

课程大纲

贝叶斯数据分析

课程编号：02800150

学 分：2

课程类型：选修

先修课程：数学分析/高等数学、线性代数、概率论与数理统计

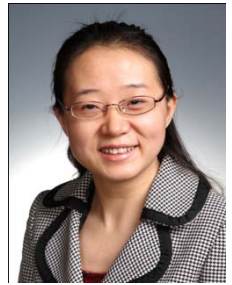
授课对象：学术研究生

任课教师：张俊妮

开课学期：2015 年春

任课教师简历（500 字左右）：

张俊妮为北京大学光华管理学院统计学副教授。1998 年毕业于中国科学技术大学，获计算机软件学士学位；2002 年毕业于美国哈佛大学，获统计学博士学位。研究领域为因果推断、贝叶斯分析、小区估计、数据挖掘以及文本挖掘。



任课教师联系方式：

光华管理学院 2 号楼 473 办公室，电话 010-62757922，邮箱 zjn@gsm.pku.edu.cn

助教姓名及联系方式：

董晶，jingdong1989@pku.edu.cn

一、项目培养目标

- 1 **Learning Goal 1** Graduates will be thoroughly familiar with the specialized knowledge and theories required for the completion of academic research.
 - 1.1 Objective 1 Graduates will have a deep understanding of basic knowledge and theories in their specialized area.
 - 1.2 Objective 2 Graduates will be familiar with the latest academic findings in their specialized area and will be knowledgeable about related areas.
 - 1.3 Objective 3 Graduates will be familiar with research methodologies in their specialized area, and will be able to apply them effectively.
- 2 **Learning Goal 2** Graduates will be creative scholars, who are able to write and publish high-quality graduation dissertation and research papers.
 - 2.1 Objective 1 Graduates will write and publish high-quality graduation dissertation and research papers
 - 2.2 Objective 2 Graduates will be critical thinkers and innovative problems solvers.
- 3 **Learning Goal 3** Graduates will have a broad vision of globalization and will be able to communicate and cooperate with international scholars
 - 3.1 Objective 1 Graduates will have excellent oral and written communication skills
 - 3.2 Objective 2 Graduates will be able to conduct efficient academic communication in at least one foreign language
- 4 **Learning Goal 4** Graduates will be aware of academic ethics and will have a sense of social responsibility.
 - 4.1 Objective 1 Graduates will have a sense of social responsibility.

- 4.2 Objective 2 Graduates will be aware of potential ethical issues in their academic career.
4.3 Objective 3 Graduates will demonstrate concern for social issues.

二、课程概述

本课程讲述贝叶斯数据分析的基本理论和算法，及其在营销、经济等学科中的应用。课程中将使用统计软件 R。

三、课程目标

在学习完本课程之后，学生应理解贝叶斯数据分析的基本理论和算法，并能使用统计软件 R 实现对一些常见类型的实际数据的贝叶斯分析。

四、内容提要及学时分配

课号	主题
第一讲	<p>Contents:</p> <p>(1) Bayesian Analysis of Simple Binomial and Normal Models (2) Prior Distributions, Posterior Distributions and Posterior Predictive Distributions (3) Summarizing Posterior Inference</p> <p>Case Study:</p> <p>(1) Estimating the Probability of a Female Birth (2) Modeling Height of Five-year-old Boys</p> <p>Reference: Gelman et al. (2014), Chapters 2-3</p>
第二讲	<p>Contents:</p> <p>(1) Introduction to Bayesian Hierarchical Model and Exchangeability (2) Gibbs Sampling</p> <p>Case Study: Educational Testing Experiments in Eight Schools</p> <p>Reference: Gelman et al. (2014), Chapter 5, Section 11.1</p>
第三讲	<p>Contents:</p> <p>(1) Bayesian Updating (2) Model Checking (3) Evaluating, comparing, and expanding models</p> <p>Case Study: Educational Testing Experiments in Eight Schools</p> <p>Reference: Greenberg (2008), Chapter 3; Gelman et al. (2014), Chapter 6-7</p>
第四讲	<p>Contents: Bayesian Analysis of Normal Linear Model</p> <p>Case Study: Estimating Demand Elasticities</p> <p>Reference:</p> <p>(1) Gelman et al. (2014), Chapters 14; Rossi et al. (2005), Section 2.8; Greenberg, E. (2012), Chapter 4.1-4.4, Section 8.1; (2) H. Schultz (1933). "A Comparison of Elasticities of Demand Obtained by Different Methods," <i>Econometrica</i>, Vol.1, #3, pp. 274-308. (3) H. Schultz (1925). "Appendix 2," <i>Journal of Political Economy</i>, Vol.33, #6, pp.634-637.</p>

