《贝叶斯统计》教学大纲

“Bayesian Statistics” Course Outline

课程编号： 2021012B

课程类型：专业选修课

总学时： 32 讲课学时：32 实验（上机）学时：0

学　　分： 2

适用对象：金融学（数据与计量分析）

先修课程：数学分析、概率论与数理统计、计量经济学

**Course Code：**2021012B

**Course Type：** Discipline Elective

**Total Hours：** 32 **Lecture：**32

**Experiment(Computer)：**0

**Credit：** 2

**Applicable Major：** Finance

**Prerequisite：** Mathematical Analysis, Probability Theory and Statistics, Econometrics

**一、课程的教学目标**

本课程旨在向学生介绍贝叶斯统计理论、贝叶斯统计方法及其在实证研究中的应用。贝叶斯统计理论与传统统计理论遵循着不同的基本假设，为我们处理数据信息提供新的角度和解读思路，并在处理某些复杂模型上（如，估计动态随机一般均衡模型、带时变参数的状态空间模型等）相比传统方法具有相对优势。

本课程要求学生在选课前具备基本的微积分、概率统计以及计量经济学知识。以此为起点，我们将主要就贝叶斯统计理论知识、统计模型的应用以及基于计算机编程的实证能力三方面对学生进行训练。经过对本课程的学习，学生应了解贝叶斯框架的基本思想，掌握基本的贝叶斯理论方法及其主要应用，并掌握实证研究中常用的贝叶斯数值抽样方法以及相关的计算机编程技能。尤其是，学生应能明确了解贝叶斯统计方法与传统统计方法在思想和应用上的区别以及各自的优缺点，以便能在实际应用中合理选择统计分析工具。通过课程内容的学习以及实际应用，学生应建立逐步建立严谨的辩证思维框架与学术规范，形成科学的世界观与价值观。

This course introduces the basic concepts of Bayesian statistics and the use of Bayesian econometric methods in empirical study. Bayesian statistics has different fundamental assumptions from the classical (frequentist) framework, providing us with an alternative way in analyzing and interpreting data information. Bayesian methods also have relative advantages, and thus are widely used, in dealing with certain complicated models (for example, the estimation of Dynamic Stochastic General Equilibrium model, state space models with time-varying parameters, etc.).

Students should have had basic trainings on calculus, probability theory and statistics, and preferably econometrics prior to this course. The major trainings offered in this course focus on Bayesian theories, Bayesian statistical models with applications and computational skills required for empirical analysis. After the course, students should develop their understanding on the philosophy of Bayesian framework, understand basic Bayesian theories, Bayesian estimation methods and their applications, and master the computer skills for the practical use of Bayesian methods. Specifically, students should understand the differences between the Bayesian viewpoint and the classical frequentist perspective in order to be able to choose appropriate analyzing tools in empirical use. Throughout the course study, students should establish rigorous framework of critical thinking and academic norms, and form scientific world conceptions and value systems.

**二、教学基本要求**

贝叶斯统计学和计量方法在近年得到越来越广泛的关注和应用，主要得益于计算机技术的发展使得贝叶斯数值抽样方法在实际应用中得以实现。因此，除了对贝叶斯相关理论的讲授，计算机数值方法的介绍与相关实践也是同等重要的。

统计理论的部分，本课程主要涵盖贝叶斯定理以及贝叶斯方法的基本数据分析框架，以及运用贝叶斯方法进行回归模型的估计、预测和模型比较的基本方法。数值方法部分，主要介绍Markov Chain Monte Carlo后验分布抽样方法（Gibbs抽样法和Methopolis-Hastings抽样法），以及边缘似然函数和贝叶斯因子的估计方法等。这两部分的内容应紧密结合并辅以实例分析，以课堂讲授结合计算机演示的方式进行教学。另外，贝叶斯方法与传统统计方法的对比也应贯穿课程各章节，以帮助学生了解两种方法的优缺点及适用范围，为实际应用提供指导，并帮助学生形成严谨的治学态度与规范。

对应课堂内容，课后作业也应强调理论和实践的结合，编程练习应占重要比例。教师应提供必要的编程软件使用指导，并指导学生根据提供的范例练习运用贝叶斯方法进行数据分析。

课程考核方式及其权重如下：

|  |  |
| --- | --- |
| 出勤 | 10% |
| 作业 | 20% |
| 期末闭卷考试 | 30% |
| 课程论文 | 30% |
| 课程论文宣讲 | 10% |

Thanks to the development of computer technology, Bayesian statistics and econometrics has become more popular as Bayesian computational methods have become practical. Computational methods are as important as theories to Bayesian statistics, so as to this course.

