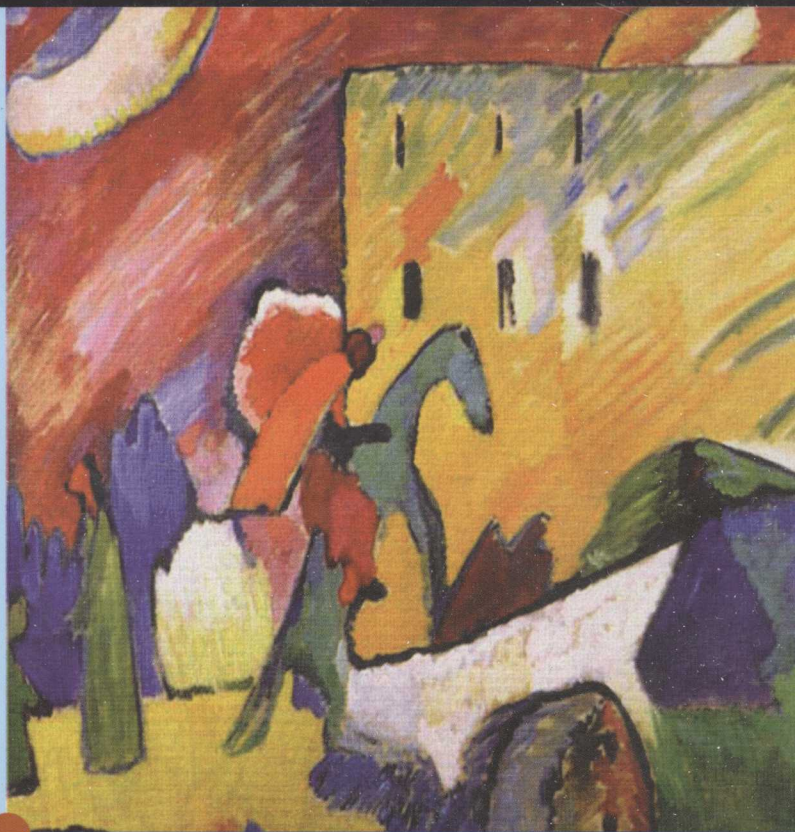


(原书第 13 版)

商务与经济统计方法 全球数据集

Statistical Techniques in Business and Economics: with Global Data Sets (13th Edition)



道格拉斯 A. 林德 (Douglas A. Lind)

卡罗来纳海岸大学和托莱多大学

威廉 G. 马歇尔 (William G. Marchal)

托莱多大学

塞缪尔 A. 沃森 (Samuel A. Wathen)

卡罗来纳海岸大学

冯燕奇 叶光 聂巧平 译



附光盘



机械工业出版社
China Machine Press

经济教材译丛

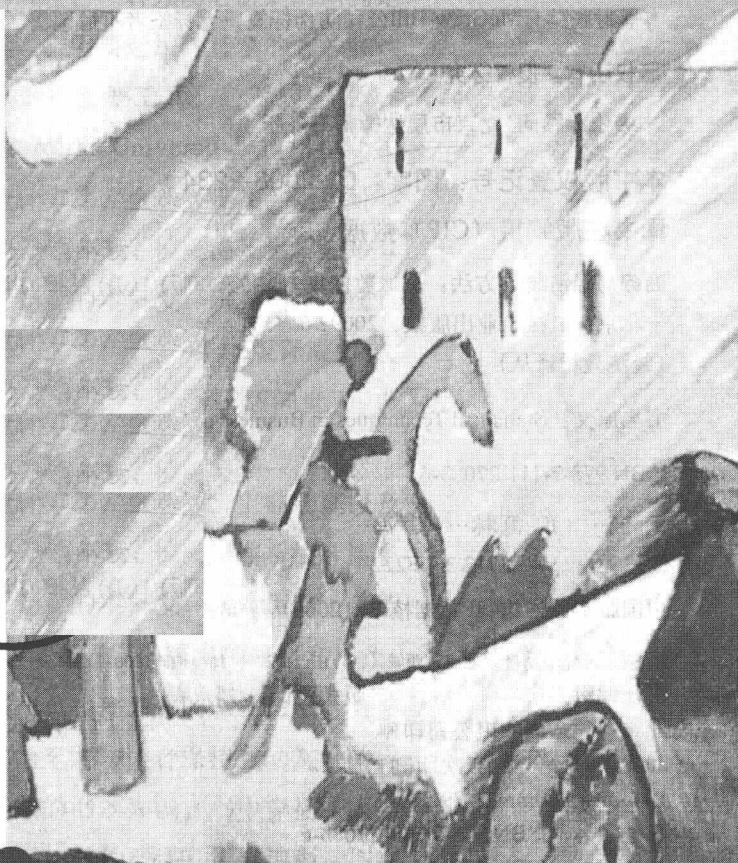
F222.1
L491-2

(原书第 13 版)



