Lab 8.R

## ANOVA

Kevin George-2448030

INTRODUCTION:

This report analyzes two distinct datasets using inferential statistical methods in R. The first dataset evaluates the consistency and performance of three analytical methods (M1, M2, M3) used by different analysts in determining the amount of a certain constituent in a sample. The second dataset explores the effect of noise intensity and gender on the academic performance of students, aiming to identify the main and interaction effects of these factors. These analyses are conducted using two-way ANOVA and post-hoc tests to derive meaningful conclusions. It emphasizes the role of ANOVA in determining significant differences among groups, as well as interactions between multiple factors.

OBJECTIVE:

Objective 1:

* To evaluate whether there is significant variation in results across the three analytical methods (M1, M2, M3).
* To determine if there is significant variation between analysts in their measurements.
* To identify which pairs of methods significantly differ, using post-hoc analysis.

Objective 2:

* To find the effect of noise intensity on student marks.
* To find the impact of gender on academic performance.
* To analyze the interaction between gender and noise intensity in affecting student marks.

*#Objective 1*  
data1=**data.frame**(Analyst = **rep**(1**:**5, 3),Method = **rep**(**c**("M1", "M2", "M3"), each = 5),Value = **c**(7.5, 7.4, 7.3, 7.6, 7.4, 7.0, 7.2, 7.0, 7.2, 7.1, 7.1, 6.7, 6.9, 6.8, 6.9));data1

## Analyst Method Value  
## 1 1 M1 7.5  
## 2 2 M1 7.4  
## 3 3 M1 7.3  
## 4 4 M1 7.6  
## 5 5 M1 7.4  
## 6 1 M2 7.0  
## 7 2 M2 7.2  
## 8 3 M2 7.0  
## 9 4 M2 7.2  
## 10 5 M2 7.1  
## 11 1 M3 7.1  
## 12 2 M3 6.7  
## 13 3 M3 6.9  
## 14 4 M3 6.8  
## 15 5 M3 6.9

*# Two-way ANOVA*  
anova1=**aov**(Value **~** Method **+** Analyst, data = data1)  
**summary**(anova1)

## Df Sum Sq Mean Sq F value Pr(>F)   
## Method 2 0.7960 0.3980 24.37 9.09e-05 \*\*\*  
## Analyst 1 0.0003 0.0003 0.02 0.889   
## Residuals 11 0.1797 0.0163   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

*# Tukey's HSD test*  
tukey1=**TukeyHSD**(anova1, "Method");tukey1

## Warning in replications(paste("~", xx), data = mf): non-factors ignored:  
## Analyst

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = Value ~ Method + Analyst, data = data1)  
##   
## $Method  
## diff lwr upr p adj  
## M2-M1 -0.34 -0.5583077 -0.121692347 0.0038476  
## M3-M1 -0.56 -0.7783077 -0.341692347 0.0000680  
## M3-M2 -0.22 -0.4383077 -0.001692347 0.0482493

INFERENCE:

(a)

H01: There is no significant difference in the results between the three methods M1,M2,M3

H11: There is significant difference in the results between atleast two methods.

H02: There is no significant difference in the results between the analysts.

H12: There is significant difference in the results between atleast two analysts.

For between methods, the F-value is 24.37 and p-value is less than 0.001. Hence we reject H01, therefore there is significant variation between the methods M1,M2,M3.

For between analysts, the F-value is 0.02 and p-value is 0.889. Hence we do not reject H02 (0.889>0.05) therefore there is no significant variation between analysts.

(b)

H01: Mean values of M1 and M2 are equal.

H11: Mean values of M1 and M2 are not equal.

H02: Mean values of M1 and M3 are equal.

H12: Mean values of M1 and M3 are not equal.

H03: Mean values of M2 and M3 are equal.

H13: Mean values of M2 and M3 are not equal.

M2-M1: Difference= -0.34, p=0.0038. We reject H01 as p-value is less than 0.05 hence mean values of M1 and M2 differ significantly.

M3-M1: Difference= -0.56, p<0.001. We reject H02 as p-value is less than 0.05 hence mean values of M1 and M3 differ significantly.

M3-M2: Difference= -0.22, p=0.0482. We reject H01 as p-value is less than 0.05 hence mean values of M2 and M3 differ significantly.

*#Objective 2*  
data2=**data.frame**(  
Gender = **rep**(**c**("Male", "Female"), each = 12),  
Noise = **rep**(**c**("Low", "Medium", "Loud"),each= 4),  
Marks = **c**(10, 12, 11, 9, 7, 9, 8, 12, 4, 5, 6, 5,12, 13, 10, 13, 13, 15, 12, 12, 6, 6, 4, 4));data2

## Gender Noise Marks  
## 1 Male Low 10  
## 2 Male Low 12  
## 3 Male Low 11  
## 4 Male Low 9  
## 5 Male Medium 7  
## 6 Male Medium 9  
## 7 Male Medium 8  
## 8 Male Medium 12  
## 9 Male Loud 4  
## 10 Male Loud 5  
## 11 Male Loud 6  
## 12 Male Loud 5  
## 13 Female Low 12  
## 14 Female Low 13  
## 15 Female Low 10  
## 16 Female Low 13  
## 17 Female Medium 13  
## 18 Female Medium 15  
## 19 Female Medium 12  
## 20 Female Medium 12  
## 21 Female Loud 6  
## 22 Female Loud 6  
## 23 Female Loud 4  
## 24 Female Loud 4

*# Two-way ANOVA*  
anova2 <- **aov**(Marks **~** Gender **+** Noise **+** Gender**:**Noise, data = data2)  
**summary**(anova2)

## Df Sum Sq Mean Sq F value Pr(>F)   
## Gender 1 20.17 20.17 9.811 0.00576 \*\*   
## Noise 2 200.33 100.17 48.730 5.44e-08 \*\*\*  
## Gender:Noise 2 16.33 8.17 3.973 0.03722 \*   
## Residuals 18 37.00 2.06   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

INFERENCE:

H01: Noise intensity does not significantly affect the marks scored by students.

H11: Noise intensity significantly affect the marks scored by students.

H02: Gender does not significantly affect the marks scored by students.

H12: Gender significantly affect the marks scored by students.

H03**:​** There is no significant interaction between gender and noise intensity in affecting marks.

H13: There is a significant interaction between gender and noise intensity in affecting marks.

(a)

The F-value=48.73 and p-value<0.001. We reject H01 since p-value is less than 0.05 hence noise intensity significantly affects the marks scored by students.

(b)

The F-value=9.81 and p-value=0.0058. We reject H02 since p-value is less than 0.05 hence gender significantly affect the marks scored by students.

(c)

The F-value =3.97 and p-value=0.0372. We reject H03 since p-value is less than 0.05 hence there is a significant interaction between gender and noise intensity.

Interaction between gender and noise affects scores of students.