多元统计分析课程大纲

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

课程详述

COURSE SPECIFICATION

以下课程信息可能根据实际授课需要或在课程检讨之后产生变动。如对课程有任何疑问，请联系授课教师。

The course information as follows may be subject to change, either during the session because of unforeseen circumstances, or following review of the course at the end of the session. Queries about the course should be directed to the course instructor.

1.	**课程名称 Course Title**	多元统计分析 Multivariate Statistical Analysis
2.	**授课院系** **Originating Department**	数学系 Department of Mathematics
3.	**课程编号** **Course Code**	MA304
4.	**课程学分 Credit Value**	3
5.	**课程类别** **Course Type**	专业选修课 Major Elective Courses
6.	**授课学期** **Semester**	春 spring【2015 春季学期——2021 春季学期】
7.	**授课语言** **Teaching Language**	中文 Chinese
8.	**授课教师、所属学系、联系方式（如属团队授课，请列明其他授课教师）** **Instructor(s), Affiliation& Contact** **（For team teaching, please list all instructors）**	蒋学军， 数学系，0755-88018687 Xuejun Jiang, Department of Mathematics
9.	**实验员/助教、所属学系、联系方式** **Tutor/TA(s), Contact**	To be announced
10.	**选课人数限额(可不填)** **Maximum Enrolment（Optional）**	60

11. 授课方式 Delivery Method	讲授 Lectures	习题/辅导/讨论 Tutorials	实验/实习 Lab/Practical	其它(请具体注明) Other（Please specify）	总学时 Total
学时数 Credit Hours	48	12	6		48

12. 先修课程、其它学习要求 Pre-requisites or Other Academic Requirements	概率论（MA215），数理统计（MA204）） Probability (MA215)， Mathematical Statistics(MA204))
13. 后续课程、其它学习规划 Courses for which this course is a pre-requisite	统计计算，统计机器学习，高等统计学 Statistical computing, Statistical machine learning, advanced statistics
14. 其它要求修读本课程的学系 Cross-listing Dept.	

教学大纲及教学日历 SYLLABUS

15. 教学目标 Course Objectives

能够让学生掌握多元统计分析的基本概念、统计思想和常见的数据处理方法，并学会使用 SAS 软件包处理数据，分析实验结果。

This course aims to enable undergraduate students to master some basic concepts and theories in multivariate statistical analysis, to lay a solid foundation for the research in statistics and to master basic methods of data processing. The students should learn to use SAS to process data and analyze experimental results.

16. 预达学习成果 Learning Outcomes

通过学习此课程，学生能够做到如下：

1 掌握多元统计分析的基本概念、统计思想和数据处理方法

2 具备使用 SAS 软件包处理数据，分析实验结果的能力

On successful completion of the course, students should be able to:

1. master basic concepts and theories in multivariate statistical analysis;
2. master basic methods of data processing;
3. can use SAS to process data and analyze experimental results.

17. 课程内容及教学日历 （如授课语言以英文为主，则课程内容介绍可以用英文；如团队教学或模块教学，教学日历须注明主讲人）

Course Contents (in Parts/Chapters/Sections/Weeks. Please notify name of instructor for course section(s), if this is a team teaching or module course.)

第一章矩阵代数 (review 2 hours)

1.1 定义

1.2 矩阵的运算

1.3 行列式
1.4 矩阵的逆
1.5 矩阵的秩
1.6 特征值、特征向量和矩阵的迹
1.7 正定矩阵和非负定矩阵
1.8 特征值的极值问题

Ch1. Matrix algebra

1.1 Definition
1.2 Matrix manipulations
1.3 Determinant
1.4 Inverse of a matrix
1.5 Rank of a matrix
1.6 Eigenvalue, eigenvector and the trace of a matrix
1.7 Positive definite matrix and nonnegative definite matrix
1.8 Extremum problem of eigenvalue

第二章 随机向量 (4 hours)
2.1 一元分布

2.2 多元分布 (2.1-2.2, 2 hours)
2.3 数字特征

2.4 欧氏距离和马氏距离 (2.3-2.4，2 hours)

Ch2. Random vector
2.1 One dimensional distribution
2.2 Multivariate distribution
2.3 Digital features
2.4 Euclidean distance and Mahalanobis distance

第三章 多元正态分布 (6 hours)
3.1 多元正态分布的定义

3.2 多元正态分布的性质 (3.1-3.2, 2 hours)
3.3 复相关系数和偏相关系数 (2 hours)

