

张建同 孙昌言 编著

以Excel和SPSS为工具 的管理统计



清华大学出版社

F₂₂₂
57

以 Excel 和 SPSS 为 工具的管理统计

张建同 孙昌言 编著

清华大学出版社

北 京

内 容 简 介

本书依托目前最为普及的 Excel 软件和社会科学领域最为普及的 SPSS 统计软件为基本工具,结合具体案例讲述管理科学中的统计学原理和方法。与计算机软件相结合是本书最大的特点。另一大特点是与实际联系密切,每一章节都提供了案例以辅助教学。本书包括数据收集和整理、统计数据的汇总、概率基础、抽样分布和参数估计、假设检验、方差分析、正交试验设计和回归分析等基本内容,还包括多元统计分析的基本内容,即因子分析和主成分分析、聚类分析和判别分析。

本书可作为高等院校经济管理类本科、硕士、MBA 和工程硕士等专业相应课程的教学用书,也可作为各类经济管理人员的参考用书。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

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。

本书防伪标签采用特殊防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

以 Excel 和 SPSS 为工具的管理统计/张建同,孙昌言编著. —北京:清华大学出版社,2005.9
ISBN 7-302-11702-0

I. 以… II. ①张…②孙… III. ①电子表格系统, Excel ②统计分析—软件包, SPSS IV. ①TP391.13
②C819

中国版本图书馆 CIP 数据核字(2005)第 098850 号

出 版 者:清华大学出版社 地 址:北京清华大学学研大厦
http://www.tup.com.cn 邮 编:100084
社 总 机:010-62770175 客户服务:010-62776969

组稿编辑:张立红(zlh-zlq@263.net)

文稿编辑:崔 伟(cuiwei80@163.com)

封面设计:鼎典智造(北京)企业策划公司

版式设计:康 博

印 刷 者:北京市世界知识印刷厂

装 订 者:北京市密云县京文制本装订厂

发 行 者:新华书店总店北京发行所

开 本:185×260 印张:24 字数:525 千字

版 次:2005 年 9 月第 1 版 2005 年 9 月第 1 次印刷

书 号:ISBN 7-302-11702-0/F·1316

印 数:1~4000

定 价:36.00 元

前 言

在经济全球化和信息化的环境下,管理决策者面临着高度不确定的外部环境和巨大的风险。管理统计学是定量分析非确定性问题的规律,帮助管理决策者进行科学决策、规避风险、获取最优经济和社会效益的科学方法,已成为现代科学管理中必不可少的强有力工具。

本书在确保知识的系统性和正确性的基础上,尽量使用通俗易懂的语言,阐述管理统计的基本概念、基本原理及解决问题的基本方法和应用条件,略去了繁琐的推导证明过程,使读者更容易理解和掌握。鉴于计算机是求解统计问题的有效手段,本书中对主要的概率统计方法,都以常用的计算机软件 Excel 和 SPSS 作为求解运算和分析的主要工具,以大幅度提高本课程的实用价值和学生的计算机应用能力。与计算机软件相结合是本书最大的特点。本书另一大特点是与实际联系密切,每一章节都提供了案例辅助教学。

本书包括数据收集和整理、统计数据的汇总、概率基础、抽样分布和参数估计、假设检验、方差分析、正交试验设计和回归分析等基本内容,还包括多元统计分析的基本内容,即因子分析和主成分分析、聚类分析和判别分析。

本书可作为高等院校经济管理类本科、硕士、MBA 和工程硕士等专业相应课程的教学用书,也可作为各类经济管理人士的参考用书。其中在章节前没加注*号的,适用于本科和 MBA 的教学和本科层次的读者,在章节前加*号的,适用于硕士研究生的教学和硕士研究生层次的读者。

全书共分 13 章,本书第 3, 4, 5, 6, 8, 9 章由孙昌言编写;第 7, 10, 11, 12, 13 章由张建同编写;第 1, 2 章由张建同和孙昌言共同编写。全书由张建同总纂并定稿。由于水平有限,书中难免错误或不足之处,恳请读者批评指正。

