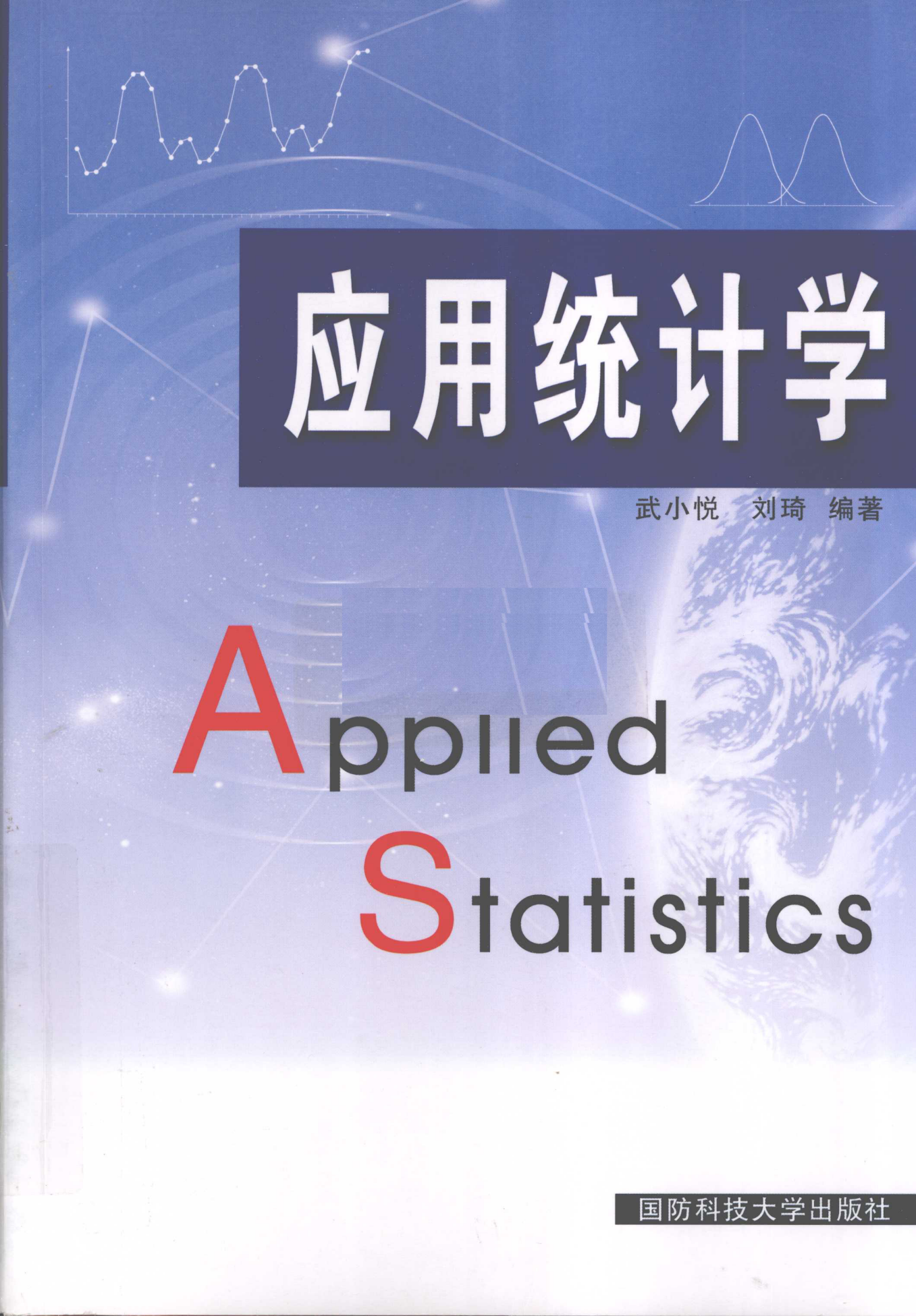




应用统计学

武小悦 刘琦 编著

Applied Statistics

国防科技大学出版社

应用统计学

武小悦 刘琦 编著

国防科技大学出版社

·长沙·

内 容 简 要

本书系统深入地介绍了应用统计学的基本概念、理论和方法。全书共分为 12 章,包括:绪论,描述性统计,参数估计与假设检验,方差分析,回归分析,时间序列分析,统计指数,统计调查,抽样技术,非参数检验,主成分分析与因子分析,判别分析与聚类分析。