商务与经济统计方法 全球数据集

Statistical Techniques in Business and Economics: with Global Data Sets (13th Edition)



道格拉斯 A. 林德 (Douglas A. Lind)

卡罗来纳海岸大学和托莱多大学

威廉 G. 马歇尔 (William G. Marchal)

托莱多大学

塞缪尔 A. 沃森 (Samuel A. Wathen)

卡罗来纳海岸大学

冯燕奇 叶光 聂巧平



机械工业出版社
China Machine Press

前言

写作本书的目的，是为主修管理、市场营销、金融、会计、经济以及工商管理其他领域的学生提供一本介绍描述统计和推断统计诸多应用概况的教材。当我们致力于商务应用时，我们还利用了很多学生们提出的问题和例题，而这些问题和例题的解决并不需要学生学过商业课程。

1967年本书的第1版出版。在那时，收集相关的商业数据是很困难的一件工作。现在这种状况已经改变了！今天，收集数据已不再是一个难题了。你在食品店购买的食物数量在收银台已被自动地记录下来。电话公司能记录下通话的次数、通话的时间和通话人的电话号码。信用卡公司能够记录下卡号、时间和日期以及购买金额等信息。医疗设施能自动地监控患者的心率、血压和体温。大量的商业信息几乎可以即时记录和报告。例如，在CNN、《今日美国》和MSNBC的网站上，你都能浏览到滞后时间不超过20分钟的股票价格。

今天，我们需要能够处理大量数字信息的技术。首先，我们需要由其他人提供关键的客户信息。其次，我们必须能够把大量信息简化为更有意义的形式，以便我们能够有效地解释、判断和决策。

现在，所有的学生都有计算器，并且大部分学生都有电脑，或者能在校园实验室里使用电脑。统计软件，例如Excel和MINITAB都可在电脑上应用。实现软件运行，得到计算结果所必需的命令，见本书所附光盘。在各章里，我们使用了很多屏幕截图使学生熟悉软件输出的特征。因为电脑和软件的有效性，我们不必再长久地陷于烦琐的计算。我们用了许多解释性的例题来代替计算性的例题，这样可以帮助学生更好地理解 and 解释统计结果。此外，我们现在更加重视统计问题的概念上的特征。在做出上述改变的同时，我们仍然继续为关键性概念提供实证的例题。

本书第13版是众人共同努力的结果：学生、同事、评论家和McGraw-Hill/Irwin的工作人员。我们感谢他们。我们要向调查和项目

小组的参与者，以及评论家们表示诚挚的感谢。

下面是被感谢者的名单（略）。

他们的建议，以及对上一版和本版原稿的详尽审阅，使得这本教材更加完善。

特别需要感谢的有：William Paterson 大学的 Leonard Presby 博士，加州州立大学 Hayward 分校的 Julia Norton 和迈阿密达得（Dade）学院的 Christopher Rogers，他们审阅了原稿和清样，检查了习题以确保准确无误；南卡罗来纳州大学的 Kathleen Whitcomb 教授准备了学习指南；卡罗来纳海岸大学的塞缪尔 A. 沃森博士准备了试题库；南俄勒冈大学的 Rene Ordonez 教授准备了 PowerPoint 演示；Denise Heban 女士和其他作者准备了指导者手册。

我们还想要向 McGraw-Hill/Irwin 的工作人员表示感谢，包括策划编辑 Richard T. Hercher, Jr., 编辑 Christina Sanders，销售经理 Sankha Basu，项目经理 James Labeots，以及那些我们虽然不知道姓名，但为本书做出宝贵贡献的人。

一、课程简介

统计学是对数据进行收集、整理、展示、分析和解释,以帮助人们更为有效地做出正确决策的一门方法论科学。统计学在众多的专业、学科领域中,特别是在商务和经济活动中,都有着重要的应用。因此,统计学历来就是高等院校工商管理、经济类各个专业学生的一门应用性很强的基础课程。