3.4 极大似然估计及估计量的性质
3.5 样本均值和（n-1）S 的抽样分布 (3.4-3.5, 2 hours)

Ch3. Multivariate normal distribution

3.1 The definition of the multivariate normal distribution
3.2 The properties of the multivariate normal distribution
3.3 Multiple correlation coefficient and partial correlation coefficient
3.4 Maximum likelihood estimation and properties of estimators
3.5 Sample mean and (n-1)S sampling distribution

（第 1-4 周完成前三章教学内容）

第四章多元正态总体的统计推断 (6 hours)

4.1 一元情形的回顾 (skip)

4.2 单个总体均值的推断

4.3 单个总体均值分量间结构关系的检验 (4.2-4.3, 2 hours)

4.4 两个总体均值的比较推断

4.5 两个总体均值分量间结构关系的检验 (4.4-4.5, 2 hours)

4.6 多个总体均值的比较检验(多元方差分析)

4.7 总体相关系数的推断 (4.6-4.7, 2 hours)

Ch4. Statistical inference on Multivariate normal population

4.1 Review a case
4.2 A single population mean inference
4.3 Structure relationship test between single population mean component
4.4 Comparison and inference between two population mean
4.5 Structure relationship test between two population mean component
4.6 Comparative test of multiple population mean (multivariate analysis of variance)
4.7 The population correlation coefficient estimation

第五章判别分析 (6 hours)

5.1 引言

5.2 距离判别 (5.1-5.2, 2 hours)

5.3 贝叶斯判别 (2 hours)

5.4 费希尔判别 (2 hours)

Ch5. Discriminant analysis

5.1 Introduction
5.2 Distance discriminant
5.3 Bias discriminant
5.4 Fisher discriminant

(第 1 到 8 周完成前五章教学内容)

第八周, 期中考试,

第六章 聚类分析 (2 hours)

6.1 引言

6.2 距离和相似系数

6.3 系统聚类法

6.4 动态聚类法

Ch6. Cluster analysis

6.1 Introduction
6.2 Distance and similarity coefficient
6.3 Hierarchical clustering method
6.4 Dynamic clustering method

第七章 主成分分析 (4 hours)

7.1 引言

7.2 总体的主成分 (2 hours)

7.3 样本的主成分 (2 hours)

Ch7. Principal component analysis

7.1 Introduction
7.2 Principal component of population
7.3 Principal component of sample

完成前 7 章后，第十三周布置一个 Project

第八章 因子分析 (6 hours)
8.1 引言
8.2 正交因子模型
8.3 参数估计 （ 8.1-8.3, 2 hours）

8.4 因子旋转 (2 hours)
8.5 因子得分 (2 hours)

Ch8. Factor analysis

8.1 Introduction
8.2 Orthogonal factor model
8.3 Parameter estimation
8.4 Factor rotation
8.5 Factor score

第九章 对应分析 (自我阅读)
9.1 行轮廓和列轮廓
9.2 独立性的检验和总惯性
9.3 行、列轮廓的坐标
9.4 对应分析图

Ch9. Correspondence analysis

9.1 Line profile and column profile
9.2 Independence test and Inertia
9.3 Row, column profile coordinates
9.4 Correspondence analysis graph

第十章 典型相关分析 （6 hours）

10.1 引言
10.2 总体典型相关 (2 hours)
10.3 样本典型相关 (2 hours)
10.4 典型相关系数的显著性检验 (2 hours)

第十六周， 期末复习小结（2 hours）

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

教材(Required): 应用多元分析第四版，上海财经大学出版社，王学民，2014 年 09 月

参考资料(Recommended)：
Applied Multivariate Mehtods for Data Analysis, Dallase. Johnson. Higher Education Press(影印版)

课程评估 ASSESSMENT

19.

评估形式 Type of Assessment	评估时间 Time	占考试总成绩百分比 % of final score	违纪处罚 Penalty	备注 Notes
出勤 Attendance		5%		One point penalized for absence without leave each time
课堂表现 Class Performance				
小测验 Quiz				
课程项目 Projects		20%		用 SAS 编程及实验结果分析
平时作业 Assignments		15%		
期中考试 Mid-Term Test		30%		
期末考试 Final Exam		30		
期末报告 Final Presentation				
其它（可根据需要改写以上评估方式） Others (The above may be modified as necessary)				

20. 记分方式 GRADING SYSTEM

☑ A. 十三级等级制 Letter Grading
☐ B. 二级记分制（通过/不通过） Pass/Fail Grading

课程审批 REVIEW AND APPROVAL

21. 本课程设置已经过以下责任人/委员会审议通过
This Course has been approved by the following person or committee of authority

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

课程详述

COURSE SPECIFICATION

以下课程信息可能根据实际授课需要或在课程检讨之后产生变动。如对课程有任何疑问，请联系授课教师。

The course information as follows may be subject to change, either during the session because of unforeseen circumstances, or following review of the course at the end of the session. Queries about the course should be directed to the course instructor.