本书内容全面,结构清晰,注重对统计分析思想的阐述,相关理论方法表述严谨。各章都编配了小结、中英文关键词、思考题,习题及统计实验题。

本书是主要面向管理类专业高年级本科生编写的教材,也可作为非数学专业本科生或研究生相关课程的教学参考书。

图书在版编目(CIP)数据

应用统计学/武小悦,刘琦编著. —长沙:国防科技大学出版社,2009.10
ISBN 978-7-81099-717-1

I. 应… II. ①武… ②刘… III. 应用统计学—高等学校—教材 IV. C8

中国版本图书馆 CIP 数据核字(2009)第 187675 号

国防科技大学出版社出版发行

电话:(0731)84572640 邮政编码:410073

<http://www.gfkdcbs.com>

责任编辑:耿 筠 责任校对:石少平

新华书店总店北京发行所经销

国防科技大学印刷厂印装

*

开本:787×1092 1/16 印张:47.5 字数:1126 千

2009 年 10 月第 1 版第 1 次印刷 印数:1-1000 册

ISBN 978-7-81099-717-1

定价:76.00 元

前 言

自 2005 年以来,我们一直担任国防科技大学管理类专业本科生“应用统计学”课程不同教学班的主讲教师。上课的学生都已学过公共基础课“概率论与数理统计”,具备了初步的数理统计基础,但还需要进一步学习掌握针对实际问题进行数据收集、分析、表述和解释的统计理论与方法。目前,多数国内外教材都没有充分考虑这一类学生学习“应用统计学”的基础和需求特点。作者在教学中一直感到难以在市面上找到针对这一类学生的适用教材。为了解决这一问题,我们在多年来教学实践的基础上编写了这本教材。

本书是主要面向管理类专业高年级本科生编写的教材,也可作为非数学专业本科生或研究生相关课程的教学参考书。本书注重对统计分析思想的阐述,相关理论方法表述严谨,并尽可能给出了应用示例。各章都编配了内容小结、中英文关键词、思考题,习题及统计实验题。当今应用统计学教学的明显特征就是在统计思想和理论方法讲解的同时,结合统计软件进行教学实践。在编写本书时,我们曾想结合 SPSS 或其他统计分析软件进行介绍,但考虑到本书篇幅所限以及各教学单位使用的统计分析软件不同,因此,最终放弃了这一尝试。建议在使用本书时,根据实际情况选择合适的统计软件实施教学,以提高学生运用统计分析方法解决实际问题的能力。

本书初稿曾以补充讲义的形式在 2008 年课程教学中进行了试用。本书的出版得到了国防科技大学本科自编教材建设项目经费支持,在编写和出版过程中,得到了国防科技大学信息系统与管理学院、国防科技大学出版社的大力支持。本书编写中参考了大量的国内外相关文献,主要的参考文献已列在各章之后。在此,向上述单位、个人和参考文献作者表示衷心的感谢。

本书第 1、5、6、7、11 章及第 12 章的聚类分析由武小悦编写，第 2、3、4、8、9、10 章及第 12 章的判别分析由刘琦编写。在编写过程中，我们共同研讨各章内容，并对书稿进行了多次相互修改、完善和认真校对。由于作者水平和时间有限，书中可能还会有不妥之处，欢迎广大师生和读者提出宝贵意见。

