

南京信息工程大学滨江学院

2021 一 2022 学年 第 二 学期

多元统计分析 课程试卷

试卷类型 A 卷 (注明 A、B 卷) 考试类型 闭卷 (注明开、闭卷)

- 注意：1、本课程为必修 (注明必修或选修)， 学时为 48， 学分为 3
- 2、本试卷共 8 页；考试时间 120 分钟； 出卷时间： 2022 年 5 月
- 3、姓名、学号等必须写在指定地方； 考试时间： 2022 年 6 月 日
- 4、本考卷适用专业年级： 20 大数据 任课教师： 刘伟

题 号	一	二	三	四	总 分
得 分					
阅卷人					

(以上内容为教师填写)

请仔细阅读以下内容：

1、 考生必须遵守考试纪律，详细内容见《南京信息工程大学滨江学院考试纪律规定》。

2、 所有考试材料不得带离考场。

3、 考生进入考场后，须将学生证或身份证放在座位的左上角。

4、 考场内不许抽烟、吃食物、喝饮料。

5、 考生不得将书籍、作业、笔记、草稿纸袋入考场，主考教师允许带入的除外。

6、 考试过程中，不允许考生使用通讯工具。

7、 开考 15 分钟后不允许考生进入考场，考试进行 30 分钟后方可离场。

8、 考生之间不得进行任何形式的信息交流。

9、 除非被允许，否则考生交卷后才能离开座位。

10、 考试违纪或作弊的同学将被请出考场，其违纪或作弊行为将上报学院。

被人郑重承诺：我已阅读上述 10 项规定，如果考试是违反了上述 10 项规定，本人将自愿接受学校按照有关规定所进行的处理。上面姓名栏所填姓名即表示本人已阅读本框的内容并签名。

一、填空题（每空 1 分，共 20 分）

- 1、多元线性模型的回归模型总体显著性检验又称为，回归系数的显著性检验又称为。
- 2、拟合优度的检验中总离差平方和 TSS=，多样本可决系数=。
- 3、多元线性模型的回归预测包括和。
- 4、多元线性模型的参数估计方法称为，广义线性模型的参数估计方法称为。
- 5、在聚类分析中，对样本进行聚类称为，对变量进行聚类称为。
- 6、设  $X, Y$  是来自均值向量为，协方差阵为的总体  $G$  的两个样品，则  $X$  与  $Y$  之间的马氏距离， $X$  与总体  $G$  的马氏距离。
- 7、Bayes 判别准则分为和。
- 8、主成分分析主要起着和的作用。
- 9、建立因子模型首先要估计载荷矩阵与特殊方差，常用的估计方法有、和以及极大似然法。
- 10、因子正交旋转包括：和斜交旋转，一般认为 KMO 超过 适合做因子分析。

二、选择题（每题 2 分，共 20 分）

- 11、正态随机向量的（ ）等价于协方差为 0 矩阵
- A . 线性相关性 B. 非线性相关性 C. 独立性 D. 相似性
- 12、下列不属于多元线性回归经典假设的是（ ）。

A . 解释变量是随机的

B. 随机干扰项与解释变量不相关

C. 随机干扰项服从正态分布

D. 正确设定回归模型

13、可决系数的数值越接近于（ ），残差平方和越小，样本回归线于样本观测值的拟合程度越高。

A . 0            B. 1            C. -1            D. 2

14、聚类分析是（ ）学习过程。

A . 有监督        B. 无监督        C. 半监督        D. 自组织

15、下列不能体现降维思想的分析方法是（ ）

A . 判别分析        B. 主成分分析        C. 因子分析        D. 对应分析

16、k-means 聚类的主要因素不包括（ ）。

A . 初始中心点

B. 输入数据及 k 值得选择

C. 距离度量

D. 收敛值

17、因子分析中正交变换是否改变公因子的共同度（ ）。

A . 改变        B. 不改变        C. 可能改变        D 不确定

18、利用主成分分析得到的各个主成分之间（ ）。

A . 相互独立        B. 线性相关        C. 不完全相关        D. 互不相关

19、R 语言中，将矩阵 B 按列求均值正确的是（ ）。

A. aply(B,1,mean)

B. aply(B,2,mean)

C. aply(B, mean,1)

D. aply(B, mean,2)

20、R 语言中，求矩阵 A 的逆矩阵的函数是（ ）。

A . chol        B. svd        C.diag        D solve

三、计算题（每题 10 分，共 30 分）

21、设两个二维总体有公共协方差，从二总体中分别抽取了容量为 9 和 8 的样本，其数据如下，采用马氏距离判别法确定（24.5， 9.8）属于哪一类。

		1	2	3	4	5	6	7	8	9
G1	X1	20.2	28.5	24.6	26.5	29.0	36.7	36.0	27.6	25.0
	X2	14.2	8.4	14.8	15.2	11.9	9.6	18.0	9.5	16.0
G2	X1	21.6	23.2	17.8	14.4	11.0	18.9	15.4	16.1	
	X2	7.5	4.5	9.2	7.2	5.5	6.3	8.0	7.8	

22、一所学校中男生占 70%，女生占 30%，女生 80%留长发，男生 5%留长发，随机选择一个长发的学生，他(她)是女生的概率多大？

23、（1）构建一个函数相应 R 语言实现程序（指定函数名为 normalize），

（2）假设将该函数保存在当地目录“C:\\Users\\Desktop\\Output”下，文件名为 Norm.R，给出调用该函数的 R 命令。

四、简答题（每题 6 分，共 30 分）

24、简述分类与聚类的主要区别

25、阐述系统聚类法的基本思想

26、论述贝叶斯判别法的基本思想

27、简述主成分分析与因子分析的异同点

28、(1) 论述因子分析的基本思想  
(2) 简述因子分析的目的