**云南大学数学与统计学院**

**实验报告**

**实验课名称： 应用多元统计分析实验**

**指导教师： 李会琼**

**专业（年级）： 统计学2021级**

**学生姓名： 枫叶 学号:**

**实验名称： 实验8**

**实验成绩：**

**《应用多元统计分析实验》实验报告 8**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验名称 |  | | 实验成绩 |  | |
| 学号 |  | | 姓名 | 枫叶 | |
| 实验时间 | 2024年5月21日 | 实验地点 | 格物楼3508 | 指导教师 | 李会琼 |
| 1. **实验目的**   学习使用R软件进行距离判别   1. **实验要求**   1. 对所使用的方法与所得到的结果进行适当的文字描述。  2. 在实验结果的相应部分附上完整的代码与适当的注释。  3. 采用一定的可视化方法体现出对应计算结果。   1. **实验内容**   习题5-10 习题5-11   1. **实验软件**   R语言   1. **实验结果**  加载包 library(readxl) library(MASS) library(dplyr) #加载相关函数 source("正态总体假设检验函数汇集.R") source("距离判别函数汇集.R") 第一题 data <- read\_xlsx("D:/预删除文件夹/大三下/多元统计/多元统计习题5-10数据.xlsx") %>%  dplyr::select(-1) %>%  mutate(across(5,as.factor)) data\_new <- read\_xlsx("D:/预删除文件夹/大三下/多元统计/多元统计习题5-10数据.xlsx",sheet = 2) %>%  dplyr::select(-1) #Fisher model <- lda(类别号~.,data,prior=rep(1,3)/3) table(predict(model)$class==data$类别号)  ##  ## FALSE TRUE  ## 2 15  predict(model,data\_new)  ## $class ## [1] 1 2 3 ## Levels: 1 2 3 ##  ## $posterior ## 1 2 3 ## 1 9.033438e-01 0.09184014 0.004816059 ## 2 1.570495e-03 0.53009994 0.468329564 ## 3 4.316357e-05 0.14198518 0.857971660 ##  ## $x ## LD1 LD2 ## 1 1.684228 -0.3136291 ## 2 -1.401069 2.1863343 ## 3 -2.531001 -0.2899144  基于Fisher判别的回判结果正确率为88.2%，新样本分别被判为1,2,3 第二题 data <- read\_xlsx("D:/预删除文件夹/大三下/多元统计/多元统计习题5-11数据.xlsx") %>%  dplyr::select(-1) %>%  mutate(across(4,as.factor)) data\_new <- read\_xlsx("D:/预删除文件夹/大三下/多元统计/多元统计习题5-11数据.xlsx",sheet = 2) %>%  dplyr::select(-1) model <- lda(类别~.,data,prior=rep(1,3)/3) table(predict(model)$class == data$类别)  ##  ## TRUE  ## 14  predict(model,data\_new)  ## $class ## [1] 2 3 ## Levels: 1 2 3 ##  ## $posterior ## 1 2 3 ## 1 4.383828e-08 9.999889e-01 1.100647e-05 ## 2 4.709248e-24 1.390760e-08 1.000000e+00 ##  ## $x ## LD1 LD2 ## 1 0.1553246 -0.2377262 ## 2 4.3475918 2.7264613  基于Fisher判别的回判结果正确率为100%，对新样本分别判为2,3 | | | | | |