河北经贸大学课程水平认定

《统计建模与应用》课程大纲

|  |  |  |  |
| --- | --- | --- | --- |
| **课程名称** | 统计建模与应用 | **课程类型** | 专业必修课程 |
| **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 |

1. **课程性质**

统计建模与应用是一门集理论、方法与实践与一体的统计应用学科。旨在培养学生利用所学统计知识，基于实际问题的大量统计数据，结合专业的统计分析软件，以建立统计模型作为分析工具和手段，完成分析实际问题和解决实际问题的过程。通过学习该门课程，使同学们获得从实际问题入手，到利用软件进行求解，以及对计算结果进行分析的全面训练。

**二、学习目的**

本课程主要采用多媒体学习方式，在讲授理论与方法的同时，更加注重理论与实际相结合，培养学生运用所学知识对现实经济问题进行分析和处理的能力。

另外，学生在学习时，要侧重于对课程理论与方法原理的理解，并学会运用相关软件进行数据处理，包括模型的建立、参数估计、模型精度评价以及模型的应用，通过对结果的分析以达到探索社会经济现象的动态结构和发展变动规律的目的。

**三、学习要求**

根据学习内容、学生特点及学时安排，采取自学、调研等相结合方式组织学习，要求学生阅读大量的相关资料，完成相关的自学、调研等学习环节，培养学生的独立研读能力。

**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 概率统计的基本知识** | **2** |  |
| **第二讲 统计数据的预处理** | **2** | **2** |
| **第三讲 统计数据的基本统计分析** | **2** | **2** |
| **第四讲 相关分析及回归分析方法建模及应用** | **4** | **2** |
| **第五讲 多元统计分析方法建模与应用** | **8** | **2** |
| **第六讲 时间序列分析方法建模与应用** | **4** | **2** |

**五、课程考核及成绩评定**

课程考核为上机开卷考试。

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** |
| 《统计建模与R软件》 | 薛毅 | 清华大学出版社，2007 |
| **学习参考书** | **作者** | **出版社及出版时间** |
| SPSS统计分析方法及应用(第三版) | 薛薇 | 电子工业出版社，2014 |
| 时间序列分析 | ［美］詹姆斯 D. 汉密尔顿 | 中国社会科学出版社，1999 |