张建同

目 录

第 1 章 统计数据的收集与整理	1
§1.1 统计学概述	1
1.1.1 统计的广泛应用	1
1.1.2 统计与统计学	2
1.1.3 统计研究对象的特点	2
1.1.4 统计学的分类	3
§1.2 数据的收集	4
1.2.1 普遍调查	4
1.2.2 重点调查	5
1.2.3 典型调查	5
1.2.4 抽样调查	6
1.2.5 网上调查	6
§1.3 问卷设计	9
1.3.1 合理性	9
1.3.2 一般性	10
1.3.3 逻辑性	10
1.3.4 明确性	11
1.3.5 非诱导性	11
1.3.6 便于整理、分析	11
§1.4 统计表和统计图	12
1.4.1 频数分布表	12
1.4.2 统计图	15
习题	24
第 2 章 统计数据的汇总	27
§2.1 度量中心趋势的指标	27
2.1.1 算术平均数	27
2.1.2 中位数	28
2.1.3 众数	29
2.1.4 算术平均数、中位数和众数间的关系	31
2.1.5 几何平均数	32

§2.2	度量离散程度的指标	33
2.2.1	极差	33
2.2.2	平均差	33
2.2.3	方差和标准差	34
2.2.4	变异系数	36
2.2.5	标准差系数	36
§2.3	度量偏斜程度的指标	36
2.3.1	用标准差为单位计量的偏度系数	36
2.3.2	使用三阶中心矩计量的偏度系数	36
§2.4	利用 Excel 数据分析功能求各种统计指标	37
	习题二	39
第 3 章	概率基础	41
§3.1	引言	41
§3.2	随机试验与随机事件	43
3.2.1	随机现象	43
3.2.2	随机试验	43
3.2.3	随机事件	44
3.2.4	事件间的关系和运算	44
§3.3	概率	47
3.3.1	频率与概率	47
3.3.2	条件概率	50
3.3.3	事件的独立性	53
§3.4	随机变量及其分布函数	54
3.4.1	随机变量	54
3.4.2	随机变量的分布函数	55
§3.5	离散型随机变量	56
3.5.1	离散型随机变量的概率分布	56
3.5.2	几种重要的离散型分布	57
§3.6	连续型随机变量	61
3.6.1	概率密度	61
3.6.2	几种重要的连续型分布	62
§3.7	随机变量的数学期望和方差	67
3.7.1	数学期望	67
3.7.2	方差	69
§3.8	大数定律和中心极限定理及其计算机模拟验证	71

3.8.1	大数定律	71
3.8.2	中心极限定理	72
3.8.3	中心极限定理的动态模拟验证	72
§3.9	新产品投资决策案例分析	74
3.9.1	投产后各种销售状况下的项目净现值	74
3.9.2	不考虑试生产时的最优决策分析	75
3.9.3	考虑试生产并获取用户试用反馈信息的方案分析	75
3.9.4	追加信息的价值	77
	习题三	78
第 4 章	抽样分布和参数估计	81
§4.1	简单随机抽样和统计量	81
4.1.1	随机样本	81
4.1.2	统计量及其分布	84
§4.2	其他抽样方法	90
4.2.1	分层随机抽样	90
4.2.2	整群抽样	90
4.2.3	系统抽样	91
4.2.4	方便抽样	91
4.2.5	判断抽样	91
§4.3	参数的点估计	91
4.3.1	参数的点估计	92
4.3.2	点估计的方法	92
4.3.3	估计量的评价标准	93
§4.4	区间估计	95
4.4.1	单个正态总体均值和方差的区间估计	96
4.4.2	总体比率的区间估计	99
*4.4.3	两个正态总体的均值差和方差比的区间估计	100
4.4.4	单侧置信限的估计	103
4.4.5	区间估计小结	104
	习题四	105
第 5 章	假设检验	107
§5.1	案例介绍	107
§5.2	假设检验的基本原理	110
5.2.1	假设检验的基本原理和步骤	110
5.2.2	检验中可能犯的两类错误	111

