.4.3 线性回归的统计检验           | (264) |
| 8.4.4 计量经济学               | (267) |
| 【思考与练习】                   | (271) |
| 部分思考与练习题参考答案              | (272) |
| 参考文献                      | (278) |



## 第1章 绪论

### 【学习目标】

通过对统计学学科构成与学科简史的学习，形成对统计学学科的初步印象，并领悟统计学课程学习的基本要求。

### 【关键术语】

描述统计，推断统计。

### 【本章要点】

1. 统计学的学科构成；
2. 统计学科形成的过程；
3. 统计学的课程学习要求。

## 1.1 统计学简介

统计是一种具有悠久历史的社会实践活动。可以说，自从有了国家就有了统计工作。最初的统计活动是为统治者管理国家的需要而进行的收集资料的工作，涉及计算国家的人力、物力、财力等活动。今天，统计已经发展成为各行各业开展活动时必不可少的一项基础工作。所有收集信息和处理信息的活动，都可以归结为统计工作。

古往今来的统计学者对统计学给予了不同的定义。根据美国统计学家 David Freedman 等著《统计学》中的定义：统计学是对令人困惑的问题做出数字设想的艺术。

把统计学称为艺术显然有些夸张，但这一定义的目的是提示统计工作者，应当创造性地提出和解决统计问题，不应囿于某些条条框框去理解统计这门科学。

## 【案例 1-1】

## 如何统计一个水库里的鱼

在一个水库中养着许多鱼，管理人员希望了解鱼的大致数量，这就是一个实践中的统计学问题。

由于鱼是不听从指挥，会在各处自由游动的，因此，在进行统计时，必须创造性地提出解决方案。

一种解决方法是先从水库的不同位置捕捞 1000 条鱼，在每条鱼的尾部作上一个标记，应当保证标记不会影响鱼的自由游动。然后，将鱼全部放回水库。几天后，从水库中再捕上来 2000 条鱼，检查其中尾巴上有标记的鱼的数目。假定在第二次捕上来的 2000 条鱼中，有 20 条尾巴上做了标记，则可以推断，水库中鱼的总数大致为

$$1000 \div \frac{20}{2000} = 10 \text{ (万条)}$$

上述这个案例在实践中是经常见到的，对于一个统计工作者来说，做出一个 10 万条鱼的估计是不够的，他还应当对这一估计的精度做出判断，因为捕鱼的过程中存在着许多随机性的因素，而这些因素会导致按照简单的数学公式计算出来的结果存在着一定的误差。在假定捕鱼过程满足随机性要求的情况下，可以计算出在一定的可靠性条件下，水库中鱼的总数大致在哪个区间之内。而如果随机性要求受到破坏，例如因为做标记的原因，尾巴上有标记的鱼比没有标记的鱼更容易被捕捞到，这时候就需要评估这样的原因造成的误差。

上述所举的例子，可以被视为统计艺术性的一种表现。在这个推断过程中，所使用到的数学方法并不难，难的是提出这样一个解决方案。在实践工作中，统计的应用方面是十分复杂的，只有将统计理解为一种艺术，创造性地提出新的方法去解决新的问题，才是真正地掌握了统计的精髓。

实践中涉及数据收集和数据处理的工作，大多数都可以归入统计学研究的范畴，例如，以下一些问题都属于统计学问题。

(1) 出于经济管理的目的，地方政府需要了解辖区范围内有多少家企业，规模如何，分别从事哪些行业的活动。

(2) 期末统计之后，老师需要了解各班学生考试成绩的基本情况，包括各班的平均分、学生成绩的差异。

(3) 物价部门调查了市场上若干种商品的价格，希望用一个综合的指标来反