In the theoretical statistics part, this course involves lectures on Bayes Theorem, the general Bayesian analysis framework, and the general approaches to estimate models, develop predictions and compare models in the Bayesian way. In terms of computational methods, we will mainly introduce the MCMC posterior sampling (the Gibbs Sampling and the Metropolis-Hasting Algorithm), and the approximation technique for marginal likelihood and the Bayes factor. The two parts should introduce along with each other. Empirical and computer illustrations are also necessary for students to understand the approach. Last but not the least, instructor should have the Bayesian framework compared with the Classical frequentist framework in all aspects when appropriate in order to illustrate the advantages/disadvantages of the two frameworks, providing students with empirical instructions and help on establishments of academic norms.

Homework assignments should be composed of derivations and computer exercises. Proper software instructions and examples should be provided to facilitate students’ practice on Bayesian analysis techniques.

The grading weights are as follows:

|  |  |
| --- | --- |
| Attendancy | 10% |
| Homework Assignments | 20% |
| Final Exam (Closed Book) | 30% |
| Term Paper | 30% |
| Term Paper Presentations | 10% |

**三、各教学环节学时分配**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 章节内容 | 讲课 | 实验 | 其他 | 合计 |
| 1 | 课程简介及贝叶斯定理  Course Overview and Bayes Theorem | 2 | 0 | 0 | 2 |
| 2 | 贝叶斯统计理论初步  Elements of Bayesian Inference | 4 | 0 | 0 | 4 |
| 3 | 线性回归模型（假设共轭先验分布）  Linear Regression Models with Natural Conjugate Priors | 6 | 0 | 0 | 6 |
| 4 | 线性回归模型（假设其他先验分布）  Linear Regression Models with Other Priors | 9 | 0 | 0 | 9 |
| 5 | 非线性回归模型  Nonlinear Regression | 6 | 0 | 0 | 6 |
| 6 | 潜在变量模型  Models with Latent Variables | 6 | 0 | 0 | 6 |
| 7 | 时间序列模型  Time Series Models | 6 | 0 | 0 | 6 |
| 8 | 贝叶斯模型比较与模型平均  Bayesian Model Comparisons and Model Averaging | 6 | 0 | 0 | 6 |
| 9 | 论文宣讲  Term Paper Presentations | 3 | 0 | 0 | 3 |
| 合计 |  | 48 | 0 | 0 | 48 |

**四、教学内容**

**第一章课程简介及贝叶斯定理**

第一节课程简介

第二节贝叶斯定理

1. 贝叶斯定理
2. 贝叶斯定理的简单应用

教学重点、难点：贝叶斯定理的涵义并通过实例阐释，阐释应体现贝叶斯框架所隐含的先验讯息与客观样本证据科学结合的思辨方式。

课程考核要求：掌握贝叶斯定理的涵义，并能运用定理对对简单的事件作出推论。

**Chapter 1 Course Overview and Bayes Theorem**

Section 1 Course Overview

Section 2 Bayes Theorem

1. Bayes Theorem
2. Simple Applications of Bayes Theorem

Key and Difficult Points: the key idea of Bayes Theorem presented with examples. The illustration should present the way of critical combination of prior information and objective sample evidence implied by the Bayesian framework.

Evaluation Requirements: master the key idea of Bayes Theorem, and make inferences with the theorem in simple scenarios.

**第二章贝叶斯统计理论初步**

第一节贝叶斯统计理论的基本要素

1. 先验分布与后验分布
2. 贝叶斯定理的应用
3. 贝叶斯统计与传统统计方法的系统性差别

第二节点估计和置信区间

1. 点估计
2. 最高后验密度区间

第三节贝叶斯决策理论

第四节模型比较

1. 边际似然函数
2. 预测密度函数

教学重点、难点：贝叶斯统计各要素的相互关联，与传统统计方法的根本区别。

课程考核要求：掌握贝叶斯统计体系各要素的定义及相互联系，了解贝叶斯统计体系与传统统计学的根本区别，了解贝叶斯统计推断的常见内容。

**Chapter 2 Elements of Bayesian Inference**

Section 1 Basic Elements of Bayesian Statistics

1. Prior and Posterior
2. How Bayes Theorem is used
3. Systematic Differences between the Bayesian and Frequentist View

Section 2 Point Estimation and Credible Sets

1. Point Estimation
2. Highest Posterior Density Interval

Section 3 Bayesian Decision Theory

Section 4 Model Comparison

1. Marginal Likelihoods
2. Predictive Densities

Key and Difficult Points: the relationships between elements of Bayesian Statistics, the fundamental differences between the Bayesian view and the frequentist view.