二、选课建议

本教材在教学时可根据具体情况适当选取,以适应各种层次学生的需要,既可作为研究生、MBA和本科生的教材,也可供从事工商管理 and 经济分析的各类专业人员参考。

对于本科生,教学内容以描述统计为主,推断统计主要介绍区间估计和假设检验一些基本概念和方法。

对于研究生、MBA层次的学生,主要介绍推断统计方法,描述统计内容原则上不在课上讲授,学生可在课下复习。

三、课程目标

通过本课程的学习,使学生能够掌握统计学的基本概念和基本方法,能利用描述统计方法对商务与经济问题中的各种数据进行收集、整理、图示以及分布特征测度的计算,能利用推断统计方法对商务与经济中的实际问题进行估计、预测、评价和综合分析。

四、教学内容、学习要点及课时安排

教 学 内 容	学 习 要 点	课时安排	
		本科生	研究生 MBA
第 1 章 什么是统计学	(1) 理解为什么要学习统计学 (2) 了解描述统计学和推断统计学的内涵 (3) 区分定性变量和定量变量 (4) 描述离散变量如何不同于连续变量 (5) 区分名义、顺序、区间和比率测量尺度	2	
第 2 章 描述数据 频数表、频数分布 和图形表示	(1) 将定性数据整理成频数表 (2) 利用条形图和饼形图表示数据的频数表 (3) 将定量数据整理成频数分布 (4) 利用直方图、频数折线图和累积频数折线图表示数据的频数分布	3	
第 3 章 描述数据 数值度量	(1) 计算算术平均数、加权平均数、中位数、众数和几何平均数 (2) 解释每种位置度量的特征、用途、优点和缺点 (3) 对于对称分布和有偏分布，确定平均数、中位数和众数的位置 (4) 计算和解释极差、平均离差、方差和标准差 (5) 理解每种离散程度度量的特征、用途、优点和缺点 (6) 理解与一组数据相联系的切比雪夫定理和经验法则	6	
第 4 章 描述数据 数据展示和数据开发	(1) 绘制和解释点状图 (2) 构造和解释茎叶图 (3) 计算和理解四分位数、十分位数和百分位数 (4) 构造和解释箱形图 (5) 计算和理解偏度系数 (6) 绘制和解释散点图 (7) 构造和解释列联表	3	
第 5 章 概率论概述	(1) 概率的定义 (2) 描述概率的古典方法、经验方法和主观方法 (3) 理解试验、事件、结果、排列和组合的含义 (4) 条件概率和联合概率的定义 (5) 使用加法法则和乘法法则计算概率 (6) 应用树形图组织和计算概率 (7) 利用贝叶斯定理计算概率	4	
第 6 章 离散型概率分布	(1) 掌握概率分布和随机变量的定义 (2) 区分离散型和连续型概率分布 (3) 计算离散型概率分布的均值、方差和标准差 (4) 描述二项概率分布的特征并利用二项概率分布计算概率 (5) 描述超几何概率分布的特征并利用超几何概率分布计算概率 (6) 描述泊松概率分布的特征并利用泊松概率分布计算概率	4	
第 7 章 连续型概率分布	(1) 理解离散型分布和连续型分布之间的区别 (2) 计算均匀分布的均值和标准差 (3) 利用均匀分布计算概率 (4) 列举正态概率分布的主要特征	4	

(续)

