《经济数据建模与分析（英语）》教学大纲

课程编号：151403B

课程类型：□通识教育必修课 □通识教育选修课

□专业必修课 ☑专业选修课

□学科基础课

总学时：48 讲课学时：48 实验（上机）学时：0

学　　分：3

适用对象：金融学（数据与计量分析）

先修课程：概率论与数理统计、中级微观经济学、计量经济学

一、教学目标

《经济数据建模与分析》是针对金融学专业高年级本科生开设的专业选修课。本课程的教学目的是培养学生对经济数据进行分析的基本能力。本课程以计算机统计分析软件为工具，利用各种统计分析方法对批量经济数据建立统计模型并进行估计和分析。本课程注重统计经济数据的分析与统计软件的实际操作相结合。通过本课程的学习，学生会运用常用的统计软件进行数据的建模与分析。本课程有助于培养学生利用数据和统计方法解决实际问题的能力，特别是提高从复杂问题中提炼统计模型的能力。这门课程能够帮助培养学生的核心价值观和职业道德素养，具体的教学目标如下：

目标1：了解和应用国内常用的经济数据库

目标2：能够熟练运用STATA/R语言进行实际数据的模型估计与分析

目标3：能够运用数据模型分析解决实际的社会经济学问题

**二、**教学内容及其与毕业要求的对应关系

本课程重点讲授统计分析的基本方法和模型以及其在计算统计软件（STATA 和R语言）上的实现，并结合国内的经济学数据库，对相关中国社会经济学问题进行实证分析。其中，应该重点讲授方法和模型以及计算经济软件上的实际操作，采取课堂授课和计算机软件的实际操作相结合。线性模型和广义线性模型有非常广泛的应用，会进行较为细致的讲解。STATA和R语言是经济学领域常用的统计软件，对于第一次使用这些软件的学生是一个难点。根据学生的实际情况，进行问题引领式教学，并加强上机操作练习。

在教学过程中，采取统计理论讲解和软件操作相结合的教学方法。介绍学生了解国内相关的经济学数据库，并进行实际的数据分析练习。每节课的理论模型都会有相应的软件操作练习，并在学期期中对学生进行软件编程的考核。在使用实际数据的过程中，会对数据的清理进行详细的讲解。该课程从培养学生分析数据的实际动手能力等方面促进了毕业要求的实现。通过对具体的中国经济问题的研究并结合时事热点的分析，本课程有利于增强学生的责任感和使命感，培养学生的家国情怀。

**三、**各教学环节学时分配

以表格方式表现各章节的学时分配，表格如下：（宋体，小四号字）

**教学课时分配**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 章节内容 | 讲课 | 实验 | 其他 | 合计 |
| 1 | 第一章统计模型概述 | 2 |  |  | 2 |
| 2 | 第二章经济数据库介绍 | 2 |  |  | 2 |
| 3 | 第三章STATA基础应用 | 6 |  |  | 6 |
| 4 | 第四章统计描述 | 4 |  |  | 4 |
| 5 | 第五章R语言基本操作 | 6 |  |  | 6 |
| 6 | 第六章估计与检验 | 4 |  |  | 4 |
| 7 | 第七章线性模型 | 7 |  |  | 7 |
| 8 | 第八章广义线性模型 | 7 |  |  | 7 |
| 9 | 第九章工具变量和2SLS | 7 |  |  | 7 |
| 10 | 第十章统计模型的实证分析 | 3 |  |  | 3 |
| 合计 |  | 48 |  |  | 48 |