Evaluation Requirements: master the concepts of the elements of Bayesian Statistics and the relationship among them, understand the fundamental differences between the Bayesian and frequentist view, and understand the common components of Bayesian inferences.

**第三章线性回归模型（假设共轭分布）**

第一节后验分布的推导

1. 用矩阵表示的线性回归模型
2. 似然函数
3. 共轭先验分布
4. 后验分布

第二节预测

第三节贝叶斯数值计算方法：蒙特卡罗模拟积分

第三节实证范例

教学重点、难点：后验分布的推导与应用，蒙特卡罗模拟积分，建立规范的科学模拟方法论。

课程考核要求：理解共轭先验分布假设下的后验分布的推导，掌握运用后验分布进行实证分析，掌握蒙特卡罗模拟积分方法。

**Chapter 3 Linear Regression Models with Natural Conjugate Priors**

Section 1 Derivation of the Posterior

1. The Linear Regression Model in Matrix Notation
2. The Likelihood Function
3. The Natural Conjugate Prior
4. The Posterior

Section 2 Prediction

Section 3 Bayesian Computational Methods: Monte Carlo Integration

Section 4 Empirical Illustrations

Key and Difficult Points: the derivation of the posterior and its applications, Monte Carlo Integration, introducing standard scientific simulation methodology.

Evaluation Requirements: understand the derivation of the posterior, analyze real data based on the posterior, master the methodology of Monte Carlo integration.

**第四章线性回归模型（假设其他先验分布）**

第一节独立正态－伽玛先验分布

1. 独立正态－伽玛先验分布
2. 后验分布

第二节贝叶斯数值计算方法：后验分布Gibbs随机抽样法

1. Gibbs随机抽样法
2. 马尔可夫链蒙特卡罗收敛判定

第三节带限制的线性回归模型

1. 后验分布
2. 贝叶斯数值计算方法：重要性随机抽样法

第四节实证范例

教学重点、难点：Gibbs随机抽样法，重要性随机抽样法。教授马尔可夫链蒙特卡罗收敛判定时应强调其对正确科学应用贝叶斯方法的重要性。

课程考核要求：掌握Gibbs随机抽样法和重要性随机抽样法，了解如何根据实际情况对抽样方法做出选择。

**Chapter 4 Linear Regression Models with Other Priors**

Section 1 The Independent Normal-Gamma Prior

1. The Independent Normal-Gamma Prior
2. The Posterior

Section 2 Bayesian Computational Methods: Gibbs Sampling

1. Gibbs Sampling
2. MCMC Convergence Diagnostics

Section 3 Linear Regression Models with Restrictions

1. The Posterior
2. Bayesian Computational Methods: Importance Sampling

Section 4 Empirical Illustrations

Key and Difficult Points: Gibbs sampling, importance sampling. When introducing the MCMC convergence diagnostics, instructor should emphasize its importance on correct application of Bayesian methods.

Evaluation Requirements: master the methodology of Gibbs sampling and importance sampling, understand how to choose appropriate methodology given different scenarios.

**第五章非线性回归模型**

第一节非线性回归模型

1. 模型设定
2. 先验分布与后验分布

第二节贝叶斯数值计算方法：Metropolis-Hasting抽样法

1. 一般性步骤
2. 独立链MH抽样法
3. 随机游走链MH抽样法
4. Metropolis-within-Gibbs抽样法

第三节实证范例

教学重点、难点：MH抽样法，Metropolis-within-Gibbs抽样法。

课程考核要求：理解MH抽样法的一般步骤，掌握独立链与随机游走链MH抽样法，了解Metropolis-within-Gibbs抽样法。

**Chapter 5 Nonlinear Regression**

Section 1 Nonlinear Regression

1. Model Setup
2. The Prior and The Posterior

Section 2 Bayesian Computational Methods: Metropolis-Hasting Algorithm

1. The General Algorithm
2. The Independent Chain MH Algorithm
3. The Random Walk Chain MH Algorithm
4. Metropolis-within-Gibbs

Section 3 Empirical Illustrations

Key and Difficult Points: Metropolis-Hasting Algorithm, Metropolis-within-Gibbs.