教 学 内 容	学 习 要 点	课时安排	
		本科生	研究生 MBA
第 7 章 连续型概率分布	(5) 定义和计算 z 值 (6) 确定位于正态分布两点之间的观测值的概率 (7) 确定大于 (或小于) 正态分布一个点的观测值的概率 (8) 利用正态概率分布近似二项分布	4	
第 8 章 抽样方法和中心极限定理	(1) 解释为什么样本通常仅是认识总体的切实可行的方法 (2) 描述抽取一个样本的一些方法 (3) 定义和构造样本均值的抽样分布 (4) 理解和解释中心极限定理 (5) 利用中心极限定理得到取自一个指定总体的所有可能的样本均值的概率	4	4
第 9 章 估计和置信区间	(1) 定义一个点估计 (2) 定义一个置信水平 (3) 当总体标准差已知时, 构造总体均值的置信区间 (4) 当总体标准差未知时, 构造总体均值的置信区间 (5) 构造总体比例的置信区间 (6) 为属性和变量抽样确定样本容量	8	4
第 10 章 单样本假设检验	(1) 定义假设和假设检验 (2) 描述五步假设检验程序 (3) 区分单侧和双侧假设检验 (4) 对总体均值进行假设检验 (5) 对总体比例进行假设检验 (6) 定义第 I 类错误和第 II 类错误 (7) 计算第 II 类错误的概率	10	8
第 11 章 两样本假设检验	(1) 对两独立总体的均值之差进行假设检验 (2) 对两总体的比例之差进行假设检验 (3) 对成对或相依观测值之间的均值差进行假设检验 (4) 理解相依与独立样本之间的区别	4	4
第 12 章 方差分析	(1) 列举 F 分布的特征 (2) 进行假设检验来确定两个总体的方差是否相等 (3) 讨论方差分析的一般思想 (4) 将数据整理成单因素和双因素的 ANOVA 表 (5) 对三个或三个以上的处理均值进行假设检验 (6) 建立处理均值差的置信区间 (7) 利用区组化变量对处理均值进行假设检验 (8) 进行具有交互作用的双因素 ANOVA		12
第 13 章 线性回归与相关	(1) 理解和解释术语因变量和自变量 (2) 计算和解释相关系数、判定系数和估计的标准差 (3) 通过假设检验来确定总体的相关系数是否为零 (4) 计算最小二乘回归直线 (5) 构造及解释因变量的置信区间和预测区间	8	4
第 14 章 多元回归与相关分析	(1) 利用多元回归分析描述若干个自变量与一个因变量之间的关系 (2) 建立、解释和应用 ANOVA 表		8

(续)

教 学 内 容	学 习 要 点	课时安排	
		本科生	研究生 MBA
第 14 章 多元回归与相关分析	(3) 计算和解释多元估计的标准误差、多元判定系数和修正多元判定系数 (4) 进行假设检验来确定相关系数是否为零 (5) 对每一个回归系数进行假设检验 (6) 利用残差分析来评估多元回归分析的假定 (7) 评估相关自变量的影响 (8) 利用和理解定性自变量 (9) 理解和解释逐步回归法 (10) 理解和解释自变量中可能存在的交互作用		8
第 15 章 指数	(1) 对术语指数进行解释 (2) 理解加权与未加权指数之间的区别 (3) 构造并解释拉氏价格指数 (4) 构造并解释帕氏价格指数 (5) 构造并解释价值指数 (6) 说明如何构造消费价格指数并对其进行解释	8	6
第 16 章 时间序列与预测	(1) 定义时间序列的构成 (2) 计算移动平均 (3) 确定线性趋势方程 (4) 计算非线性趋势的趋势方程 (5) 利用趋势方程对未来时期进行预测, 并建立季节调整预测 (6) 确定并解释季节指数 (7) 利用季节指数消除数据的季节影响 (8) 检验自相关	12	8
第 17 章 非参数方法: 检验	(1) 列出 χ^2 分布的特征 (2) 进行假设检验, 对一组观察频数与期望分布进行比较 (3) 进行假设检验, 确定两种分类标准是否相关		6
第 18 章 非参数方法 分级数据分析	(1) 用二项分布和标准正态分布作为检验统计量, 对相依样本进行符号检验 (2) 利用威尔科克森符号秩检验对相依样本进行假设检验 (3) 对独立样本进行威尔科克森秩和检验并解释 (4) 对多个独立样本进行克鲁斯卡尔-沃利斯检验并解释 (5) 计算并解释斯皮尔曼秩相关系数 (6) 通过假设检验确定总体中秩之间的相关是否与 0 有显著差异		6
第 19 章 统计过程控制与质量管理	(1) 了解质量控制在生产和服务运营中的作用 (2) 定义和理解术语偶然原因、可指出的原因、受控、失控、属性和变量 (3) 绘制和说明帕累托图 (4) 绘制和说明鱼刺图 (5) 绘制和说明均值和极差图 (6) 绘制和说明缺陷比例图与 c -条形图 (7) 阐述接受抽样 (8) 为不同的抽样计划构造抽样特征曲线		6

(续)

教 学 内 容	学 习 要 点	课时安排	
		本科生	研究生 MBA
第 20 章 决策论基础	(1) 定义术语：自然状态、事件、备选方案和收益 (2) 把信息组织成收益表或决策树 (3) 确定备选方案的预期收益 (4) 计算机会损失和预期机会损失 (5) 评估信息的期望值		4
课时总计		80	80

