

# 中国海洋大学本科生课程大纲

课程名称	统计软件 (R) Statistical Software (R)	课程代码	075104202111
课程属性	专业知识	课时/学分	1 学分
课程性质	选修	实践学时	32
责任教师	孙志华	课外学时	32

**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：

统计软件是统计学专业的一门重要课程，通过本课程的学习，使学生掌握一门实用而又专业性很强的技能，同时为学生学习其他专业课程、从事科学研究奠定前提和基础。统计软件在介绍了相关统计软件的基础上，系统阐述了 R 语言统计软件的基本理论和方法，侧重培养使用 R 语言统计软件进行数据处理和统计分析的能力。通过本课程学习，要求学生掌握 R 语言统计软件的基本统计分析方法和应用，能够利用 R 语言软件对统计问题进行建模、分析和基本的编程技术，为分析复杂统计问题奠定基础，并为深入学习其他统计理论提供有力的实际分析工具。

### 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统计软件学习，提高统计分析的实践能力。课程内容的选取基于学生“掌握了数理统计、回归分析等内容”。课程内容包括三个模块：R 语言基本原理与基本操作、概率分布与数据分析、统计推断的 R 语言实现等。

R 语言基本原理与基本操作模块主要来介绍 R 语言，主要包括 R 的特点，程序包安装，R 的基本操作与编程，数据读取，R 的图形功能等。概率分布与数据分析模块的任务是：掌握抽样与常见分布的 R 的实现，R 函数进行探索性数据分析，直方图以及数据的描述性统计分析等。统计推断的 R 语言实现模块主要来介绍参数估计、假设检验、

方差分析、回归分析等统计推断问题的 R 语言实现

### 3. 课程与其他课程的关系：

先修课程：编程语言、概率论、数理统计；

并行课程：回归分析、时间序列分析及应用（R 语言）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 R 语言统计软件来描述和解决统计数据分析问题的能力。

到课程结束时，学生应能：

- (1) 掌握 R 语言统计软件的若干基本内容：R 语言软件的基本原理、基本操作、R 语言的数据结构，R 语言编程的方法等。
- (2) 掌握 R 语言统计软件的操作，独立完成从建立数据文件到各种统计分析的操作等。
- (3) 掌握 R 语言统计软件的基本统计分析方法和应用，能够利用 R 语言软件对统计问题进行建模、分析和基本的编程技术。
- (4) 了解用 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。

## 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上机练习、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及上机操作练习。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：

《R 语言与统计分析》 ， 汤银才编， 高等教育出版社， 2008 年。

##### 2、主要参考书：

[1] 《 统计建模与 R 软件》 ， 薛毅， 陈立萍， 清华大学出版社， 2007 年。

[2] 《R 语言实战》 ， Robert I. Kabacoff 著， 图灵程序设计图书； 高涛， 肖楠， 陈钢译， 人民邮电出版社， 2013 年。

[3] 《R 语言统计入门（第 2 版）》 ， Peter Dalgaard 著， Springer 系列图书； 郝智恒， 何通， 邓一硕， 刘旭华译， 人民邮电出版社， 2014 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	统计软件介绍	1	主要统计软件的介绍， R 语言介绍、安装等	
2	R 的基本原理与核心	R 的基本原理	1	R 的基本原理， 在线帮助， 数据结构	
		数据存储读取与图形功能	2	R 语言的存储与读取， 常用统计作图函数	
		R 函数编程	1	R 的分支结构， 循环结构， 程序调试	上机实践本章学习内容及相关习题
3	概率与分布	随机抽样， 排列组合与概率的计算	1	随机抽样， choose()函数	
		概率分布与内嵌分布	2	密度函数， 分位数函数， 随机数函数， 内嵌的分布等	
		中心极限定理	1	中心极限定理	上机实践本章学习内容及相关习题
4	探索性数据分析	直方图与密度函数估计	2	概率函数图， 直方图进行数据分析	
		单组数据描述性统计	1	位置度量、分散度量、分散形状等的统计量函数实现	

		多组数据描述性统计	2	样本协方差相关系数及相关性检验函数	
		分类数据的描述性统计	1	列联表, 边际列表, 频率列联表等	上机实践本章学习内容及相关习题
5	参数估计	矩估计与最大似然估计	2	矩估计, 最大似然估计	
		区间估计	2	单个正态总体、两正态总体参数的区间估计	上机实践本章学习内容及相关习题
6	参数的假设检验	假设检验	3	单正态、两正态总体的假设检验与 $p$ 值, 成对数据的检验, 样本比率检验	上机实践本章学习内容及相关习题
7	方差分析	方差分析	3	单因子、双因子方差分析, 协方差分析	上机实践本章学习内容及相关习题
8	回归分析与相关分析	相关性及度量	1	相关性及度量	
		线性回归	2	一元线性回归, 多元线性回归	
		回归诊断	2	回归诊断	
		Logistic 回归	2	Logistic 回归	上机实践本章学习内容及相关习题

## 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	20
2. 平时测验成绩	30
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，未及时完成，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，未及时完成，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：