武小悦，刘琦

2009 年 9 月于长沙

目 录

第1章 绪 论

1.1 统计与统计学	(1)
1.1.1 统计的内涵	(1)
1.1.2 统计学的重要性	(2)
1.1.3 统计学的分科	(3)
1.1.4 统计学与其他学科的关系	(4)
1.1.5 统计工作的基本职能	(5)
1.2 统计学的产生与发展	(6)
1.2.1 古典统计学时期	(6)
1.2.2 近代统计学时期	(7)
1.2.3 现代统计学时期	(8)
1.2.4 统计在我国的发展	(8)
1.3 统计学的若干基本概念	(9)
1.3.1 统计数据类型及来源	(9)
1.3.2 总体与样本	(10)
1.3.3 参数与统计量	(11)
1.3.4 标志与指标	(12)
1.3.5 变量及其分类	(13)
1.3.6 变量的测量水平	(13)
1.4 统计工作的过程	(15)
1.5 常用的统计分析软件	(17)
1.6 统计学的学习	(20)
小结	(21)
关键词	(21)
思考题	(22)
习题	(22)
实验题	(23)
参考文献	(23)

第2章 描述性统计

2.1 数据的预处理	(25)
2.1.1 数据的审核	(25)
2.1.2 数据的筛选	(26)
2.1.3 数据的排序	(26)
2.1.4 数据的分组	(27)
2.2 数据的位置特征	(33)
2.2.1 平均数	(33)
2.2.2 中位数	(35)
2.2.3 众数	(38)
2.2.4 平均数、中位数与众数的比较	(41)
2.3 数据的离散特征	(43)
2.3.1 极差	(43)
2.3.2 四分数间距	(43)
2.3.3 方差与标准差	(45)
2.3.4 平均绝对离差	(46)
2.3.5 变异系数	(47)
2.3.6 数据离散特征的比较	(48)
2.4 分布的形态特征	(49)
2.4.1 偏倚程度	(49)
2.4.2 高陡程度	(51)
2.5 样本数据的图形描述	(52)
2.5.1 茎叶图	(52)
2.5.2 五数概括与箱线图	(57)
2.5.3 相关关系与散点图	(61)
2.5.4 P-P图和Q-Q图	(65)
小结	(69)
关键词	(70)
思考题	(71)
习题	(71)
实验题	(73)
参考文献	(76)

第3章 参数估计与假设检验

3.1 估计与检验概述	(78)
3.1.1 估计与检验的基本概念	(78)
3.1.2 常用分布介绍	(82)
3.2 总体参数的点估计	(86)
3.2.1 矩估计法	(86)
3.2.2 极大似然法	(88)
3.2.3 估计量的评价	(90)
3.3 区间估计	(92)
3.3.1 区间估计的基本概念	(92)
3.3.2 总体均值的置信区间	(95)
3.3.3 总体方差的置信区间	(97)
3.3.4 总体比例的置信区间	(98)
3.3.5 两总体均值之差的置信区间	(99)
3.3.6 两总体方差之比的置信区间	(101)
3.4 假设检验	(103)
3.4.1 显著性检验与 P 值检验	(103)
3.4.2 一个总体均值的假设检验	(107)
3.4.3 一个总体比例的假设检验	(110)
3.4.4 一个总体方差的假设检验	(111)
3.4.5 两个总体均值的假设检验	(113)
3.4.6 两个总体方差的假设检验	(115)
3.5 Bayes 估计与检验	(117)
3.5.1 Bayes 公式	(118)
3.5.2 验前分布	(120)
3.5.3 Bayes 点估计	(126)
3.5.4 Bayes 区间估计	(130)
3.5.5 Bayes 假设检验	(131)
小结	(134)
关键词	(135)
思考题	(136)
习题	(137)
实验题	(138)