说明：(1) 第 1~7 章为本科生的学习内容，对于研究生、MBA 层次的学生，不再安排课时讲授，学生可在课下复习。

(2) 案例讨论的课时可由教师灵活掌握分配使用。

目 录

译者序

前 言

教学建议

第1章 什么是统计学	1
1.1 引言	1
1.2 为什么学习统计学	2
1.3 什么是统计学	3
1.4 统计学的类型	4
1.5 变量的类型	6
1.6 测量尺度	7
1.7 伦理与统计	11
1.8 计算机应用	11
小结	12
本章习题	13

第2章 描述数据

频数表、频数分布和

图形表示

2.1 引言	14
2.2 构造频数表	15
2.3 构造频数分布：定量数据	19
2.4 一个软件例子	22
2.5 相对频数分布	23
2.6 频数分布的图形表示	24
小结	29
本章习题	29

第3章 描述数据

数值度量

3.1 引言	31
3.2 总体均值	32
3.3 样本均值	33
3.4 算术平均数的性质	34
3.5 加权平均数	35
3.6 中位数	37
3.7 众数	38
3.8 软件解法	39
3.9 均值、中位数和众数的 相对位置	40
3.10 几何平均数	41
3.11 为什么要研究离散程度	43
3.12 离散程度的度量	44
3.13 软件解法	49
3.14 标准差的解释与应用	50
3.15 分组数据的均值和标准差	52
3.16 伦理和结果报告	54
小结	55
符号读法	56
本章习题	57

第4章 描述数据

数据展示和数据开发

4.1 引言	58
4.2 点状图	58
4.3 茎叶图	60

4.4 离散程度的其他度量	63	第7章 连续型概率分布	132
4.5 偏度	67	7.1 引言	132
4.6 描述两个变量之间的关系	70	7.2 均匀概率分布族	133
小结	73	7.3 正态概率分布族	135
符号读法	74	7.4 标准正态概率分布	137
本章习题	74	7.5 二项分布的正态近似	145
第1~4章复习	76	小结	148
术语	76	本章习题	149
习题	78	第5~7章复习	151
案例	79	术语	152
第5章 概率论概述	81	习题	153
5.1 引言	81	案例	154
5.2 何谓概率	82	第8章 抽样方法和 中心极限定理	157
5.3 分配概率的方法	84	8.1 引言	157
5.4 计算概率的几个法则	87	8.2 抽样方法	158
5.5 列联表	94	8.3 抽样误差	163
5.6 树形图	95	8.4 样本均值的抽样分布	165
5.7 贝叶斯定理	97	8.5 中心极限定理	167
5.8 计数原理	101	8.6 样本均值抽样分布的应用	173
小结	105	小结	175
符号读法	106	符号读法	175
本章习题	106	本章习题	175
第6章 离散型概率分布	108	第9章 估计和置信区间	178
6.1 引言	108	9.1 引言	178
6.2 何谓概率分布	108	9.2 均值的点估计和置信区间	179
6.3 随机变量	109	9.3 比例的置信区间	189
6.4 概率分布的均值、方差和 标准差	111	9.4 有限总体修正因子	191
6.5 二项概率分布	113	9.5 选择适当的样本容量	193
6.6 超几何分布	120	小结	196
6.7 泊松概率分布	122	本章习题	196
6.8 协方差	126	第8~9章复习	198
小结	129	术语	198
本章习题	130	习题	199
		案例	200

第 10 章 单样本假设检验	201	12.2 F 分布	246
10.1 引言	201	12.3 比较两总体的方差	247
10.2 什么是假设	202	12.4 ANOVA 的假定	250
10.3 什么是假设检验	202	12.5 ANOVA 检验	251
10.4 假设检验五步法	202	12.6 关于成对处理均值的推断 ...	256
10.5 单侧与双侧显著性检验	206	12.7 双因素方差分析	258
10.6 总体均值的检验:		12.8 具有交互作用的双因素	
已知总体标准差	207	方差分析	262
10.7 假设检验中的 p -值	210	小结	266
10.8 总体均值的检验:		符号读法	267
总体标准差未知	212	本章习题	267
10.9 关于比例的检验	217	第 10~12 章复习	269
10.10 第 II 类错误	219	术语	270
小结	221	习题	271
符号读法	222	案例	272
本章习题	222		
第 11 章 两样本假设检验	224	第 13 章 线性回归与相关	274
11.1 引言	224	13.1 引言	274
11.2 两样本假设检验:		13.2 何谓相关分析	275
独立样本	225	13.3 相关系数	276
11.3 比例的两样本检验	229	13.4 判定系数	279
11.4 总体均值的比较:		13.5 相关系数的显著性检验	280
总体标准差未知		13.6 回归分析	282
(合并 t 检验)	232	13.7 估计的标准误差	287
11.5 总体均值的比较:		13.8 线性回归的假定	289
总体标准差不等	235	13.9 置信区间和预测区间	290
11.6 两样本的假设检验:		13.10 有关判定系数的更多内容 ...	293
相依样本	238	13.11 相关系数、判定系数与估计的	
11.7 相依样本与独立样本的		标准误差之间的关系	295
比较	240	13.12 变换数据	297
小结	242	13.13 协方差	299
符号读法	243	小结	301
本章习题	243	符号读法	302
		本章习题	302
第 12 章 方差分析	246	第 14 章 多元回归与相关分析 ...	305
12.1 引言	246	14.1 引言	305