Evaluation Requirements: understand the general algorithm of Metropolis-Hasting, master the independence chain/random walk chain MH algorithm, understand the methodology of Metropolis-within-Gibbs.

**第六章潜在变量模型**

第一节删截线性回归模型

第二节Probit模型

第三节Tobit模型

第四节带混合正态的模型

教学重点、难点：各模型的特点及相对应的数据特点，各模型适用的后验分布抽样方法。

课程考核要求：理解各模型的特点及相对应的数据特点，了解适用的后验分布抽样方法。

**Chapter 6 Models with Latent Variables**

Section 1 Censored Linear Models

Section 2 Probit Models

Section 3 Tobit Models

Section 4 Modeling with Mixtures of Normals

Key and Difficult Points: features of the above models and their corresponding datasets, posterior sampling with the above models.

Evaluation Requirements: understand the features of the above models and their corresponding datasets, understand the methodology for posterior sampling.

**第七章时间序列模型**

第一节线性时间序列模型

1. 常见模型
2. 一般性估计方法

第二节状态空间模型

1. 一般模型设定
2. Gibbs抽样法
3. MH抽样法

教学重点、难点：状态空间模型的贝叶斯估计法。

课程考核要求：掌握一般时间序列模型的贝叶斯估计方法，了解状态空间模型及其贝叶斯估计法。

**Chapter 7 Time Series Models**

Section 1 Linear Time Series Models

1. Common Time Series Models
2. General Bayesian Approach

Section 2 State-Space Models

1. General Model Setup
2. Gibbs Sampling
3. Metropolis-Hasting Algorithm

Key and Difficult Points: Bayesian estimation algorithm of state-space models.

Evaluation Requirements: master the general estimation approaches for linear time series models, understand the setup of state-space models and their Bayesian estimation algorithm.

**第八章贝叶斯模型比较与模型平均**

第一节贝叶斯模型比较

1. 贝叶斯因子
2. 用Gelfand－Dey方法估计边际似然函数
3. 用Chib方法估计边际似然函数
4. （选讲）贝叶斯假设检验

第二节贝叶斯模型平均

教学重点、难点：贝叶斯因子的应用，边际似然函数的估计。教授过程应以实证例子展示如何科学进行模型选择，以及模型误用带来的学术和实证风险。

课程考核要求：掌握贝叶斯因子的定义及应用，了解边际似然函数的估计方法, 了解模型平均的意义。

**Chapter 8 Bayesian Model Comparisons and Model Averaging**

Section 1 Bayesian Model Comparisons

1. Bayes Factor
2. Marginal Likelihood Approximation with the Gelfand-Dey Method
3. Marginal Likelihood Approximation with the Chib Method

Section 2 Bayesian Model Averaging

Key and Difficult Points: Bayes factor and its application, approximation of marginal likelihood. Empirical examples should be provided to illustrate of the scientific model selection process and the risk of model mispesification.

Evaluation Requirements: master the concept and application of Bayes factor, understand the methodology to approximate the marginal likelihood, understand the idea of Bayesian model averaging.

**第九章论文宣讲**

以PPT汇报分组课程论文成果。论文内容应具备科学性，论文撰写应严格符合相关学术规范。

课程考核要求：英文宣讲及PPT，要求宣讲结构清晰、有条理，内容细节见课堂要求。

**Chapter 9 Term Paper Presentations**

A short presentation on term papers in groups. The content of term paper should include scientific aspects and rigorously follow academic norms.

Evaluation Requirements: presentation in English and well organized, details as announced in class.

**五、其它**

授课老师应根据学生的能力和接受程度适当调节课程进度和内容，但第一至五章及第八章第一节建议优先讲授。

Instructor should adjust the course schedule and content according to the capability of students. But Chapter 1 to 5 and the Section 1 in Chapter 8 should be covered in priority.

**六、主要参考书**

[1] Gary Koop. Bayesian Econometrics. England: Wiley. 2003.

[2] John Geweke. Contemporary Bayesian Econometrics and Statistics. NJ: Wiley-Interscience. 2005