**七、学习具体内容和要求**

**第一讲 概率统计的基本知识**

**一、基本要求**

复习并掌握数理统计的基本知识。

**二、授课方法**

自学与答疑相结合。

1. **学习内容**
2. 随机事件与概率

1、随机事件、概率、古典概率、几何概率、条件概率

2、概率的乘法公式、全概率公式、Bayes公式

3、独立事件

1. 随机变量及其布

1、随机变量的定义、分布函数

2、离散随机变量、连续随机变量

3、随机向量

（三）随机变量的数字特征

1、数学期望、方差

2、几种常用的随机变量分布的期望与方差

3、协方差与相关系数、矩与协方差矩阵

（四）极限定理

1、大数定律

2.中心极限定理

（五）数理统计的基本概念

1、总体、个体、简单随机样本

2.参数空间及分布族，统计量与抽样分布

3、正态总体样本均值与样本方差的分布

**四、重点难点**

随机变量的分布、极限定理、数理统计的基本概念。

**五、思考与讨论**

1、随机变量的定义、随机变量的分布函数？

2、常用的随机变量分布的期望与方差？

3、概率的乘法公式、全概率公式、Bayes公式？

**第二讲 统计数据的预处理**

**一、基本要求**

掌握统计数据处理的意义及数据预处理的类型及方法。

**二、授课方法**

理论与实际问题相结合

**三、学习内容**

（一）统计数据管理

1、数据的排序、数据选取、数据分组

2、变量的计算、分类汇总、计数

3、数据的转置、数据加权、数据拆分

（二）数据清理

1、缺失值数据的处理

2、异常值数据的处理

（三）数据转换

1、数据标准化

2、数据的离散化

**四、重点难点**

数据的管理及数据的标准化方法，数据缺失

**五、思考与讨论**

1、统计数据的管理有哪些方法？

2、如何处理数据的缺失值问题和异常值问题？

3、常用的对数据标准化方法？

**第三讲 统计数据的基本统计分析**

**一、基本要求**

掌握数据的基本统计分析，把握数据的基本统计特征及总体的分布形态，对后续的建模将起到重要的支撑作用。

**二、授课方法**

理论与实际问题相结合

**三、学习内容**

（一）频数分析

1、频数分析的目的和基本任务

2、频数分析的软件实现

3、交叉分组下的频数分析的目的及任务

4、交叉列联表行列变量间关系的分析

5、案例分析

（二）计算基本描述统计量及软件实现

1. 多选项分析
   1. 多选项分析的目的
   2. 多选项分析的上机实现
   3. 多选项分析的应用案例

（四）数据的分布

1、分布函数

2、直方图、经验分布图与QQ图

3、正态性检验及分布拟合检验

**四、重点难点**

交叉分组下的频数分析目的及实现过程，多选项分析的目的及实现过程。

**五、思考与讨论**

1、简单频数分析和交叉分组下的频数分析的目的及任务，如何上机实现？

2、基本描述统计量的种类及软件实现？

3、什么是多选项分析？如何上机实现？

4、数据序列的正态性检验及分布拟合检验方法及实现？

**第四讲 相关分析及回归分析方法建模及应用**

**一、基本要求**

相关分析和回归分析都是分析客观事物之间相关性的数量分析方法。掌握相关分析及回归分析的概念、区别与联系，两种方法的适用范围，掌握相关分析及回归分析的相关理论与方法，能够熟练使用专业统计分析软件进行相关的操作，并对结果做出正确的分析。

**二、授课方法**

理论与实际问题相结合

**三、学习内容**

1. 相关分析、偏相关分析

1、散点图、相关系数及应用案例

2、偏相关分析的意义、偏相关系数及上机实现

（二）典型相关分析

1. 典型相关分析的目的、任务和基本理论
2. 典型相关分析的数学描述

3、典型相关分析的上机实现及案例

（三）回归分析

1、回归分析概述、线性回归模型

2、回归参数的最小二乘估计、统计检验、计量经济检验及预测

3、线性回归分析的上机操作及案例

（四）多重多元回归分析

1、多重多元回归分析的目的、任务

2、多重多元回归的数学模型

3、多重多元回归分析的基本思想

4、多重多元回归分析的的上机实现及案例

**四、重点难点**

相关与回归分析的基本理论、方法，上机实现及结果分析。

**五、思考与讨论**

1、相关分析、偏相关分析的目的、任务，联系与区别，适用范围？

2、线性回归模型的统计检验、计量经济检验理论、方法及上机实现？

3、典型相关分析的目的、任务和基本理论？

4、多重多元回归分析的目的、任务和基本理论？

**第五讲多元统计分析方法建模与应用**

**一、基本要求**

多元分析是多变量的统计分析方法，是数理统计中应用广泛的一个重要分支。通过本讲的学习，学生需要掌握多元统计分析的多种应用方法，主要包括因子分析、对应分析、聚类分析、判别分析等重要内容。

**二、授课方法**

理论与实际问题相结合

**三、学习内容**

1. 因子分析

1、因子分析的目的、任务

2、因子分析的基本思想、数学模型

3、因子分析的实现步骤、上机实现及案例分析

1. 聚类分析

1、聚类分析的目的、任务

2、聚类分析的基本思想、数学模型

3、聚类分析的实现步骤、上机实现及案例分析

（三）判别分析

1、判别分析的目的、任务

2、判别分析的基本思想、数学模型

3、判别分析的实现步骤、上机实现及案例分析

（四）对应分析

1、对应分析的目的、任务

2、对应分析的基本思想、数学模型

3、对应分析的实现步骤、上机实现及案例分析

**四、重点难点**

因子分析、对应分析、聚类分析、判别分析等多元分析方法的目的、任务、基本理论、数学模型及上机实现过程、结果分析。

**五、思考与讨论**

1、因子分析的目的、任务、基本理论、数学模型及上机实现过程、结果分析？

2**、**对应分析的目的、任务、基本理论、数学模型及上机实现过程、结果分析？

3、聚类分析的目的、任务、基本理论、数学模型及上机实现过程、结果分析？

4、判别分析的目的、任务、基本理论、数学模型及上机实现过程、结果分析？

**第六讲 时间序列分析方法建模与应用**

**一、基本要求**

掌握时间序列分析方法的建模过程、步骤、上机实现及结果解释。

**二、授课方法**

理论与实际问题相结合

**三、学习内容**

（一）时间序列的平稳性检验

1、数据图检验法

2、自相关、偏自相关检验法

3、单位根检验法

（二）时间序列的预处理方法

1、差分、季节差分、对数变换与差方运算的结合运用

2、季节调整法

3、序列缺失值处理的基本方法

（三）齐次非平稳序列模型分析

1、ARIMA分析的基本思想和模型

2、ARIMA分析的基本操作及应用分析

（四）指数平滑

1、指数平滑的基本思想、指数平滑模型

2、指数平滑的基本操作及应用案例分析

（五）协整与误差修正模型

1、协整的概念

2、协整的检验

3、误差修正模型

**四、重点难点**

非平稳性时间序列的检验及非平稳时间序列建模的方法、过程、上机实现及结果分析；协整与误差修正模型的相关概念、理论方法与应用。

**五、思考与讨论**

1、如何对时间序列进行平稳性检验？如何上机实现？

2、时间序列建模之前进行预处理的方法有哪些？

3、ARIMA模型的建立过程、上机实现及结果的分析？

4、指数平滑的基本思想、模型、上机实现及结果的分析？

5、协整与误差修正模型的相关概念、理论方法及应用？