14.2 多元回归分析	305	16.6 最小二乘法	368
14.3 回归方程对数据的 拟合好吗	309	16.7 非线性趋势	370
14.4 多元线性回归的推断	314	16.8 季节波动	371
14.5 评价多元回归的假定	319	16.9 消除数据的季节影响	376
14.6 逐步回归	325	16.10 杜宾-瓦特森统计量	379
14.7 具有交互作用的回归模型	327	小结	383
小结	330	本章习题	383
符号读法	331	第 15~16 章复习	386
本章习题	331	术语	387
第 13~14 章复习	336	习题	387
术语	336		
习题	337	第 17 章 非参数方法: 检验	388
案例	338	17.1 引言	388
第 15 章 指数	340	17.2 拟合优度检验: 期望频数相等	388
15.1 引言	340	17.3 拟合优度检验: 期望频数不等	392
15.2 简单指数	341	17.4 卡方的局限性	393
15.3 为什么要把数据转换成 指数	342	17.5 列联表分析	395
15.4 指数的构成	343	小结	398
15.5 未加权指数	344	符号读法	398
15.6 加权指数	346	本章习题	399
15.7 价值指数	349		
15.8 特殊用途指数	350	第 18 章 非参数方法	
15.9 消费者物价指数	353	分级数据分析	400
15.10 调整基期	356	18.1 引言	400
小结	357	18.2 符号检验	401
本章习题	358	18.3 相依样本的威尔科克森 符号秩检验	407
第 16 章 时间序列与预测	360	18.4 独立样本的威尔科克森秩和 检验	410
16.1 引言	360	18.5 克鲁斯卡尔-沃利斯检验: 秩方差分析	412
16.2 时间序列的构成	361	18.6 秩顺序相关	415
16.3 移动平均	363	小结	418
16.4 加权移动平均	365	符号读法	419
16.5 线性趋势	367		

本章习题	419	本章习题	443
第 17~18 章复习	420		
术语	421	第 20 章 决策论基础	445
习题	421	20.1 引言	445
案例	422	20.2 决策要素	446
		20.3 不确定条件下决策的案例 ...	446
第 19 章 统计过程控制与质量		20.4 最大最小策略、最大最大策略	
管理	423	及最小最大遗憾策略	450
19.1 引言	423	20.5 完全信息的价值	450
19.2 质量控制简史	423	20.6 敏感性分析	451
19.3 差异产生的原因	426	20.7 决策树	452
19.4 诊断图	427	小结	453
19.5 质量控制图的用途和类型 ...	429	本章习题	454
19.6 受控和失控情形	433		
19.7 属性控制图	435	附录 [⊖]	
19.8 接受抽样	439		
小结	442	Excel 命令与自测题答案 [⊖]	
符号读法	442		

⊖ 附录部分见本书配套光盘。
⊖ Excel 命令与自测题答案见本书配套光盘。

第 1 章

什么是统计学

学习目标

1. 理解为什么要学习统计学；
2. 了解描述统计学和推断统计学的内涵；
3. 区分定性变量和定量变量；
4. 区分离散变量与连续变量；
5. 区分名义尺度、顺序尺度、区间尺度和比率尺度。

1.1 引言

100 多年以前，英国作家、历史学家威尔斯（H.G.Wells）曾经说过，终有一日，定量分析对于有效地行使公民的权利和义务而言，将如同阅读能力一样必不可少。然而他却没有提到商业，因为当时工业革命才刚刚开始。威尔斯先生无疑是正确的。随着“商业经验”、对未来的“精心预测”和“直觉”成为成功管理者的三大关键特质，单就管理过程中的量化决策而言，今天的商业问题已经变得十分复杂了。