参考文献	(139)
------	-------

第4章 方差分析

4.1 基本概念和原理	(140)
4.1.1 方差分析的基本概念	(140)
4.1.2 方差分析的基本原理	(144)
4.2 单因素方差分析	(145)
4.2.1 单因素方差分析模型	(145)
4.2.2 固定效应模型	(146)
4.2.3 多重比较	(154)
4.2.4 方差分析模型的检验	(158)
4.2.5 随机效应模型	(160)
4.3 无交互作用的双因素方差分析	(163)
4.3.1 双因素方差分析模型	(163)
4.3.2 固定效应模型	(165)
4.3.3 随机效应模型和混合效应模型	(175)
4.4 有交互作用的双因素方差分析	(180)
4.4.1 固定效应模型	(180)
4.4.2 随机效应模型和混合效应模型	(187)
4.5 试验设计概述	(195)
4.5.1 试验设计的基本概念和原则	(195)
4.5.2 随机化完全区组设计	(196)
4.5.3 因子设计	(199)
小结	(200)
关键词	(201)
思考题	(202)
习题	(202)
实验题	(206)
参考文献	(207)

第5章 回归分析

5.1 一元线性回归	(210)
5.1.1 一元线性回归模型	(210)
5.1.2 一元线性回归模型参数的求解	(212)

5.1.3 最小二乘参数估计的性质	(214)
5.1.4 回归参数的方差	(215)
5.1.5 一元线性回归模型的检验	(216)
5.1.6 回归模型假设的检验	(221)
5.1.7 一元回归预测	(225)
5.2 多元线性回归	(226)
5.2.1 多元线性回归模型	(226)
5.2.2 多元线性回归模型参数的求解	(228)
5.2.3 最小二乘参数估计的性质	(231)
5.2.4 多元线性回归模型的检验	(234)
5.2.5 多元线性回归模型假设的检验	(237)
5.2.6 多重共线性问题	(245)
5.2.7 多元回归预测	(247)
5.3 非线性回归模型	(248)
5.3.1 Logistic 回归	(249)
5.3.2 其他非线性回归模型	(254)
小结	(259)
关键词	(260)
思考题	(261)
习题	(261)
实验题	(266)
参考文献	(267)

第 6 章 时间序列分析

6.1 时间序列的基本概念	(269)
6.1.1 时间序列的定义及分类	(269)
6.1.2 时间序列的编制原则	(273)
6.1.3 时间序列的影响因素	(273)
6.2 时间序列的指标分析	(274)
6.2.1 时间序列的水平指标	(275)
6.2.2 时间序列的速度指标	(281)
6.3 时间序列的分解	(285)
6.3.1 时间序列的分解模型	(285)
6.3.2 长期变动趋势的分析	(286)

6.3.3 季节变动的分析	(294)
6.3.4 循环变动的分析	(301)
6.3.5 不规则变动的分析	(302)
6.4 时间序列的预测	(303)
6.4.1 趋势外推法	(303)
6.4.2 移动平均法	(305)
6.4.3 指数平滑法	(310)
6.4.4 预测方法的评价	(313)
小结	(314)
关键词	(315)
思考题	(316)
习题	(316)
实验题	(318)
参考文献	(319)

第7章 统计指数

7.1 统计指数概述	(320)
7.1.1 统计指数的概念	(320)
7.1.2 统计指数的性质和作用	(322)
7.1.3 统计指数的种类	(324)
7.2 指数的编制	(327)
7.2.1 综合指数的编制	(327)
7.2.2 平均指数的编制	(330)
7.2.3 平均指标指数的编制	(333)
7.3 指数体系和因素分析	(336)
7.3.1 指数体系的概念	(336)
7.3.2 综合指数的因素分析	(337)
7.3.3 平均指标指数的因素分析	(341)
7.3.4 指数数列	(342)
7.4 价格统计指数	(343)
7.4.1 居民消费价格指数和商品零售价格指数	(343)
7.4.2 股票价格指数	(344)
7.5 综合评价指数	(346)
7.5.1 综合评价的基本过程	(346)

7.5.2 单项指标的标准化方法	(347)
7.5.3 综合评价指数的综合方法	(349)
7.5.4 权重的确定方法	(349)
7.5.5 社会发展综合评价指数	(351)
小结	(353)
关键词	(354)
思考题	(355)
习题	(355)
实验题	(356)
参考文献	(357)

