

# 中国海洋大学本科生课程大纲

课程名称	多元统计分析 Multivariate Statistical Analysis	课程代码	075302101213
课程属性	专业知识	课时/学分	48/3
课程性质	选修	实践学时	
责任教师	类淑河	课外学时	96

**课程属性：**公共基础/通识教育/学科基础/专业知识/工作技能，**课程性质：**必修、选修

## 一、课程介绍

### 1. 课程描述

多元统计分析是统计学的一个重要分支，主要讨论多变量（多指标）观测数据的统计推断方法和多元数据的分析处理技术。课程教学内容包括理论与实践两个方面。

理论方面可分为两部分，一部分为多元统计推断的理论与方法，包括多元总体与样本、多元统计量（均值向量、协方差矩阵、相关矩阵）、多元正态分布及多元抽样分布理论、多元正态总体参数的估计和假设检验、多元方差分析等；另一部分为多元数据处理方法与技术，包括主成分分析、因子分析、聚类分析、判别分析等。前一部分中的多元统计推断，尤其是多元正态总体参数的各种假设检验，主要在正态总体的分布假定之上展开，而后一部分中的多元统计分析方法与技术，往往可以不依赖于分布正态性的假定。

实践方面，对应的，也可分为两部分，一是多元统计推断理论结果的模拟验证，可以借助软件产生随机数，结合分布拟合检验方法验证多元抽样分布理论；另一部分是各种多元分析方法的实现，要求学生在理解理论方法的基础上，尽量自己编写程序实现相应的多元分析方法，结合模拟或实际数据，与应用软件提供的分析函数的处理结果进行比较，给出合理的结果分析。

### 2. 设计思路

随着数据采集、存储技术的不断进步，各行各业、各学科领域积累的数据量越来越庞大，人们对于数据分析方法的需求也越来越强烈。在各种实际观测数据中，多元数据，即多指标观测数据，是最常见的数据类型。因而，多元统计分析方法一直是数

据分析方法中最普遍使用的方法。

课程内容的选择，兼顾课程的宽度与深度，侧重理论方法及其原理。对于大部分常用多元推断方法和多元分析技术，强化理论方法的推演和讨论，使学生在学过本课程之后，具备较坚实的多元分析理论基础。

在内容编排上，前面三章先讨论多元统计推断方法及其理论基础，结构上与一元数理统计平行，从总体、样本、统计量，到多元抽样分布、参数估计方法、各类检验问题，概念、方法、结论，有相似之处，更有多元统计的特有新内容。后面几章讨论常用多元数据分析方法和技术，包括降维技术、分类与判别技术等。

### 3. 课程与其他课程的关系

先修课程：高等代数、概率论、数理统计学；

并行课程：随机过程、非参数统计、应用回归分析；

后置课程：数据挖掘、时间序列分析。

本课程需要学生具备概率论、数理统计以及矩阵代数的基础，与随机过程、时间序列分析、数据挖掘等课程构成了统计方向的专业课程群。

## 二、课程目标

本课程旨在培养学生熟悉多元统计的思想方法和分析解决实际问题的能力，掌握多元统计推断方法和多元数据分析处理技术，与概率论、数理统计学、随机过程、时间序列分析等课程一起，培养学生形成看待分析实际问题的统计视角和统计思维。

到课程结束时，学生应能：

- (1) 理解多元抽样分布理论，掌握多元正态总体的参数估计和假设检验方法；
- (2) 掌握各种多元数据分析方法的原理、统计思想、适用问题及应用限制；
- (3) 利用自己熟悉的统计软件（如 R 软件包、Matlab 等）实现多元统计推断和多元数据分析方法；
- (4) 面对多元数据或多指标实际问题，能够借助多元统计的概念与模型，正确的抽象、描述实际问题，并能够选择恰当的多元分析方法，合理的分析处理多元数据；针对实际问题让学生开展交流学习，并通过口头报告或书面作业的形式提供学习讨论结果；激发同学深入理解多元统计分析所表达的处理实际问题时所遵循的理念，提升提出问题并解决问题的能力。

### 三、学习要求

要完成所有的课程任务，学生必须：

(1) 按时上课,上课认真听讲,积极参与课堂讨论、随堂练习和测试。本课程将包含较多的课堂提问、讨论等课堂活动,课堂表现和出勤率是成绩考核的组成部分。

(2) 按时完成常规练习作业。这些作业要求学生按书面形式提交,只有按时提交作业,才能掌握课程所要求的内容。延期提交作业需要提前得到任课教师的许可。

(3) 完成教师布置的一定量的文献阅读和背景资料、案例分析、理论探讨和较大量的实际数据分析、方法软件实现、随机模拟等作业,其中部分内容允许以小组合作形式完成。这些作业能加深对课程内容的理解、促进同学间的相互学习、并能引导对某些问题和理论的更深入探讨。

