河北经贸大学课程水平认定

《多元统计分析》课程大纲

|  |  |  |  |
| --- | --- | --- | --- |
| **课程名称** | 多元统计分析 | **课程类型** | 必修课 |
| **总 学 时** | 32学时 | **学 分** | 2学分 |
| **适用专业** | 统计学 | **开课单位** | 数统学院 |

**一、课程性质**

多元统计分析是统计学的一个重要分支，是处理多维数据不可缺少的重要工具，随着电子计算机的普及和发展，多元统计分析方法已愈来愈多地应用于社会经济各个方面的数据分析之中。多元统计分析是利用统计学和数学方法，将隐没在大规模原始数据群体中的重要信息集中提炼出来，简明扼要的把握系统的本质特征，分析数据系统中的内在规律性。利用多元分析中不同的方法还可以对研究对象进行分类和简化。多元分析是实现做定量分析的有效工具。

**二、学习目的**

通过本课程的学习，让学生会应用多元统计分析中的诸多方法进行数据分析，通过和不同的学科知识相结合，对所考虑具体问题给出合理的推断。

**三、学习要求**

要求学生掌握各种判别分析、聚类分析、主成分分析、相关分析和因子分析等各种多元分析方法的思想及统计分析方法。

**四、学习内容及学时分配**

课程内容与学时分配

|  |  |  |
| --- | --- | --- |
| **课 程 内 容** | **学时分配** | |
| **自学** | **自学研讨调研** |
| **第一讲 应用****多元统计方法简介** | **2** | **2** |
| **第二讲** **主成分分析** | **2** | **2** |
| **第三讲 因子分析** | **2** | **4** |
| **第四讲 判别分析** | **2** | **2** |
| **第五讲 Logistic回归模型** | **2** | **4** |
| **第六讲 聚类分析** | **2** | **2** |
| **第七讲 多元回归分析** | **2** | **2** |

**五、课程考核及成绩评定**

课程考核为闭卷考试。

成绩评定：考试成绩实行百分制，其中基础知识测试题的分值掌握在40分左右；综合能力测试题的分值掌握在60分左右。60分为及格。

**六、推荐教材和学习参考书**

|  |  |  |
| --- | --- | --- |
| **推荐教程** | **作者** | **出版社及出版时间** |
| 实用多元统计分析 | DALLASE.JOHNSON | 高等教育出版社，2008 |
| **学习参考书** | **作者** | **出版社及出版时间** |
| 应用多元统计分析 | 李春林、陈旭红 | 清华大学出版社，2013 |
| 多元统计分析 | 于秀林﹑任雪松 | 中国统计出版社，1999 |
| 应用多元分析 | 王学民 | 上海财经大学出版社，1999 |
| SPSS for Windows统计分析 | 卢纹岱 | 电子工业出版社，2001 |