第8章 统计调查

8.1 统计调查概述	(358)
8.1.1 统计调查的基本要求	(358)
8.1.2 数据的来源	(359)
8.1.3 统计数据的误差	(360)
8.1.4 统计调查的流程	(362)
8.2 统计调查的实施	(364)
8.2.1 组织方式	(364)
8.2.2 统计调查方法	(371)
8.3 统计表	(373)
8.3.1 统计表的作用	(374)
8.3.2 统计表的设计	(374)
8.3.3 统计表的分类	(376)
8.4 调查问卷	(380)
8.4.1 调查问卷及问题的分类	(380)
8.4.2 量表设计	(381)
8.4.3 问卷的基本结构	(383)
8.4.4 问卷设计的原则	(384)
8.4.5 问卷设计流程	(387)
8.5 信度与效度分析	(389)
8.5.1 信度	(389)
8.5.2 效度	(394)
8.5.3 信度和效度的关系	(396)

8.6 统计整理	(397)
小结	(406)
关键词	(407)
思考题	(408)
习题	(408)
实验题	(410)
参考文献	(410)

第9章 抽样技术

9.1 抽样概述	(412)
9.1.1 抽样的相关概念	(412)
9.1.2 总体参数与样本参数	(415)
9.1.3 抽样分布	(416)
9.1.4 抽样误差	(418)
9.2 简单随机抽样	(419)
9.2.1 简单随机抽样的基本概念	(419)
9.2.2 简单随机抽样的实施方法	(421)
9.2.3 不放回简单随机抽样的参数估计	(424)
9.2.4 放回简单随机抽样的参数估计	(430)
9.2.5 简单随机抽样的设计效应	(431)
9.2.6 简单随机抽样的样本量确定	(431)
9.3 分层抽样	(435)
9.3.1 分层抽样与分层随机抽样	(435)
9.3.2 分层随机抽样的相关符号说明	(437)
9.3.3 分层随机抽样的参数估计	(438)
9.3.4 分层随机抽样的样本量分配	(442)
9.3.5 事后分层	(445)
9.3.6 分层随机抽样的样本量确定	(447)
9.3.7 分层随机抽样的设计效应	(452)
9.4 整群抽样	(454)
9.4.1 整群抽样的特点与群的划分原则	(454)
9.4.2 群规模相等时的整群抽样	(455)
9.4.3 群规模不等时的整群抽样	(460)
9.4.4 不等概率抽样	(468)

9.4.4 不等概率抽样	(468)
9.4.5 整群抽样的设计效应	(478)
9.5 系统抽样	(482)
9.5.1 系统抽样的特点	(482)
9.5.2 系统抽样的实施方法	(482)
9.5.3 估计量与参数估计	(486)
小结	(491)
关键词	(492)
思考题	(493)
习题	(493)
实验题	(497)
参考文献	(497)

第 10 章 非参数检验

10.1 非参数检验概述	(499)
10.1.1 非参数检验的特点	(499)
10.1.2 秩统计量	(500)
10.1.3 非参数检验的类型	(501)
10.2 符号检验	(507)
10.2.1 检验的基本原理和步骤	(507)
10.2.2 配对样本的符号检验	(509)
10.2.3 McNemar 检验	(512)
10.3 秩和检验	(514)
10.3.1 检验的基本原理和步骤	(514)
10.3.2 大样本的秩和检验	(519)
10.3.3 Mann-Whitney U 检验	(522)
10.4 符号秩检验	(524)
10.4.1 检验的基本原理和步骤	(524)
10.4.2 配对样本的符号秩检验	(528)
10.5 Kruskal-Wallis 检验	(531)
10.5.1 检验的基本原理和步骤	(531)
10.5.2 大样本的 Kruskal-Wallis 检验	(536)
10.6 Friedman 检验	(537)
10.6.1 检验的基本原理和步骤	(537)

