课程大纲
COURSE SYLLABUS

1.	**课程代码/名称** **Course Code/Title**	**统计前沿选讲 II / Selected topics in frontiers of Statistics II** **STA5102**
2.	**课程性质** **Compulsory/Elective**	专业选修课 Major Elective Courses
3.	**开课单位** **Offering Dept.**	**统计系** Department of Statistics
4.	**课程学分/学时** **Course Credit/Hours**	48
5.	**授课语言** **Teaching Language**	英文 English
6.	**授课教师** **Instructor(s)**	陈欣 Chen Xin
7.	**开课学期** **Semester**	秋季 Fall Semester
8.	**是否面向本科生开放** **Open to undergraduates or not**	是
9.	**先修要求** **Pre-requisites**	（如面向本科生开放，请注明区分内容。 If the course is open to undergraduates, please indicate the difference.） 统计线性模型（MA329） Statistical Linear Models (MA329)

10. 教学目标
Course Objectives

（如面向本科生开放，请注明区分内容。 If the course is open to undergraduates, please indicate the difference.）

本课程首先讲述半参数回归里面的基本概念，它们是现代统计学的基石。本课程也为进一步学习其他领域比如大数据，深度学习，人工智能 打下良好的基础。半参数回归降低数据的复杂度，并且保留数据的重要特征，可应用于许多学科领域。基本教学目标是掌握比如正则化回归分析，样条和局部平均法，可加模型，混合模型，分位数半参数回归，以及半参数回归中的模型选择等。基本目标是教会学生掌握统计学习和现代统计方法,培养学生的统计学思维和分析数据的能力,并为后续课程打下良好的基础。

This course begins with an introduction to the basic concepts of semiparametric regression, which is the cornerstone of modern statistics. This course also lays a good foundation for further study in other areas such as big data, deep learning, and artificial intelligence. Semiparametric regression reduces complex data sets to summaries that we can understand well, while keeping essential features of the data. The basic teaching objectives are to master such as regularized regression analysis, Spline and local averaging methods, additive model, mixed models, quantile semiparametric regression and model selection in semiparametric regression. The basic goal is to teach students to master semiparametric regression and modern statistical methods, to develop students' statistical thinking and ability to analyse data, and to lay a good foundation for follow-up courses.

11. 教学方法
Teaching Methods

完成本课程后,学生应掌握半参数回归的基本概念和方法，熟悉各种半参数回归的方法和技巧，并能解决现实中的各类

数据分析问题。特别是,在学习本课程后,学生应该能够

1.掌握基本知识,深入理解和掌握半参数回归的各种概念和定理以及公式的本质。学生应该能够不仅记住概念，也要学会基本现代统计学方法， 同时能深刻理解如何利用统计深度学习解决问题。

2.掌握基本技能, 并能正确的进行数据分析。培养思维能力，提高对数据的分析能力，乃至概括的能力。

3.提高解决实际问题的能力。学习本课程后,学生应该能够使用学到的知识对实际问题建立合理模型， 从而解决相关的半参数回归问题。

After completing this course, students should master the basic concepts and methods of semiparametric regression, be familiar with various semiparametric regression methods and techniques and solve various types of data analysis problems. After studying this course, students should be able to

1. Master the basic knowledge, deeply understand and master the various concepts and theorems of semiparametric regression and the essence of the formula. Students should be able to not only remember concepts, but also basic modern statistical methods, while also having a deep understanding of how to use semiparametric regression to solve problems.

2. Master basic skills and perform data analysis correctly. Develop thinking skills, improve the ability to analyse data, and even generalize.

3. Improve the ability to solve practical problems. After studying this course, students should be able to use the knowledge they have acquired to develop a reasonable model of the actual problem to solve the relevant semiparametric regression problems.

12. 教学内容
Course Contents
（如面向本科生开放，请注明区分内容。 If the course is open to undergraduates, please indicate the difference.）

Section 1	1 介绍 （2 hours） 1. Introduction (2 hours)
Section 2	2 参数回归 （4 hours） 1.1 岭回归 1.2 LASSO 1.3 偏最小二乘法 2 Parametric Regression（4 hours） 2.1 Ridge Regression 2.2 Lasso 2.3 Partial Least Square
Section 3	3 样条和局部平均法 (8 hours) 3.1 方法 3.2 惩罚样条 3.3 CART 3.4 K-最近邻 3.5 核平滑方法 3 Splines & Local Averaging Methods （8 hours） 3.1 Methods 3.2 Penalized Splines 3.3 CART 3.4 K-nearest neighbours 3.5 Kernel smoothing

Section 4	4　可加模型 （16 hours） 4.1 方法 4.2 变系数模型 4.3 部分线性模型 4.4 广义可加模型 4 Additive Models （16 hours） 4.1 Methods 4.2 Varying Coefficient Model 4.3 Partial Linear Model 4.4 Generalized Additive Model
Section 5	5　混合模型（12 hours） 5.1 线性混合模型 5.2 半参数混合模型 5　Mixed Models (12 hours) 5.2 Linear Mix Models 5.2 Semiparametric Mix Models
Section 6	6. Selection of additional topics (6 hours) 6.1 Quantile Semiparametric Regression 6.2 Model selection in Semiparametric regression 6. 节选额外内容（6 小时） 6.1 分位数半参数回归 6.2 半参数回归中的模型选择
Section 7	
Section 8	
Section 9	
Section 10	
……	

13. 课程考核
Course Assessment

（①考核形式 Form of examination；②.分数构成 grading policy；③如面向本科生开放，请注明区分内容。If the course is open to undergraduates, please indicate the difference.）

平时成绩 General assessment（小测验以及作业, Quiz & Homework） 30%
期中考试 Midterm exam 30%
期末项目 Final project 40%

14. 教材及其它参考资料
Textbook and Supplementary Readings

1. Semiparametric Regression (Cambridge Series in Statistical and Probabilistic Mathematics) David Ruppert , M. P. Wand , R. J. Carroll.　2003

2.　Semiparametric Regression with R (Springer) Jaroslaw Harezlak, David Ruppert, Matt P. Wand. 2018