§5.3	单个正态总体均值的检验	112
§5.4	单个正态总体方差的检验 (χ^2 检验)	114
§5.5	两个独立正态总体均值的检验	116
§5.6	成对样本试验的均值检验	120
§5.7	两个正态总体方差的检验 (F 检验)	122
§5.8	总体比例的检验	124
5.8.1	单个总体比例的检验	124
5.8.2	两个总体比例的检验	124
§5.9	假设检验小结	126
*§5.10	总体分布的 χ^2 检验	127
5.10.1	检验的基本原理	127
5.10.2	检验的具体步骤	128
§5.11	SPSS 软件的使用	130
5.11.1	SPSS 简介	130
5.11.2	独立样本 T 检验 (两独立总体均值的检验)	131
5.11.3	成对样本的 T 检验	135
	习题五	137
第 6 章	方差分析	139
§6.1	引言	139
6.1.1	问题的提出	139
6.1.2	方差分析的基本概念	141
6.1.3	方差分析的基本假设条件	141
6.1.4	方差分析的目的	142
§6.2	单因子方差分析	142
6.2.1	单因子试验的数学模型	142
6.2.2	方差分析的基本方法	143
6.2.3	检验 H_0 的统计量	144
6.2.4	方差分析表	144
6.2.5	进一步的分析	146
§6.3	双因子方差分析	147
6.3.1	不考虑交互作用的双因子方差分析	147
6.3.2	考虑交互作用时的双因子方差分析	150
§6.4	SPSS 软件的使用	154
6.4.1	单因子方差分析	154
6.4.2	双因子和多因子方差分析	157
	习题六	159

*第 7 章 正交试验设计	163
§7.1 试验设计的基本概念	163
§7.2 无交互作用的正交设计	165
7.2.1 试验的步骤	165
7.2.2 数据分析	167
§7.3 有交互作用的正交设计	175
7.3.1 试验的步骤	175
7.3.2 数据分析	177
习题七	180
第 8 章 回归分析	181
§8.1 引言	181
8.1.1 变量间的两类关系	182
8.1.2 线性回归的数学模型	183
8.1.3 线性回归模型的经典假设条件	185
8.1.4 回归分析的内容和分析步骤	185
§8.2 一元线性回归	185
8.2.1 一元线性回归的数学模型	185
8.2.2 参数 β_0 和 β_1 的最小二乘估计	186
8.2.3 最小二乘估计 $\hat{\beta}_0, \hat{\beta}_1$ 的性质	188
8.2.4 回归方程的显著性检验	188
8.2.5 预测和控制	192
§8.3 曲线回归	196
8.3.1 曲线回归的分析步骤	196
8.3.2 常用曲线的线性化方法	196
§8.4 多元线性回归	202
8.4.1 多元线性回归的数学模型	202
8.4.2 参数 β 的最小二乘估计	203
8.4.3 回归方程的显著性检验	204
8.4.4 回归系数的显著性检验	204
8.4.5 预测和控制	207
§8.5 质量控制应用案例分析	208
§8.6 SPSS 软件的使用	211
8.6.1 线性回归分析	211
8.6.2 曲线回归	213
习题八	215

*第 9 章 违背经典假设的经济计量模型	219
§9.1 引言	219
§9.2 异方差	220
9.2.1 异方差的概念	220
9.2.2 异方差产生的原因	222
9.2.3 异方差的后果	223
9.2.4 异方差的识别和检验	224
9.2.5 消除异方差的方法	227
§9.3 自相关	237
9.3.1 自相关的概念	237
9.3.2 产生自相关的原因	238
9.3.3 自相关的后果	240
9.3.4 自相关的识别和检验	240
9.3.5 自相关的处理方法	243
§9.4 多重共线性	249
9.4.1 多重共线性的概念	249
9.4.2 多重共线性的后果	250
9.4.3 产生多重共线性的原因	251
9.4.4 多重共线性的识别和检验	252
9.4.5 消除多重共线性的方法	253
习题九	259
*第 10 章 主成分分析	263
§10.1 主成分的概念	263
§10.2 主成分计算	264
§10.3 样本主成分	268
§10.4 主成分的选取	268
§10.5 案例分析	271
§10.6 主成分分析的上机实现	272
习题十	275
*第 11 章 因子分析	277
§11.1 因子分析模型	277
11.1.1 数学模型	277
11.1.2 因子模型的性质	278
§11.2 模型参数的统计意义	279
11.2.1 A 的元素 a_{ij} 的统计意义	279