1.	**课程名称 Course Title**	多元统计分析 Multivariate Statistical Analysis
2.	**授课院系** **Originating Department**	统计与数据科学系 Department of Statistics and Data Science
3.	**课程编号** **Course Code**	MA304
4.	**课程学分 Credit Value**	3
5.	**课程类别** **Course Type**	专业选修课 Major Elective Courses
6.	**授课学期** **Semester**	春 spring【2022 春季学期起】
7.	**授课语言** **Teaching Language**	中文 Chinese
8.	**授课教师、所属学系、联系方式（如属团队授课，请列明其他授课教师）** **Instructor(s), Affiliation& Contact** **（For team teaching, please list all instructors）**	蒋学军， 统计与数据科学系，0755-88018687 Xuejun Jiang, Department of Statistics and Data Science
9.	**实验员/助教、所属学系、联系方式** **Tutor/TA(s), Contact**	To be announced
10.	**选课人数限额(可不填)** **Maximum Enrolment（Optional）**	60

11. 授课方式 Delivery Method	讲授 Lectures	习题/辅导/讨论 Tutorials	实验/实习 Lab/Practical	其它(请具体注明) Other（Please specify）	总学时 Total
学时数 Credit Hours	48	12	6		48

12. 先修课程、其它学习要求 Pre-requisites or Other Academic Requirements	概率论与数理统计（MA212）
13. 后续课程、其它学习规划 Courses for which this course is a pre-requisite	统计计算，统计机器学习，高等统计学 Statistical computing, Statistical machine learning, advanced statistics
14. 其它要求修读本课程的学系 Cross-listing Dept.	

教学大纲及教学日历 SYLLABUS

15. 教学目标 Course Objectives

能够让学生掌握多元统计分析的基本概念、统计思想和常见的数据处理方法，并学会使用 SAS 软件包处理数据，分析实验结果。

This course aims to enable undergraduate students to master some basic concepts and theories in multivariate statistical analysis, to lay a solid foundation for the research in statistics and to master basic methods of data processing. The students should learn to use SAS to process data and analyze experimental results.

16. 预达学习成果 Learning Outcomes

通过学习此课程，学生能够做到如下：

1 掌握多元统计分析的基本概念、统计思想和数据处理方法

2 具备使用 SAS 软件包处理数据，分析实验结果的能力

On successful completion of the course, students should be able to:

1. master basic concepts and theories in multivariate statistical analysis;
2. master basic methods of data processing;
3. can use SAS to process data and analyze experimental results.

17. 课程内容及教学日历 （如授课语言以英文为主，则课程内容介绍可以用英文；如团队教学或模块教学，教学日历须注明主讲人）

Course Contents (in Parts/Chapters/Sections/Weeks. Please notify name of instructor for course section(s), if this is a team teaching or module course.)

第一章矩阵代数 (review 2 hours)

1.1 定义

1.2 矩阵的运算

1.3 行列式
1.4 矩阵的逆
1.5 矩阵的秩
1.6 特征值、特征向量和矩阵的迹
1.7 正定矩阵和非负定矩阵
1.8 特征值的极值问题

Ch1. Matrix algebra

1.1 Definition
1.2 Matrix manipulations
1.3 Determinant
1.4 Inverse of a matrix
1.5 Rank of a matrix
1.6 Eigenvalue, eigenvector and the trace of a matrix
1.7 Positive definite matrix and nonnegative definite matrix
1.8 Extremum problem of eigenvalue

第二章 随机向量 (4 hours)
2.1 一元分布

2.2 多元分布 (2.1-2.2, 2 hours)
2.3 数字特征

2.4 欧氏距离和马氏距离 (2.3-2.4，2 hours)

Ch2. Random vector
2.1 One dimensional distribution
2.2 Multivariate distribution
2.3 Digital features
2.4 Euclidean distance and Mahalanobis distance

第三章 多元正态分布 (6 hours)
3.1 多元正态分布的定义

3.2 多元正态分布的性质 (3.1-3.2, 2 hours)
3.3 复相关系数和偏相关系数 (2 hours)

3.4 极大似然估计及估计量的性质
3.5 样本均值和（n-1）S 的抽样分布 (3.4-3.5, 2 hours)

