## 附件3

## 《应用多元统计分析》教学大纲

**一、基本信息**

|  |  |  |  |
| --- | --- | --- | --- |
| **课程编号**： | 1042019 | **课程性质**： | 学位专业基础课 |
| **课程名称**： | 应用多元统计分析 | **英文名称**： | Applied Multivariate Statistical Analysis |
| **学时/学分**： | 32/2 | **开课时间**： | 第1学期 |
| **适用对象**： | **应用统计专业硕士** | | |
| **大纲执笔人**： | 陈钰芬 | **当前版本**： | **2021** |

**二、课程概述**

随着大数据、云计算和人工智能时代的来临，一门集数学、统计学和计算机科学为一体的数据科学在全世界范围内迅速兴起。多元统计分析方法成为处理多变量数据不可或缺的重要技术和方法，是大数据分析的重要工具。应用多元统计分析主要介绍几种重要的定量分析研究方法。应用面极广，可广泛应用于地区经济社会发展综合评价、企业竞争力比较、市场研究、生产优化、质量管理等等。对金融、医疗卫生、生产管理、政府、科研和教育等领域都有重要的应用价值。学好这门课，不仅有助于提高同学的综合分析能力，实际应用能力，而且能直接服务于社会，为宏观和微观各层决策者、管理者提供依据。