第五、六讲	<p>Contents:</p> <p>(1) Bayesian Analysis of Regression Models with Multivariate (Continuous or Categorical) Responses</p> <p>(2) Latent Variables</p> <p>(3) Metropolis and Metropolis-Hastings Algorithms</p> <p>(4) Assessing Convergence of the MCMC Outputs</p> <p>Case Study: Analyzing Survey Data on Brands of Scotch Consumed</p> <p>Reference:</p> <p>(1) Gelman et al. (2014), Chapters 11; Rossi et al. (2005), Section 2.8, Chapter 4; Greenberg, E. (2012), Chapter 9.1-9.2</p> <p>(2) Edwards, Y. and G. Allenby (2003), "Multivariate Analysis of Multiple Response Data," Journal of Marketing Research, 40, 321-334.</p>
第七、八讲	<p>Contents: Bayesian Analysis of Hierarchical Linear Models</p> <p>Case Study:</p> <p>(1) Accounting-level Modeling of Trade Promotion for Sliced Cheese</p> <p>(2) Analyzing Store-Level Panel Data on Orange Juice Sales</p> <p>Reference:</p> <p>(1) Gelman et al. (2014), Chapter 15; Rossi et al. (2005), Chapter 5; Greenberg, E. (2012), Chapter 9.3</p> <p>(2) Boatwright et al (1999), "Account-Level Modeling for Trade Promotion," Journal of the American Statistical Association, 94, 1063-1073.</p> <p>(3) Alan L. Montgomery (1997), "Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data," Marketing Science 16(4) 315-337.</p>
第九、十讲	<p>Contents: Bayesian Analysis of Hierarchical Generalize Linear Models</p> <p>Case Study:</p> <p>(1) Analyzing Credit Card Data from A Conjoint Experiment</p> <p>(2) Analyzing Household Panel Data on Margarine Purchases</p> <p>Reference:</p> <p>(1) Gelman et al. (2014), Chapter 16; Rossi et al. (2005), Chapter 5, Appendix A</p> <p>(2) Allenby and Ginter (1995), "Using Extremes to Design Products and Segment Markets," Journal of Marketing Research, 392-403.</p> <p>(3) Allenby and Rossi (1991), "Quality Perception and Asymmetric Switching Between Brands," Marketing Science 10, 185-205.</p>
第十一、十二讲	小组案例报告

五、教学方式

教师在课堂讲授并与学生探讨理论、算法和 R 程序。

六、教学过程中 IT 工具等技术手段的应用

学生在上机实习、做作业和期末项目时需要使用统计软件对实际数据进行贝叶斯分析。

七、教材

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2014), Bayesian Data Analysis, 3rd edition, CRC Press.

Rossi, P. E., Allenby, G. M. and McCulloch, R. (2005), Bayesian Statistics and Marketing, John Wiley & Sons, Ltd.

Greenberg, E. (2008), Introduction to Bayesian Econometrics, 1st edition, Cambridge University Press.

八、参考书目

Gelman, A. and Hill, J. (2006), Data Analysis Using Regression and Multilevel/Hierarchical Models, Cambridge University Press.

九、教学辅助材料，如 CD、录影等

无。

十、课程学习要求及课堂纪律规范

学生分为各小组，每个组由 3 人组成，在整个学期中互相合作与支持，并共同完成课程要求的作业和期末项目。

每次作业按照规定时间上交。迟于规定时间 24 小时内上交的作业，扣除该次作业总分的 50%。迟于规定时间 24 小时以外上交的作业，以零分计。

十一、学生成绩评定办法（需详细说明评估学生学习效果的方法）

先如下计算小组成绩：

作业： 50%

期末项目： 50%

期末时每位同学需要对所在小组内各位成员的贡献进行评分，每位同学的最终成绩在小组成绩的基础上根据贡献分进行调整。