**MANOVA-Exercise**

Jitesh Rawat

3/23/2021

**library**(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3 v purrr 0.3.4  
## v tibble 3.0.4 v dplyr 1.0.2  
## v tidyr 1.1.2 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.0

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

**library**(palmerpenguins)  
data <- penguins  
**head**(data)

## # A tibble: 6 x 8  
## species island bill\_length\_mm bill\_depth\_mm flipper\_length\_~ body\_mass\_g sex   
## <fct> <fct> <dbl> <dbl> <int> <int> <fct>  
## 1 Adelie Torge~ 39.1 18.7 181 3750 male   
## 2 Adelie Torge~ 39.5 17.4 186 3800 fema~  
## 3 Adelie Torge~ 40.3 18 195 3250 fema~  
## 4 Adelie Torge~ NA NA NA NA <NA>   
## 5 Adelie Torge~ 36.7 19.3 193 3450 fema~  
## 6 Adelie Torge~ 39.3 20.6 190 3650 male   
## # ... with 1 more variable: year <int>

Y <- data **%>%** **select**(**contains**("length"))  
Y <- **as.matrix**(Y)

G <- data **%>%** **select**(species)  
G <- **as.vector**(G)

*# MANOVA test*   
model1 = **manova**(**cbind**(bill\_length\_mm, flipper\_length\_mm) **~** species,data = data)  
**summary**(model1)

## Df Pillai approx F num Df den Df Pr(>F)   
## species 2 1.3805 377.72 4 678 < 2.2e-16 \*\*\*  
## Residuals 339   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

**shapiro.test**(model1**$**residuals)

##   
## Shapiro-Wilk normality test  
##   
## data: model1$residuals  
## W = 0.98112, p-value = 1.016e-07

*#From the below statistics we can make out that the model is not normal*

**library**(heplots)

## Warning: package 'heplots' was built under R version 4.0.4

## Loading required package: car

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

## The following object is masked from 'package:purrr':  
##   
## some

res <- **boxM**(data[,**c**(3,4,5,6)], data**$**species)

## Warning in boxM.default(data[, c(3, 4, 5, 6)], data$species): 2 cases with  
## missing data have been removed.

res

##   
## Box's M-test for Homogeneity of Covariance Matrices  
##   
## data: data[, c(3, 4, 5, 6)]  
## Chi-Sq (approx.) = 76.795, df = 20, p-value = 1.365e-08

**plot**(res, gplabel="Species")



*#We can see that p-value is less than 0.05 which means the data does not follow the homogenity of variance*  
*#Even from the plot we can observe the difference in different species*