

中国海洋大学本科生课程大纲

课程名称	应用回归分析 Applied Regression Analysis	课程代码	075123201341
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	孙志华	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：

回归分析是应用极其广泛的数据分析方法之一。它基于观测数据建立变量间适当的依赖关系，以分析数据内在规律，并可用于预报、控制等问题的科学。本课程针对高年级统计学及应用数学专业学生开设，课程包括回归分析的若干基本内容：一元线性回归、多元线性回归模型的参数估计和回归方差的显著性检验，异常值和强影响值判别，异方差性、自相关性、多重共线性的诊断，逐步回归和非线性回归等。通过课程学习，要求学生掌握回归分析的若干基本理论和方法，能够利用这些理论方法并借助计算机软件对实际问题进行建模、分析和求解，进而提升对统计学的理解。

2. 设计思路：

本课程引导高年级统计学及应用数学专业学生通过回归分析来探讨和理解由实际问题所驱动的数学在理论和应用两方面的发展途径。课程内容的选取基于学生“掌握了概率论和数理统计等内容”。课程内容包括三个模块：线性回归分析、回归模型的诊断、变量选择方法等。

线性回归分析是确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法，运用十分广泛，课程以回归分析为中心展开，在变量之间的关系是线性情况下的一元线性回归分析和多元线性回归分析。主要包括：回归参数的估计、回归方程的

显著性检验、预测和控制等，同时给出线性回归分析模型的应用案例。

回归模型的诊断是对回归模型及相应的统计推断方法进行合理性研究，检查数据和模型以及统计推断方法中可能存在的问题，问题的解决方案，主要包括残差分析，异常点和强影响点判别，异方差性诊断，自相关性诊断、多重共线性诊断。课程将以实际问题为引导，强调回归模型的诊断的基本思想和解决问题方案，同时给出回归分析诊断的应用案例。

变量选择的研究是统计学的一个重要问题，随着科学研究的深入，数据复杂程度的增加，针对复杂数据和复杂模型的变量选择问题成为统计学的一个研究热点问题。变量选择的方法主要有：子集回归、逐步回归、系数压缩等方法，同时给出变量选择的应用案例。

3. 课程与其他课程的关系：

先修课程：概率论、数理统计；

并行课程：统计软件、时间序列分析、随机过程等；

后置课程：非参数统计、多元统计分析等。

二、课程目标

本课程目标是引导并培养学生用统计学回归思维来描述和解决统计问题的能力，增强沟通能力和团队合作意识。

到课程结束时，学生应能：

- (1) 理解并掌握经典的一元线性回归、多元线性回归模型的参数估计和回归方差的显著性检验的方法。
- (2) 掌握异常值和强影响值判别，异方差性、自相关性、多重共线性的诊断的方法。
- (3) 掌握子集回归、逐步回归、系数压缩等变量选择方法
- (4) 了解 Logistic 回归。
- (5) 掌握用 R 统计软件实现回归分析、回归诊断、变量选择的过程，加强对所学回归分析知识的理解。
- (6) 针对实际问题开展小组研究，并通过口头报告或书面研究报告形式提供研究结果；激发同学深入理解处理实际问题时所遵循的统计理念，提升解决

统计分析问题的能力。

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随堂练习、讨论、小组作业展示等课堂活动, 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阅读文献和背景资料、案例分析、理论探讨和软件应用等作业。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、参考教材与主要参考书

1、选用教材：

《统计学精品译丛：例解回归分析(原书第5版)》，英文书名:Regression Analysis by Example, Fifth Edition, Samprit Chatterjee, Ali S.Hadi 著；郑忠国，许静译，机械工业出版社。

2、主要参考书：

[1] 《应用回归分析》，何晓群编著，中国人民大学出版社，2015年4月。

[2] 《A Modern Approach to Regression with R》，Simon J. Sheather, Springer Texts in Statistics, 2008年。

五、进度安排

序号	专题	主题	计划课时	主要内容概述	实验实践内容
----	----	----	------	--------	--------

1	绪论	回归分析简介	2	回归分析起源、发展，回归分析实例，回归分析步骤简介等	
2	一元线性回归	一元线性回归模型	4	一元线性回归实例，模型，参数估计	
		假设检验与置信区间	2	一元线性回归的假设检验，置信区间	
		预测	2	一元线性回归的预测；拟合等	R 统计软件实现：教材作业、案例分析
3	多元线性回归	参数估计	3	多元线性回归实例，参数估计	
		模型处理	2	多元线性回归系数的解释，中心化，规范化	
		最小二乘估计	4	最小二乘估计及性质	
		相关系数、检验、预测	4	复相关系数，单个相关系数推断、线性模型的假设检验、预测	R 统计软件实现：教材作业、案例分析
4	回归诊断	回归诊断、异常点判定	4	标准回归假定，残差，图形诊断方法，杠杆、强影响点、异常值，Cook 距离	
		回归方程中变量的作用	2	添加变量及稳健回归	R 统计软件实现：教材作业、案例分析
5	定性预测变量	定性变量模型	2	定性变量回归方程，回归参数随时间的稳定性	
6	加权最小二乘	加权最小二乘	2	异方差处理的加权最小二乘	R 统计软件实现：教材作业、案例分析
7	自相关	自相关	3	自相关性及其处理方法	
8	共线性	共线性定义	2	共线性判定方法	
		主成分回归	2	主成分回归法基本思想、步骤、注意事项	
		岭回归	2	岭回归方法及注解	R 统计软件实现：教材作业、案例分析
9	变量选择	变量选择准则、方法	4	回归模型估计预测控制，变量选择的几种准则，变量选择的方法，注意事项，实例	R 统计软件实现：教材作业、案例分析
10	Logistic 回归	Logistic 回归	2	Logistic 回归简介	

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、课堂讨论及平时表现	30
2. 平时测验成绩	20
3. 期末考试成绩	50
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：