**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7**

**实验成绩：**

**《应用多元统计分析实验》实验报告 7**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验名称 |  | | 实验成绩 |  | |
| 学号 |  | | 姓名 |  | |
| 实验时间 | 2024年5月14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 （2） 习题5-11 191页 例5.2.2   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(-c(1,3)) %>%  `colnames<-`(c("X1","X3","X4","group")) %>%  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5-10数据.xlsx",sheet = 2) %>%  dplyr::select(-c(1,3)) %>%  `colnames<-`(c("X1","X3","X4")) multi.cov.test(data[,1:4],data$group,3)  ## $p.value ## [1] 0.01698739 ##  ## $correct.M ## [1] 24.57182  #Bayes table(distinguish.bayes(TrnX = data[,1:3],TrnG = data$group,var.equal = F)==data$group)  ##  ## FALSE TRUE  ## 3 14  distinguish.bayes(TrnX = data[,1:3],TrnG = data$group,TstX = data\_new,var.equal = F)  ## 1 2 3 ## blong 1 2 3  与实验6的做法相同，删去指标2再进行判别，对协方差阵的检验结果为拒绝原假设，即认为各协方差阵不全相等，广义平方距离判别的回判结果正确率为82.3%，对新样本的判别结果为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) #由于是正态总体，协方差阵相同，直接调用实验手册中的Bayes判别函数即可 table(distinguish.bayes(data[,1:3],data$类别,var.equal = T) == data$类别)  ##  ## TRUE  ## 14  distinguish.bayes(data[,1:3],data$类别,var.equal = T,TstX = data\_new)  ## 1 2 ## blong 2 3  基于广义平方距离的回判结果为全部判断正确，对新样本的判断结果为2,3 第三题 data <- read\_xlsx("D:/预删除文件夹/大三下/多元统计/多元统计例5.2.2数据.xlsx") %>%  mutate(across(1,factor,labels=c(3,2,1)))  ## Warning: There was 1 warning in `mutate()`. ## ℹ In argument: `across(1, factor, labels = c(3, 2, 1))`. ## Caused by warning: ## ! The `...` argument of `across()` is deprecated as of dplyr 1.1.0. ## Supply arguments directly to `.fns` through an anonymous function instead. ##  ## # Previously ## across(a:b, mean, na.rm = TRUE) ##  ## # Now ## across(a:b, \(x) mean(x, na.rm = TRUE))  multi.cov.test(data[,2:5],ind = data$分类,3)  ## $p.value ## [1] 0.1190175 ##  ## $correct.M ## [1] 17.89174  先做协方差阵全相等的检验，在0.05的显著性水平下不能拒绝原假设，即认为协方差阵相等，教材中直接假定协方差阵不等可能出于经验证据，这里基于假设检验结果认为协方差阵相等，直接调用Bayes判别函数  table(distinguish.bayes(data[,2:5],data$分类,var.equal = T) == data$分类)  ##  ## FALSE TRUE  ## 4 11  正确率为73.33%，下面再基于协方差阵不等的假定作检验  table(distinguish.bayes(data[,2:5],data$分类) == data$分类)  ##  ## FALSE TRUE  ## 4 11  检验结果同样只有73.33%的正确率，暂不清楚为什么与教材结果不一致 | | | | | |