Two-Way ANOVA with Interaction

Srikar

2021-08-25

Objective:

To check whether the different genders feel the same relief of pain and to check whether each drug gives same amount of relief from pain

Data Description:

The dataset contains mainly 4 variables, they are Gender, Drug type Interaction and Pain Index. Gender, Interaction and Drug type are independant variables (X) # and dependent variable is the Pain Index (Y). The two genders being Male and Female. The different type of drugs are A, B and C which are pain relieving drugs. The pain index is on a scale from 1-20 with 1 being the least amount of pain and 20 being the most amount of pain.The interaction category contains the interaction between each gender with respect to the 3 drug types,

Data Summary

```
library(readxl)
dat <- read_excel("C:/Users/Srikar/Desktop/SS/R/Sem 5/Design of Exp/Practical 4/data.xlsx")
```

```
## Registered S3 methods overwritten by 'tibble':
##   method     from
##   format.tbl pillar
##   print.tbl  pillar
```

```
head(dat,10)
```

```
## Warning: `...` is not empty.
##
## We detected these problematic arguments:
## * `needs_dots`
##
## These dots only exist to allow future extensions and should be empty.
## Did you misspecify an argument?
```

```
## # A tibble: 10 x 4
##    Gender Drug  Interaction `Pain Index`
```

```
##    <chr>  <chr> <chr>              <dbl>
##  1 male   C     MC                 12.4
##  2 female A     FA                  7.69
##  3 male   C     MC                 14
##  4 female A     FA                  9.69
##  5 male   C     MC                 11.6
##  6 female A     FA                  8.89
##  7 female A     FA                  6.94
##  8 female A     FA                  2.13
##  9 female A     FA                  7.26
## 10 female A     FA                  5.87
```

```
str(dat)
```

```
## tibble [33 x 4] (S3: tbl_df/tbl/data.frame)
##  $ Gender     : chr [1:33] "male" "female" "male" "female" ...
##  $ Drug       : chr [1:33] "C" "A" "C" "A" ...
##  $ Interaction: chr [1:33] "MC" "FA" "MC" "FA" ...
##  $ Pain Index : num [1:33] 12.4 7.69 14 9.69 11.6 8.89 6.94 2.13 7.26 5.87
...
```

Gender, Drug and Interaction are character types. Pain Index is of numeric type

```
dim(dat)
```

```
## [1] 33  4
```

There are 33 observations and 4 variables

Important terminologies:

1) Two-way Analysis- A two-way ANOVA test is a statistical test used to determine the effect of two nominal predictor variables on a continuous outcome variable.
A two-way ANOVA tests the effect of two independent variables on a dependent variable

2) Interaction-Interaction effects represent the combined effects of factors on the dependent measure. When an interaction effect is present,
the impact of one factor depends on the level of the other factor

Hypothesis Testing

1)Hypothesis with respect to gender

Null Hypothesis (Ho) : μ 1 =μ 2 (Each Gender has eqaul pain index)
Alternative Hypothesis (H1): μ 1 =/ μ 2 (Genders have different pain index)

2) Hypothesis with respect to drug

Null Hypothesis (Ho): μ 1 =μ 2 =μ 3 (Each drug gives the same amount of pain relief)
Alternative Hypothesis (H1): (Each drug gives different amount of pain relief)

3)Hypothesis with respect to interaction between factors

Null Hypothesis (Ho): There is no interaction between gender and drug type
Alternate Hypothesis (H1): There is interaction between gender and drug type

Procedure

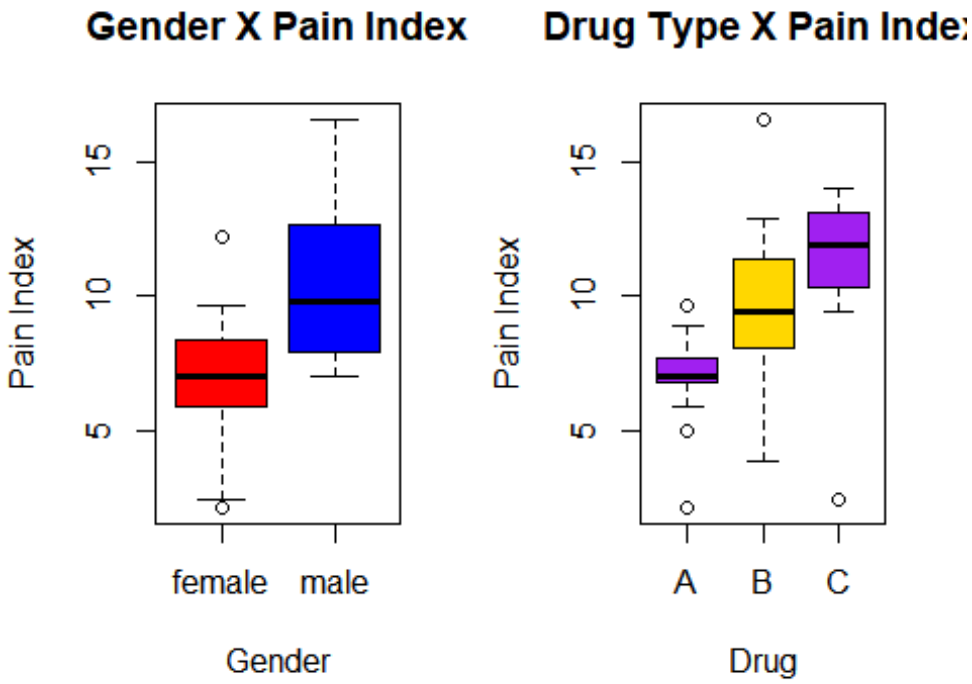
1) Converting the factor variables as factors

```
gen=as.factor(dat$Gender)
drg=as.factor(dat$Drug)
int=as.factor(dat$Interaction)
```