**四、教学内容**

第一章统计模型概述

第一节 统计模型的特点

第二节 统计模型的作用

第三节 统计模型的类型

第四节 实现统计模型的软件平台

教学重点、难点: 统计模型区别于其他模型的特征

课程的考核要求：了解统计模型的作用，理解统计模型的特征

复习思考题：

1. 什么是模型

2． 统计模型的特点是什么

第二章经济数据库介绍

第一节 调研数据及抽样设计

第二节 国家改革开放40多年的伟大成就

第三节 微观及宏观经济数据库的介绍

第四节 国外数据的介绍

教学重点、难点: 调研数据的特点

课程的考核要求：深刻理解体会国家改革开放取得的伟大成就，增强学生的爱国情怀和民族自豪感；结合我国经济现状，了解常用的经济学数据库

复习思考题：

面板数据和横截面数据的区别是什么？

第三章STATA基础应用

第一节数据的读取

第二节数据的探究与描述

第三节编程初步

教学重点、难点：运用STATA对数据进行基本的描述性分析

课程的考核要求：掌握STATA描述性分析的基本命令，运用命令分析数据

复习思考题：

描述数值类变量的命令有哪些？

第四章 统计描述

第一节 集中趋势的描述

数值平均数

位置平均数

第二节 离散程度的描述

离散程度的绝对指标

离散程度的相对指标

数据的标准化

是非标志标准差

第三节 分布的偏态与峰度

原点矩与中心矩

分布的偏态

分布的峰度

教学重点、难点：离散程度的描述

课程的考核要求:理解统计描述的相关概念

复习思考题：

1. 简述算术平均数、众数与中位数的特点及应用场合。

2. 简述全距、平均差、标准差与方差的应用场合。

第五章 R语言基本操作

第一节 R语言的数据结构

向量、索引向量、因子向量

数组和矩阵

数据框

列表

第二节 R语言的程序结构

顺序结构

选择结构

循环结构

第三节 R语言的绘图方法

高级绘图语句

低级绘图语句

交互式绘图语句

绘图语句的参数

教学重点、难点：R语言的数据结构和程序结构

课程的考核要求：理解R语言的程序结构并进行初步的编程

复习思考题：

1． 何为因子向量？

2． 数据框与列表的区别

3. 循环语句的种类有哪些

4. 高级绘图语句和低级绘图语句的关系是什么？

第六章 估计与检验

第一节 参数估计

点估计

区间估计

第二节 假设检验

Z检验

T检验

F检验

卡方检验

似然比检验

第三节 在STATA/R语言中进行估计和检验

参数估计的实现方法

假设检验的实现方法

教学重点、难点：参数估计和假设检验的原理

课程的考核要求：理解参数估计和假设检验的原理，掌握其在STATA/R中的实现方法。

复习思考题：

1． 似然比检验的思想是什么

2．如何实现卡方检验

第七章 线性模型

第一节 方差分析模型

方差分析的基本思想

方差分析的基本假设

第二节 线性回归模型

一元线性回归

多元线性回归

第三节 案例分析

教学重点、难点：多元线性回归模型的基本假设及估计结果的诠释

课程的考核要求：理解多元线性回归模型的假设，并能用统计软件实现线性回归模型分析

复习思考题：回归估计的方法

第八章 广义线性模型

第一节 logistic回归模型

logistic回归模型的设定

logistic回归模型的估计和检验

logistic回归模型的R语言实现方法

第二节 定序回归模型

定序回归模型的设定

定序回归模型的估计和检验

定序回归模型的R语言实现方法

第三节 泊松回归模型

泊松回归模型的设定

泊松回归模型的估计和检验

泊松回归模型的R语言实现方法

教学重点、难点：GLM函数的使用

课程的考核要求：掌握logistic回归模型、定序回归模型和泊松回归模

的基本原理和统计软件中模型的实现方法

复习思考题：

1． Logistic回归模型的使用范围

2． GLM函数与LM函数的区别

3． 如何实现模型的比较分析

第九章工具变量和2SLS

第一节工具变量模型和假设

第二节二阶最小二乘估计

第三节工具变量的选取

第四节工具变量模型的应用

教学重点、难点：内生性问题及工具变量的选取

课程的考核要求：理解实际模型中内生性问题的来源，掌握工具变量的检验方法。

复习思考题：

如何选取合适的工具变量？

第十章统计模型的实证分析

1. 数据清理
2. 统计模型的实证分析

教学重点、难点：运用实际数据进行建模与分析

课程的考核要求：应用STATA/R 进行模型估计与分析，并能解释模型背后的社会经济学现象

复习思考题：

模型估计过程中有什么问题？如何解决？

五、考核方式、成绩评定

本课程的考核方式为项目设计与论文。 学期期中的项目展示占30%，平时作业占30%，期末学术论文占40%。学期期中的项目展示要求学生以团队形式完成一个课题项目的数据分析并进行课堂汇报。期末学术论文要求主题明确，运用现有的数据和统计方法，对感兴趣的社会经济学问题进行实证分析，并完成一篇完整的学术论文。

六、主要参考书及其他内容

[1] 薛毅。统计建模与R软件. 清华大学出版社，2007年

[2] 吴喜之。统计学：从数据到结论（第4版）。中国统计出版社，2013年

[3] 伍得里奇。计量经济学导论：现代观点. 中国人民大学出版社，2008年

[4] 斯托克和沃森。计量经济学导论。中国人民大学出版社，2014年

[5] Cameron and Trivedi, Microeconometrics using Stata: Revised Edition. Stata

Press,2010.

执笔人：王超 教研室主任：　　　　　系教学主任审核签名：