统计学是进行决策的工具之一，不仅商业人士如此，日常生活中我们也会用到一些统计概念。例如，新的一天开始了，你打开热水器让水先流一会儿，然后用手试试水温，决定是否需要加些热水或冷水，如果水温刚好合适，你就开始淋浴。又比如你在一个食品店里准备买一张比萨饼，每个摊位都会提供小块的比萨饼以供品尝，经过品尝你才会决定是否要在这个摊位购买。在上面两个例子中，决定的做出和行为的选择都建立在抽取样本的基础上。

商业活动中也经常会遇到类似的问题。凯洛格公司必须保证 25.5 克盒装食品的平均重量与标签相符。为此，他们设定一个比标签说明稍高的“目标”重量，对每个装满食品的包装盒进行称重，称重机器每小时都会报告这一时间段内包装盒中食品重量的分布，以及由于重量达不到标签说明而被“清除”的数目。质检部门也会随机从生产线抽取样本，并检查产品质量和盒子重量，如果平均产品重量与目标重量显著不符，或者被清除的包装盒过多，就需要对整个流程进行调整。

美联储前任主席艾伦·格林斯潘（Alan Greenspan）的演说对全球股市动向和政治观点都有一定的影响力，他深知在公众演说中利用统计工具与统计技术提供精确、及时信息的重要性。格林斯潘博士在美国技术委员会(National Skills Summit)的演说中指出：“工人不仅需要掌握专业的技能知识，还需要具备创造、分析和处理信息并有效地与其他人交流的能力。也就是说，能够从众多观点和意见中筛选出事实，并按照一定方式对它们进行组织和整理，从中获取信息。”

作为一名商学或经济学专业的学生，你需要具备组织、分析和处理数据并提供信息的基本知识和技能。本书将告诉你一些基本的统计技术和方法，使你个人的能力和商业决策能力得到更好的开发。

1.2 为什么学习统计学

浏览你所在大学的课程表，你会发现许多院系统计学都是必修课。为什么会这样？工学院、文学院的心理学和社会学系以及商学院开设的统计学课程又有哪些不同？事实上，课程内容基本上是相同的，它们最大的区别在于所使用的案例。商学院的学生对利润、工作时间和工资等问题感兴趣，心理学系的学生对测试分数感兴趣，而工学院的学生感兴趣的是特定机器可以生产多少件产品。不过，这三个学院的学生对数据的典型值和离散程度都有着共同的兴趣。此外，对数学基础的要求也有所不同，工学院的统计学课程通常需要计算，商学院和教育学院则更多地强调应用。如果你已经完成了高中的代数课程，就完全具备了本书所需要的数学基础。

究竟为什么这么多专业都需要学习统计学呢？其首要原因在于数字化信息随处可见。翻开报纸（《今日美国》）、新闻杂志（《时代》、《新闻周刊》、《美国新闻与世界报道》）、商业杂志（《商业周刊》、《福布斯》）、大众杂志（《人物》）、女性杂志（《女性家庭周刊》或《Elle》）、体育杂志（*Sports Illustrated*、*ESPN The Magazine*），你将完全处于数字化信息的狂轰滥炸中。^①

下面就是几个例子：

- 2003年，美国一般家庭的年收入为43 318美元。东北部地区家庭平均收入为46 742美元，中西部地区为44 732美元，南部地区为39 823美元，西部地区为46 820美元。你可以通过网址<http://www.census.gov/hhes/incoms> 查阅最新的信息。
- 2005年7月，波音公司宣布2005年上半年共有155架飞机交付使用，其中133架是波音737。作为最大的买家，西北航空公司购买了22架波音737。你可以登录波音公司的网址www.boeing.com，在搜索框中输入英文词组“orders and deliveries”，在弹出的网页中寻找关于波音公司飞机订购和交付的最新信息。同时，网页左侧还有一个订单示意图，从中可以查看当年的飞机交付使用情况。
- 《今日美国》（www.usatoday.com）的“Snapshots”（实录）栏目会刊登不同的研究组织、基金和联邦政府的调查结果。例如，许多人相信电子邮件比普通邮件更优越，但美国邮电业通过最新的调查指出，67%的成年人认为普通邮件比电子邮件更私人化，56%的成年人乐于接收普通邮件，55%的成年人则希望能够打开一封普通邮件。