### 四、参考教材

#### 1、选用教材：

《应用多元统计分析》（第1版），高惠璇，北京大学出版社，2005年1月。

#### 2、主要参考书：

- 1) 张尧庭, 方开泰.《多元统计分析引论》.科学出版社.1997
- 2) M.肯德尔,《多元分析》.中国科学院计算中心概率统计组译.科学出版社.1983
- 3)王斌会,《多元统计分析及R语言建模》.暨南大学出版社,2016年3月第4版
- 4) Anderson, T. W. (1984), An Introduction to Multivariate Statistical Analysis (2nd Ed.), John Wiley & Sons, Inc., Reprinted by World Publishing Corp. Beijing, 1990.
- 5) Richard A. Johnson & Dean W. Wichern , Applied Multivariate Statistical Analysis, Printice Hall Inc (1998); 中文译本: 陆璇译, 清华大学出版社, 2001年4月 第四版

### 五、进度计划

序号	专题	主题	计划课时	主要内容概述	实验实践内容
1	第一章 第二章	绪论简介 多元正态分布及参数估计	10	随机向量,多元正态分布定义与性质,条件分布和独立性,随机阵的正态分布,多元正态分布的参数估计	书面作业: 习题二中的部分题目

2	第三章	多元正态总体参数的假设检验	20	几个重要统计量的分布, 单总体均值向量的检验及置信域, 多总体均值向量的检验, 协方差阵的检验, 独立性检验, 正态性检验	书面作业: 习题三中的部分题目
3	第五章	判别分析	10	距离判别, 贝叶斯判别法, 费希尔判别法	书面作业: 习题五中的部分题目
4	第六章	聚类分析	10	聚类分析的方法, 距离与相似系数, 系统聚类法, 动态聚类法, 有序样品聚类法	书面作业: 习题六中的部分题目
5	第七章	主成分分析	8	总体的主成分, 样本的主成分, 主成分分析的应用	书面作业: 习题七中的部分题目
6	第八章	因子分析	6	因子分析引言, 因子模型	第五章---第七章个别题目需要编程完成

注: 带\*号的学时为实践学时;

## 六、成绩评定

(一) 考核方式  A : A. 闭卷考试 B. 开卷考试 C. 论文 D. 考查 E. 其他

(二) 成绩综合评分体系:

成绩综合评分体系	比例%
1. 课下作业、课堂讨论及平常表现	15
2. 平时测验成绩	15
3. 期末考试成绩	70
总计	100

## 附：作业和平时表现评分标准

### 1) 作业的评分标准

作业的评分标准	得分
1.严格按照作业要求并及时完成，基本概念清晰，解决问题的方案正确、合理，能提出不同的解决问题方案。	90-100分
2.基本按照作业要求并及时完成，基本概念基本清晰，解决问题的方案基本正确、基本合理。	70-80分
3.不能按照作业要求，未按时完成，基本概念不清晰，解决问题的方案基本不正确、基本不合理。	40-60分
4.不能按照作业要求，未按时完成，基本概念不清晰，不能制定正确和合理解决问题的方案。	0-30分

### 2) 课堂讨论及平时表现评分标准

课堂讨论、平常表现评分标准	得分
1.资料的查阅、知识熟练运用，积极参与讨论、能阐明自己的观点和想法，能与其他同学合作、交流，共同解决问题。	90-100分
2.基本做到资料的查阅、知识的运用，能参与讨论、能阐明自己的观点和想法，能与其他其他同学合作、交流，共同解决问题。	70-80分
3.做到一些资料的查阅和知识的运用，参与讨论一般、不能阐明自己的观点和想法，与其他同学合作、交流，共同解决问题的能力态度一般。	40-60分
4.不能做到资料的查阅和知识的运用，不积极参与讨论，不能与其他同学合作、交流，共同解决问题。	0-30分

## 七、学术诚信

学习成果不能造假，如考试作弊、盗取他人学习成果、一份报告用于不同的课程等，均属造假行为。他人的想法、说法和意见如不注明出处按盗用论处。本课程如有发现上述不良行为，将按学校有关规定取消本课程的学习成绩。

## 八、大纲审核

教学院长：

院学术委员会签章：