11.2.2	A 的行元素平方和的统计意义	280
11.2.3	A 的列元素平方和的统计意义	280
§11.3	变量 $X_1, X_2, \dots, X_p$ 之间的相关性检验	280
§11.4	模型的参数估计方法	281
11.4.1	主成分法	281
11.4.2	主因子法	284
11.4.3	极大似然法	284
11.4.4	正交旋转	285
§11.5	因子得分	285
§11.6	因子分析的上机实现	287
	习题十	292
*第 12 章	聚类分析	293
§12.1	距离和相似性度量	293
12.1.1	距离	294
12.1.2	相似系数	297
§12.2	系统聚类法	298
12.2.1	最短距离法	299
12.2.2	最长距离法	300
12.2.3	中间距离法	300
12.2.4	类平均法	301
12.2.5	重心法	302
12.2.6	离差平方和法(Ward)	302
12.2.7	系统聚类法的统一	303
§12.3	动态聚类法	304
§12.4	有序样品的聚类	304
§12.5	聚类分析上机实现	306
12.5.1	系统聚类法	306
12.5.2	K-均值法	312
	习题十二	314
*第 13 章	判别分析	317
§13.1	距离判别	317
13.1.1	两总体情况	317
13.1.2	多总体情况	319
§13.2	Bayes 判别	320
13.2.1	贝叶斯判别准则	320

13.2.2	正态总体的贝叶斯判别	322
§13.3	Fisher 判别	323
§13.4	逐步判别	325
§13.5	SPSS 中判别分析方法和概念的介绍	325
13.5.1	建立判别函数的方法	325
13.5.2	典则判别函数	326
§13.6	判别分析的上机实现	327
13.6.1	二元变量的判别分析	327
13.6.2	多总体判别的情况	332
	习题十三	337
附表 1	泊松分布表	339
附表 2	标准正态分布表	341
附表 3	χ^2 分布表	343
附表 4	t 分布表	347
附表 5	F 分布表	349
附表 6	杜宾-瓦森检验临界值表	360
附表 7	常用正交表	363

第 1 章

统计数据的收集与整理

本章将介绍统计学的基本知识，统计数据的收集、调查的设计和数据的整理和表示的基本方法，以及使用 Office 软件 (Word、PowerPoint、Excel) 绘制各种统计图表。

§ 1.1 统计学概述

1.1.1 统计的广泛应用

在人类活动的各个方面都离不开统计工作和统计数据。

在个人生活中，人们的任何决策都依赖于有关的统计信息。无论是报考学校、选择工作单位、购买房屋、股票交易、外出旅游，还是购买日常生活用品、生活开支预算等，都离不开有关的统计信息。尽管个人生活中的许多统计信息是以非常简单粗糙的形式出现的，但正是在对过去发生的事情或经验所获得的信息进行综合的基础上，人们才能作出正确的判断和决策。

在政府的层次上，统计更是渗透于每一个部门的管理工作之中。政府部门通过不断收集经济、社会、人口等各个领域的统计数据，在综合分析的基础上对国民经济的发展进行预测、规划、指导和调控，并对全社会提供各种统计信息。