**七、学习具体内容和要求**

**第一讲 应用多元统计方法简介**

**一、基本要求**

要求学生对多元统计分析课程有一个概括的认识。

**二、授课方法**

自学。

**三、学习内容**

（一） 简述各种多元统计方法

简单介绍了主成分分析、因子分析、判别分析、典型判别分析、罗吉斯回归分析、聚类分析、多变量方差分析、典型变量分析、典型相关分析等方法。

1. 两个例子

介绍研究个体的独立性。

1. 变量的类型

（四） 数据矩阵和向量

介绍变量的数值、数据矩阵、数据向量及数据的下标符号。

（五） 多元正态分布

本节主要介绍关于多元正态分布的定义、均值向量、方差-协方差矩阵、相关矩阵、多元正态分布的密度函数以及典型的二元正态分布。

（六） 统计计算

本节主要介绍计算机的使用、缺失值的处理、取样的策略、数据的输入错误以及如何校正。

（七） 多变量的异常值

本节主要介绍如何确定异常值、处理异常值以及异常值的影响。

**四、重点难点**

多元统计分析的概念及它能解决哪些类型的实际问题。

**五、思考与讨论**

1﹑什么是多元统计分析？

2﹑多元统计分析能解决哪些类型的实际问题？

**第二讲 主成分分析**

**一、基本要求**

要求学生了解主成分分析的基本原理及其作用，掌握主成分分析的方法，能够应用这一方法分析数据，解决实际问题。

**二、授课方法**

自学

**三、学习内容**

（一）主成分分析的目的

（二）基于方差-协方差矩阵的主成分分析

本节主要介绍如何求主成分得分和因子载荷向量。

（三）主成分得分的估计

（四）两种判断主成分个数的方法

（五）基于相关矩阵的主成分分析

本节主要介绍基于相关矩阵的主成分得分、因子相关向量、样本相关矩阵以及如何确定主成分的个数。

（六）检验原始变量的独立性

（七）统计计算软件

本节主要介绍如何使用SAS和SPSS做主成分分析。

**四、重点难点**

主成分分析的作用及其应用，贡献率、累计贡献率的意义及其在实际应用中的作用。

**五、思考与讨论**

1﹑主成分分析的几何意义是什么？

2﹑主成分分析的主要作用有那些？

3﹑什么是贡献率和累计贡献率，其意义何在？

4﹑为什么说贡献率和累计贡献率能反映主成分中所包含的原始变量的信息？

5﹑为什么要用标准化数据去估计V的特征向量与特征值？

6﹑证明：对于标准化数据有S=R。

7﹑主成分分析在SPSS中如何实现？

**第三讲 因子分析**

**一、基本要求**

要求学生了解因子分析模型，理解因子载荷阵的统计意义，了解方差旋转的作用，掌握用因子分析模型分析研究实际问题的能力。

**二、授课方法**

自学

**三、学习内容**

（一）因子分析的目的

（二）因子分析模型

本节主要介绍因子分析的一些假设、因子分析模型的矩阵形式以及一些因子分析术语的定义。

（三）因子分析方程以及求解方程

（四）如何选择理想的因子数目

本节主要介绍两种方法：主观标准方法和客观标准方法。

（五）计算因子分析方程的方法

本节主要介绍基于相关矩阵的主因子法和迭代法。

（六）因子旋转

本节举例说明因子旋转的方法、方差最大化旋转方法、斜交旋转方法。

（七）因子得分

本节主要介绍巴特利特方法、回归方法及一些特定的方法。

**四、重点难点**

了解因子模型及因子载荷阵的统计意义，掌握因子分析的方法。

**五、思考与讨论**

1﹑因子得分模型与主成分分析模型有何不同？

2﹑因子载荷阵的统计意义是什么？

3﹑方差旋转的目的是什么？

4﹑因子分析有何作用？

5﹑因子模型与回归模型有何不同？

6﹑在SPSS中如何实现因子分析？

**第四讲 判别分析**

**一、基本要求**

要求学生了解判别分析的基本思想和几种常用判别分析方法，能够用这些方法分析解决实际问题。

**二、授课方法**

自学

**三、学习内容**

（一）二维正态总体的判别分析

（二）两个总体的价值函数和先验概率

（三）一般的判别准则

（四）多个总体的判别准则

（五）典型的判别函数

（六）近邻判别分析法

（七）分类树

**四、重点难点**

判别分析的基本思想及几种判别分析方法的实现与应用。

**五、思考与讨论**

1﹑简述距离判别法的基本思路，图示其几何意义。

2﹑判别分析与聚类分析有何异同？

3﹑简述贝叶斯判别的基本思路。

4﹑简述费歇判别的基本思路。

5﹑简述逐步判别法的基本思想。

6﹑在SPSS软件中如何实现判别分析？

**第五讲 Logistic回归模型**

**一、基本要求**

要求学生了解Logistic回归模型的基本思想和模型构建、分析方法，能够用这些方法分析解决实际问题。

**二、授课方法**

自学

**三、学习内容**

（一）Logistic回归模型的基本思想

（二）Logit转换

（三）Logistic模型的构建、分析及应用

**四、重点难点**

Logistic回归模型的基本思想及模型的构建、分析。

**五、思考与讨论**

1﹑简述Logistic回归模型构建的基本思路。

2﹑二元Logistic回归模型和多元Logistic回归模型有何异同？

3﹑在SPSS软件中如何实现Logistic回归模型？

**第六讲 聚类分析**

**一、基本要求**

要求学生了解聚类分析的原理和作用，了解聚类分析中常用的距离和相似系数以及它们各自的特点及适用范围。掌握系统聚类法，能够应用系统聚类法解决实际数据分析问题。

**二、授课方法**

自学

**三、学习内容**

（一）聚类的目的  
 介绍聚类分析的目的和聚类的分类

（二）距离和相似系数  
 介绍聚类常用的距离和相似系数

（三）类和类的特征  
 介绍类的四种定义、类的三个常用特征和计算类与类之间距离的五种方法  
 （四）系统聚类法  
 介绍系统聚类法的步骤、性质以及在不同类与类之间距离下距离递推公式  
 （五）有序样品的聚类  
 介绍有序样品的聚类方法和步骤

**四、重点难点**

重点是系统聚类法。

**五、思考与讨论**

1﹑简述系统聚类法的基本思路。

2﹑写出样品间相关系数公式。

3﹑常用的距离及相似系数有哪些 ？它们各有什么特点？

4﹑利用谱系图分类应注意哪些问题？

5﹑在SPSS中如何实现系统聚类分析？

**第七讲 多元回归分析**

**一、基本要求**

要求学生了解多元回归分析的目的和基本思想，掌握多重多元回归分析的建模方法。

**二、授课方法**

自学

**三、学习内容**

（一）多元回归分析  
（二）多重多元回归分析

**四、重点难点**

重点是多重多元回归分析的建模思想和实现方法。

**五、思考与讨论**

1﹑简述多元回归分析的基本思路。

2﹑简述多重多元回归分析的建模思想。

3﹑在SPSS中如何实现多元回归分析和多重多元回归分析？