10.6.2	大样本的 Friedman 检验	(542)
10.7	Cochran Q 检验	(543)
10.8	游程检验	(548)
10.8.1	检验的基本原理和步骤	(548)
10.8.2	两个总体分布的游程检验	(549)
10.9	χ^2 检验	(550)
10.9.1	χ^2 拟合优度检验	(550)
10.9.2	列联表的独立性检验	(554)
10.10	Kolmogorov-Smirnov 检验	(557)
10.10.1	Kolmogorov 检验	(557)
10.10.2	Smirnov 检验	(559)
10.11	Spearman 秩相关系数检验	(560)
10.11.1	Spearman 秩相关系数的计算	(560)
10.11.2	检验的原理与步骤	(563)
10.12	Kendall τ 相关系数检验	(564)
10.12.1	Kendall τ 相关系数的计算	(564)
10.12.2	检验的原理与步骤	(567)
小结		(568)
关键词		(569)
思考题		(570)
习题		(570)
实验题		(575)
参考文献		(576)

第 11 章 主成分分析与因子分析

11.1	主成分分析概述	(577)
11.1.1	主成分分析的基本思想	(577)
11.1.2	主成分的定义	(578)
11.2	主成分分析模型的求解	(580)
11.2.1	相关的矩阵分析理论	(580)
11.2.2	总体主成分的求解及其性质	(586)
11.2.3	主成分与多元正态分布	(588)
11.2.4	样本主成分模型的求解	(591)
11.3	主成分分析的应用	(595)

11.3.2	主成分分析的应用步骤	(595)
11.3.3	应用示例	(598)
11.4	因子分析概述	(601)
11.4.1	因子分析的基本思想	(601)
11.4.2	公共因子分析模型	(603)
11.4.3	正交因子模型	(604)
11.5	因子分析模型的求解	(607)
11.5.1	主成分法	(607)
11.5.2	极大似然法	(611)
11.5.3	因子旋转	(615)
11.6	因子分析的应用	(619)
11.6.1	因子得分计算	(619)
11.6.2	因子分析的步骤	(622)
11.6.3	应用示例	(625)
小结		(632)
关键词		(633)
思考题		(634)
习题		(635)
实验题		(636)
参考文献		(642)

第 12 章 判别分析与聚类分析

12.1	距离判别法	(643)
12.1.1	距离判别法的原理	(644)
12.1.2	各总体协方差阵相等时的判别准则	(645)
12.1.3	错判概率的估计	(649)
12.2	Fisher 判别法	(650)
12.2.1	Fisher 判别法的基本思想	(650)
12.2.2	Fisher 判别法的主要步骤	(653)
12.3	Bayes 判别法	(656)
12.3.1	Bayes 判别法的基本原理	(656)
12.3.2	正态等协方差且等损失下的 Bayes 判别	(658)
12.4	逐步判别	(660)
12.4.1	判别能力的度量与检验	(660)

12.4.2	逐步判别的基本原理	(663)
12.4.3	变量筛选过程	(665)
12.5	聚类分析概述	(670)
12.5.1	聚类分析的作用	(670)
12.5.2	聚类分析的基本模型	(670)
12.6	相似性度量	(671)
12.6.1	距离与相似性	(672)
12.6.2	常见的距离	(673)
12.6.3	定性变量的距离	(675)
12.6.4	变量的相似性度量	(676)
12.7	系统聚类法	(677)
12.7.1	系统聚类法概述	(677)
12.7.2	若干实现算法	(678)
12.7.3	系统聚类法的应用	(686)
12.8	动态聚类法	(692)
12.8.1	初始凝聚点的确定	(692)
12.8.2	K -均值聚类	(693)
12.9	有序对象的分类	(696)
12.9.1	有序对象分类的算法	(696)
12.9.2	应用示例	(697)
小结		(701)
关键词		(702)
思考题		(703)
习题		(703)
实验题		(709)
参考文献		(713)
附 录		(714)