在科学和学术研究的各个领域，包括自然科学、经济学、社会学、体育、医疗卫生、环境保护等领域，都广泛使用统计学方法进行分析 and 推断。

在商务活动中，统计学有更多的应用，经营管理人员在作决策时需要应用统计方法来归纳分析各种可以获得的统计资料。商务活动中经常使用统计方法的有以下几个方面。

1. 财务分析

企业的财务报告和各種财务分析指标都是对企业的财务状况和经营成果的统计

分析结果；管理会计各种方法的应用需要以成本和收益为基础的经营业绩的统计分析；企业的投资和融资决策更离不开各种内外部统计信息的支持。

2. 产品开发

企业产品开发计划的制定，需要对经济发展趋势、商业竞争、顾客需求、财务收支估算等方面的数据进行统计分析，并在此基础上进行产品开发的经济可行性分析。

3. 计划

企业各种计划的制定，都需要对销售、资金、人力资源需求、成本和利润等因素进行预测，而预测就是以过去和当前的统计数据为依据，运用统计推断方法对未来情况进行的估计。

4. 市场研究

在市场研究中，需要对消费者的需求偏好及其变化趋势，竞争对手的情况，本企业产品和服务的顾客满意度等方面的数据进行统计分析。

5. 工序管理和质量控制

运用统计分析方法，可以帮助确定影响产品质量的主要因素，制定有效的质量控制标准和工艺规范，达到降低成本、提高生产效率的目标。质量控制和质量管理是企业经营管理中运用统计分析方法最多的领域。

6. 人力资源管理

在企业的人力资源管理中，需要经常使用统计方法来分析人事变动、出勤状况、工作业绩等情况，并在此基础上进行业绩评定、奖励和惩罚、制定有效的激励约束机制等。

1.1.2 统计与统计学

关于“统计”一词，可以有三种理解：统计工作、统计资料、统计学。

- 统计工作是指应用科学的方法对调查研究的对象进行数据搜集、整理和分析的全部工作，其成果是形成各种统计数据资料。
- 统计资料是统计工作所取得的成果，包括两大类型：原始资料和再生信息。
- 统计学是关于如何搜集、整理和分析客观现象数量规律的一门方法论的科学，是对统计工作实践加以总结升华而产生的理论，并用以指导统计实践。

1.1.3 统计研究对象的特点

统计研究的对象有以下 4 个方面的特点。

1. 数量性

统计学研究的对象是客观现象的数量特征和规律性。统计学是通过大量的观察方法来研究事物总体数量方面的特征及其规律性,反映客观事物在一定时间、地点等条件下的数量表现及其发展趋势和变化规律,为人类活动的各种决策提供依据。凡是产生数据的领域,都需要统计工作和统计学;同样,凡是能以数量来描述的事物,都可以作为统计学的研究对象。

2. 总体性

统计学研究的是客观现象总体的数量特征与规律性,而不是个体的量。总体由具有某种相同性质的一定范围内的全体事物组成。例如:对全国农村家庭收入情况进行统计分析,此时全国所有农村家庭的收入就构成研究的总体,但统计学对总体数量规律性的认识是通过对大量个体的观察和分析后获得的。

3. 具体性

统计学研究的对象是一定时间、地点等条件下具体事物的量,而不是抽象对象的量,这是统计学和数学的一个重要区别。因此,统计工作需要具体事物进行调查研究,收集和整理特定时间、地点等条件下所研究对象中大量个体的有关数据是统计分析的基础。

4. 差异(变异)性

组成统计研究对象总体的个体应是有差异的,否则就不需要进行统计分析。客观事物是错综复杂的,受到多种因素的影响,因而不同个体在数量方面必定存在差异,这就需要通过总体中大量的个体进行观察并进行综合分析,才能获得总体的数量特征,如平均值、方差、偏斜度等方面的分布特征。

1.1.4 统计学的分类

根据研究的重点和观察问题的角度不同,统计学主要有以下两种分类方法。

1. 描述统计学和推断统计学

这一分类方法既反映了统计学发展的两个主要阶段,同时也反映了各自不同的侧重。

描述统计学 研究如何对客观现象进行数量的计量、加工、概括和表示的方法。在20世纪之前统计学基本上处于描述阶段,描述统计学是统计学的基础。

推断统计学 研究如何根据样本数据来推断总体的分布情况,概率论是推断统计学的主要理论基础。推断统计学是近代统计学的核心,也是统计学中的主要内容。

2. 理论统计学和应用统计学

理论统计学 主要研究统计学的数学原理,它基于概率论的原理,还包括不属于传统概率论的一些内容,如随机化原理的理论、各种估计的原理、假设检验的原理以及一般决策的原理。在统计实践中经常会遇到一些原有的统计方法不能解决的新问题,需要创造新的统计模型和统计分析方法,这就需要统计理论的研究与指导。