第二个学习统计学的原因在于，对我们日常生活产生影响的各种决策经常需要利用统计技术，也就是说，统计技术已经影响到个人生活的质量。下面有几个例子：

- 保险公司通过统计分析来确定家庭、汽车、人寿和健康保险的费率。利用网页www.ssa.gov/OACT/STATS/table4c.html（注意字母的大小写）上的一些表格，可以估算出20岁女性的平均剩余寿命为60.25年，80岁女性的平均剩余寿命为4.56年，50岁男性的平均剩余寿命则为27.87年。在这些关于未来寿命预期的基础上，人寿保险的保费就可以计算出来。
- 环境保护组织（Environmental Protection Agency）非常关注伊利湖的水质问题。他们定期采集湖水样本，确定湖水污染程度，以将水质控制在合格水平。
- 医学研究者对使用不同药物和不同治疗方法的病人的治愈率进行研究。例如，对膝盖受损的病人采用手术治疗或物理治疗的效果如何？如果你每天服用一片阿司匹林，患心脏病的风险是否会因此而降低？

① 我们为什么要学习统计学的例子。

第三个学习统计学的原因在于，统计方法能够帮助你了解决策是怎样做出的，并使你更加清楚这些决策会对你产生什么样的影响。

不论选择了什么样的工作，你都会发现自己面临着许许多多的决策问题，而掌握一种数据分析技术对此将大有裨益。为了做出明智的决策，你需要：

- (1) 确定现有信息是否充分，是否需要额外的信息；
- (2) 如果需要收集额外信息，确保不要被这些信息所误导；
- (3) 将信息以有效且直观的形式加以汇总；
- (4) 对可得信息进行分析；
- (5) 得出结论并进行推断，同时对错误结果的风险进行评价。

本书所讲述的各种统计方法将为你提供一个进行决策的框架。

总而言之，至少有3个原因让我们学习统计学：(1) 数据随处可见；(2) 对我们日常生活产生影响的各种决策经常需要利用统计技术；(3) 无论从事什么样的工作，你都不可避免地要与数据打交道并做出很多专业的决策。对统计方法的了解将会使你的决策更加有效。

1.3 什么是统计学

日常生活中，我们经常遇到统计 (statistics) 这个词。我们应该怎样定义它呢？它具有两层含义。在比较普遍的用法中，统计指的是数字化信息。例如，大学毕业生的平均起始工资；上一年度因为酒精中毒而死亡的人数；从昨天到今天，道琼斯工业平均指数的变化，以及2005赛季芝加哥队本垒打的次数。上述例子中，统计是一些数值或者百分比。其他的例子还有：

- 在美国，一般汽车每年行驶 11 099 英里，公共汽车每年行驶 9 353 英里，卡车每年行驶 13 942 英里。在加拿大，相应的数据为一般汽车 10 371 英里，公共汽车 19 823 英里，卡车 7 001 英里。
- 等待技术支持的时间平均为 17 分钟。
- 自 1945 以来，商业周期的平均长度为 61 个月。

这些都是统计的实例。统计就是一组数字化信息的集合。

实践中的统计

我们请你注意一个具有特色的标题——实践中的统计。仔细阅读每一篇文章，你可以充分了解统计学在管理、经济、护理、法律实务、体育及其他领域中的广泛应用。

- 2005 年，《福布斯》公布的美国富豪排名榜中，微软公司的创始人盖茨名列第一。据估计，他的净资产约为 465 亿美元。
- 2005 年，按总收益排名，美国最大的四家公司依次为：埃克森美孚石油公司 (ExxonMobil)、通用汽车 (General Motors)、福特公司 (Ford) 和雪佛龙公司 (Chevron)。
- 在美国，一般的高中生毕业生一生的总收入为 120 万美元；获得学士学位的大学毕业生一生的总收入为 210 万美元；获得硕士学位的研究生一生的总收入为 250 万美元。

统计信息还经常可以表示为图形的形式，而在吸引读者注意力和描述大量数据信息方面，图形是非常有用的。例如，图 1-1 给出在美国超市的点和薯片类食品中，Frito-Lay 公司的销售量及所占的市场份额。你只需快速一瞥，就可以看出薯片的销售总量为 8 亿磅^①，而 Frito-Lay 公司的市场份额为 64%。从图 1-1 中还可以看出，Frito-Lay 公司的玉米饼市场份额为 82%。

① 1 磅 = 0.453 千克。