Ch3. Multivariate normal distribution

3.1 The definition of the multivariate normal distribution
3.2 The properties of the multivariate normal distribution
3.3 Multiple correlation coefficient and partial correlation coefficient
3.4 Maximum likelihood estimation and properties of estimators
3.5 Sample mean and (n-1)S sampling distribution

（第 1-4 周完成前三章教学内容）

第四章多元正态总体的统计推断 (6 hours)

4.1 一元情形的回顾 (skip)

4.2 单个总体均值的推断

4.3 单个总体均值分量间结构关系的检验 (4.2-4.3, 2 hours)

4.4 两个总体均值的比较推断

4.5 两个总体均值分量间结构关系的检验 (4.4-4.5, 2 hours)

4.6 多个总体均值的比较检验(多元方差分析)

4.7 总体相关系数的推断 (4.6-4.7, 2 hours)

Ch4. Statistical inference on Multivariate normal population

4.1 Review a case
4.2 A single population mean inference
4.3 Structure relationship test between single population mean component
4.4 Comparison and inference between two population mean
4.5 Structure relationship test between two population mean component
4.6 Comparative test of multiple population mean (multivariate analysis of variance)
4.7 The population correlation coefficient estimation

第五章判别分析 (6 hours)

5.1 引言

5.2 距离判别 (5.1-5.2, 2 hours)

5.3 贝叶斯判别 (2 hours)

5.4 费希尔判别 (2 hours)

Ch5. Discriminant analysis

5.1 Introduction
5.2 Distance discriminant
5.3 Bias discriminant
5.4 Fisher discriminant

(第 1 到 8 周完成前五章教学内容)

第八周, 期中考试,

第六章 聚类分析 (2 hours)

6.1 引言

6.2 距离和相似系数

6.3 系统聚类法

6.4 动态聚类法

Ch6. Cluster analysis

6.1 Introduction
6.2 Distance and similarity coefficient
6.3 Hierarchical clustering method
6.4 Dynamic clustering method

第七章 主成分分析 (4 hours)

7.1 引言

7.2 总体的主成分 (2 hours)

7.3 样本的主成分 (2 hours)

Ch7. Principal component analysis

7.1 Introduction
7.2 Principal component of population
7.3 Principal component of sample

完成前 7 章后，第十三周布置一个 Project

第八章 因子分析 (6 hours)

8.1 引言

8.2 正交因子模型

8.3 参数估计 （ 8.1-8.3, 2 hours）

8.4 因子旋转 (2 hours)

8.5 因子得分 (2 hours)

Ch8. Factor analysis

8.1 Introduction
8.2 Orthogonal factor model
8.3 Parameter estimation
8.4 Factor rotation
8.5 Factor score

第九章 对应分析 (自我阅读)

9.1 行轮廓和列轮廓

9.2 独立性的检验和总惯性

9.3 行、列轮廓的坐标

9.4 对应分析图

Ch9. Correspondence analysis

9.1 Line profile and column profile
9.2 Independence test and Inertia
9.3 Row, column profile coordinates
9.4 Correspondence analysis graph

第十章 典型相关分析 （6 hours）

10.1 引言

10.2 总体典型相关 (2 hours)

10.3 样本典型相关 (2 hours)

10.4 典型相关系数的显著性检验 (2 hours)

第十六周， 期末复习小结（2 hours）

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

教材(Required): 应用多元分析第四版，上海财经大学出版社，王学民，2014 年 09 月

参考资料(Recommended)：
Applied Multivariate Mehtods for Data Analysis, Dallase. Johnson. Higher Education Press(影印版)

课程评估 ASSESSMENT

19.

评估形式 Type of Assessment	评估时间 Time	占考试总成绩百分比 % of final score	违纪处罚 Penalty	备注 Notes
出勤 Attendance		5%		One point penalized for absence without leave each time
课堂表现 Class Performance				
小测验 Quiz				
课程项目 Projects		25%		
平时作业 Assignments		20%		
期中考试 Mid-Term Test				
期末考试 Final Exam		50%		
期末报告 Final Presentation				
其它（可根据需要改写以上评估方式） Others (The above may be modified as necessary)				

20. 记分方式 GRADING SYSTEM

☑ A. 十三级等级制 Letter Grading
☐ B. 二级记分制（通过/不通过） Pass/Fail Grading

课程审批 REVIEW AND APPROVAL

21. 本课程设置已经过以下责任人/委员会审议通过
This Course has been approved by the following person or committee of authority