应用统计学 将统计学的基本原理应用于各个领域就形成各种应用统计学的分支。它包括适用于各个领域的一般性的统计分析方法,如参数估计、假设检验、方差分析、回归分析等,还包括在某一领域中特定的统计分析方法,如经济领域中的指数分析法等。应用统计学需要既熟悉统计知识又熟悉某一领域业务知识的专门人才,它侧重于阐明统计的基本原理,并将理论统计学的结论作为工具应用于各个领域。

§ 1.2 数据的收集

统计数据的收集是统计整理和分析,以及统计推断和预测的基础。统计数据的收集就是根据统计研究的目的和要求,有组织、有计划地向调查对象搜集原始资料的过程。在实际应用中,我们要依据特定的研究目的或工作任务,结合研究对象所具有的性质和特点,相应地选择适合的调查方法,必要时也可以是几种调查方法结合使用。这里,我们介绍几种通常采用的调查方法,包括:普遍调查、重点调查、典型调查、抽样调查和网上调查。

1.2.1 普遍调查

普遍调查简称普查,是专门组织的一次性的全面调查。如全国的人口普查、能源普查、工业普查等。普查的组织方式一般有两种:一种是建立专门的普查机构,配备大量的普查人员;另一种是利用调查单位的原始记录和核算资料,发放调查表,由登记单位填报,如物资库存普查等。普查时必须注意以下几个原则。

(1) 规定统一的标准时点。标准时点是指对被调查对象登记时所依据的统一时点。

(2) 规定统一的普查期限。在普查范围内各调查单位或调查点尽可能同时进行登记。

(3) 规定普查的项目和指标。普查时必须按照统一规定的项目和指标进行登记,不准任意改变或增减,以免影响汇总和综合,降低资料质量。

【例 1.1】 2002 年在国务院统一部署下,开展全国“基本单位普查”。关于这次普查的部分重要要求与规定如下。

调查目的: 统计分析全国和各省市区单位的单位总数、性质、隶属关系、分布情况等。

标准时间: 2001 年 12 月 31 日。

调查对象：各类法人单位，各类法人单位所属的产业活动单位。

调查项目：单位总数；

性质分类(企业法人、事业法人、社团法人、机关法人、其他法人)；

基本情况——性质划分、隶属关系(中央、省市、区县、街道等)；

主要构成(产业结构、行业分布、地区分布、经济成分、规模结构)；

分布情况。

1.2.2 重点调查

重点调查是专门组织的一种非全面调查，它是在总体中选择个别的或部分重点单位进行调查，籍以了解总体的基本情况。这些单位虽然少，但它们调查的标志值在总体标志总量中占有绝大比重，通过对这些单位的调查，就能掌握总体的基本情况。例如，要了解全国钢铁企业的生产情况，只要调查宝钢、鞍钢、马钢、包钢、首钢等大型钢铁公司，就能达到调查的目的。因为这些钢铁企业虽在全国只占少数，但它们的产量在全国钢产量中占有绝大部分的比重。因此当调查的目的只要掌握调查对象的基本情况，而在总体中却有部分单位能较集中地反映所研究的问题时，采用重点调查是比较合适的。

重点调查的特点：

(1) 重点调查适用于调查对象的标志值比较集中于某些单位的场合，这些单位的管理比较健全，统计力量比较充实，能够及时取得准确资料。

(2) 重点调查的目的在于了解总体现象某些方面的基本情况，而不要求全面准确地推算总体数字。

(3) 重点调查比实际调查的单位数目少，在满足调查目的所要求的前提下，可以比全面调查节省人力、物力和时间。

1.2.3 典型调查

典型调查是根据调查的目的，选择在同类对象中最具典型性的部分和个体进行调查。典型调查也是一种非全面调查。例如，选择第一汽车制造厂作为国有企业改革情况调查分析的样板。典型调查的作用如下：

- 典型调查可用来研究新生事物。
- 典型调查可用来研究事物变化的规律。
- 典型调查可用来分析事物的不同类型，研究它们之间的差别和相互关系。
- 典型调查的资料可用来补充和验证全面统计的数字，推论和测算有关现象的总体。