**三、教学目标**

通过本课程的理论教学和相关实验训练，使学生具备如下能力：

1、掌握数据分析处理的基本思想和数据准备的基本技能

2、掌握主成分分析的基本思想、数学原理，能够运用SAS软件（或Python软件）对实际数据进行分析。

3、掌握因子分析的基本思想、数学原理，能够运用SAS软件（或Python软件）对实际数据进行分析。

4、掌握典型相关分析的基本思想、数学原理，能够运用SAS软件（或Python软件）对实际数据进行分析。

5、掌握对应分析的基本思想、数学原理，能够运用SAS软件（或Python软件）对实际数据进行分析。

6、掌握聚类分析的基本思想、数学原理，能够运用SAS软件（或Python软件）对实际数据进行分析。

7、掌握判别分析的基本思想、数学原理，能够运用SAS软件（或Python软件）对实际数据进行分析。

8、能够综合运用多种方法进行数据分析，解决实际问题。

**四、课程内容**

**第1章 绪论**

**重点内容：**数据的组织，协方差矩阵和相关系数矩阵，矩阵的秩、特征值及特征向量。

**难点内容：**协方差矩阵和相关系数矩阵，矩阵的秩、特征值及特征向量。

**教学内容：**掌握数据的组织，掌握协方差矩阵、相关系数矩阵的定义及相互之间的关系；掌握矩阵的运算；行列式；理解矩阵的逆、秩；特征值及特征向量。

1.1多元统计分析概述

1.2描述统计量

1.3 矩阵代数

**第2章 主成分分析**

**重点内容：**主成分的几何解释，主成分的计算，主成分的含义解释，主成分的应用。

**难点内容：**主成分的含义解释，主成分的应用。

**教学要求：**了解主成分分析的基本思想，明确主成分分析的基本步骤，掌握主成分的计算、主成分的含义解释，能熟练用SAS软件或 Python软件进行分析，并能运用主成分分析法分析解决实际问题。

2.1主成分分析的基本思想

2.2主成分的计算

2.3主成分的性质

2.4主成分的应用

**第3章 因子分析**

**重点内容：**初始因子载荷矩阵，因子旋转，公共因子的含义解释，因子得分的计算。

**难点内容：**因子旋转，公共因子的含义解释。

**教学要求：**了解因子分析的基本思想，明确因子分析与主成分分析的区别与联系；明确因子分析的基本步骤，对各公共因子的含义解释正确；能用SAS软件或 Python软件进行分析，并能运用因子分析法分析解决实际问题。

3.l因子分析的基本思想

3.2因子分析的数学模型

3.3因子分析的计算

3.4因子分析的应用

**第4章 典型相关分析**

**重点内容：**典型变量的求解，典型变量的性质，典型变量的含义解释，典型冗余分析，典型相关系数的显著性检验。

**难点内容：**典型变量的求解，典型变量的含义解释，典型冗余分析。

**教学要求：**明确典型相关分析的基本原理和思想，明确典型相关分析与主成分分析的区别与联系，能运用SAS软件或 Python软件分析解决实际问题。

4.l典型相关分析的基本思想

4.2典型变量的求解

4.3典型变量的性质

4.4典型变量的含义解释

4.5典型冗余分析

4.6典型相关系数的显著性检验

**第5章 对应分析**

**重点内容：**对应分析的数学原理，行轮廓、列轮廓，独立性检验和总惯量，对应分析图

**难点内容：**对应分析的数学原理，对应分析图。

**教学要求：**理解对应分析的基本思想，掌握对应分析的数学原理，明确行轮廓、列轮廓的计算及含义，掌握对应分析图的解释，能运用SAS软件或 Python软件分析解决实际问题。

5.l对应分析的基本思想

5.2对应分析的数学原理

5.3对应分析图的解释

5.4对应分析的应用

**第6章 聚类分析**

**重点内容：**样品间的距离，系统聚类法的基本思想，几种系统聚类方法：最短距离法、最长距离法、中间距离法、类平均法、重心法、离差平方和法（Ward法），系统聚类法的基本步骤，动态聚类法的基本思想和步骤，凝聚点的确定。

**难点内容：**各种系统聚类法，动态聚类法凝聚点的确定。

**教学要求：**明确聚类分析的基本概念，了解系统聚类法和动态聚类法的基本思想，熟悉聚类分析的全过程；能运用SAS软件或 Python软件分析解决实际分类问题。

6.l聚类分析的基本概念

6.2距离和相似系数

6.3系统聚类法

6.4动态聚类法

**第7章 判别分析**

**重点内容：**判别分析与聚类分析的比较，多个总体协方差矩阵相等时距离判别函数的计算及判别准则，Bayes判别的基本思想，Fisher判别的基本思想、基本步骤和判别函数的计算。

**难点内容：**多个总体协方差矩阵相等时判别函数的计算及判别准则，Fisher判别的基本思想、基本步骤和判别函数的计算。

**教学要求：**明确判别分析的基本原理和思想，明确判别分析与聚类分析的区别与联系，能运用SAS软件或 Python软件分析解决实际问题。

7.1判别分析的基本思想

7.2距离判别法

7.3 Bayes判别

7.4 Fisher判别

**五、教学安排**

该课程每周2学时，16周，20学时为课堂授课教学时间，12学时为课内实验，同时开设开放实验。建议教学进度如下：

|  |  |
| --- | --- |
| **章节** | **学时数** |
| 第1章 绪论 | 2 |
| 第2章 主成分分析 | 6 |
| 第3章 因子分析 | 6 |
| 第4章 典型相关分析 | 3 |
| 第5章 对应分析 | 3 |
| 第6章 聚类分析 | 6 |
| 第7章 判别分析 | 6 |

**六、授课方式与手段**

以课堂理论教学为主，实验实践为辅。在教学过程中注重能力的培养，以实际应用系统为例，提高理论教学实用性，提高学生分析和解决实际问题的能力。鼓励学生参加开放实验，在平时SAS（或Python）实验作业基础上进一步提高软件的操作运用能力。

**七、考核方式及成绩评定**

**考核方式**：期末上机考试，平时作业、出勤、课堂情况，SAS实验作业。

**成绩评定标准**：总成绩（百分制）＝平时成绩×30％＋期末考试成绩×70％。

**八、教材及主要参考书**

|  |
| --- |
| **指定教材：** |
| 陈钰芬，陈骥.多元统计分析，清华大学出版社，2020年 |
| **参考书目：** |
| [1] 何晓群编著，多元统计分析（第五版），中国人民大学出版社，2019.06 |
| [2] 谢龙汉,尚涛编著，SAS统计分析与数据挖掘，电子工业出版社，2013.1 |
| [3] [美]Richard A.Johnson,Dean W.Wichern著，陆璇 等译，实用多元统计分析，清华大学出版社，2008.11 |