#The function factor has helped us encode the vector as a factor for performing the two-way analysis

#2) Visualizing the data

```
par(mfrow=c(1,2))
A.box=boxplot(dat$`Pain Index`~dat$Gender,xlab = 'Gender',ylab = 'Pain
Index',col=c('red','blue'),main="Gender X Pain Index")
B.box=boxplot(dat$`Pain Index`~dat$Drug,xlab = 'Drug',ylab = 'Pain
Index',col=c('purple','gold'),main="Drug Type X Pain Index")
```




```
Table.A <- A.box$stats
colnames(Table.A)<-A.box$names
rownames(Table.A)<-c('min','lower quartile','median','upper quartile','max')
Table.A
```

```
##                female  male
## min             2.400  7.00
## lower quartile  5.870  7.92
## median          7.025  9.81
## upper quartile  8.350 12.65
## max             9.690 16.60
```

We see that the minimum level of pain level of female and male are 2.4 and 7 respectively.The median pain levels of female and male are 7.025 and 9.81. The maximum pain levels of female and male are 9.69 and 16.60. We observe 2 outliers in female

```
Table.B <- B.box$stats
colnames(Table.B)<-B.box$names
rownames(Table.B)<-c('min','lower quartile','median','upper quartile','max')
Table.B
```

```
##                    A      B      C
## min            5.870  3.840  9.410
## lower quartile 6.810  8.095 10.305
## median         7.035  9.420 11.900
```

```
## upper quartile 7.690 11.355 13.150
## max            8.890 12.900 14.000
```

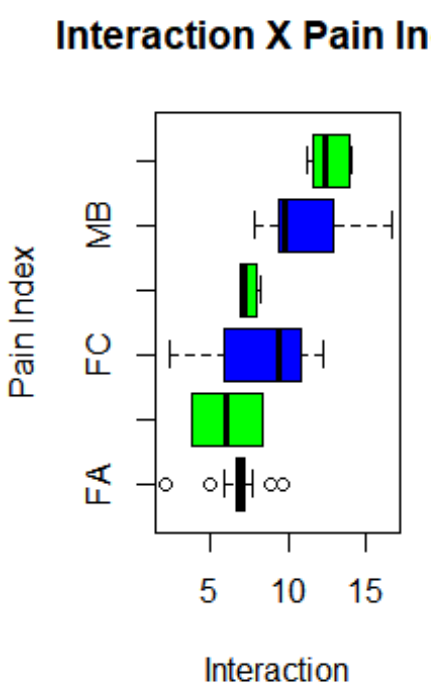
The minimum pain levels relieved by Drug A,B and C are 5.87,3.84 and 9.410. The median levels of each drug are 7.035 , 9.420, 11.900
The maximum pain levels of each drug are 8.890, 12.900, 14.000.We observe 2 outliers in Drug A. One outlier above the maximum value in Drug B and one outlier below the the minimum value of Drug C

```
C.box=boxplot(dat$`Pain Index`~dat$Interaction,xlab = 'Interaction',ylab =
'Pain Index',col=c('blue','green'),main="Interaction X Pain
Index",horizontal=TRUE)

Table.C <- C.box$stats
colnames(Table.C)<-C.box$names
rownames(Table.C)<-c('min','lower quartile','median','upper quartile','max')
Table.C
```

```
##                   FA    FB     FC   MA    MB   MC
## min             5.87 3.840  2.400 7.00  7.84 11.2
## lower quartile 6.67 3.840  5.905 7.00  9.42 11.6
## median          6.98 6.095  9.410 7.24  9.81 12.4
## upper quartile 7.26 8.350 10.805 8.00 12.90 13.9
## max             7.69 8.350 12.200 8.18 16.60 14.0
```

The table provides the minimum, first quarter, median, third quarter and maximum values of the interaction which can be graphically represented for better understanding.



Since the outliers are naturally part of the population and are not any sampled values, we will consider the outliers for both the factors.

3) Creating the ANOVA model

```
model=aov(dat$`Pain Index`~dat$Gender+dat$Drug+dat$Interaction)
summary(model)
```

```
##                 Df Sum Sq Mean Sq F value   Pr(>F)
## dat$Gender       1 103.78  103.78  17.803 0.000247 ***
## dat$Drug         2  45.97   22.98   3.943 0.031457 *
## dat$Interaction  2  30.41   15.20   2.608 0.092143 .
## Residuals       27 157.40    5.83
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As The p-value of of gender is below the significance level of 5% (0.000397),we will reject the Null-Hypothesis. Similarly the p-value of drug (0.041804) is below the 5% significance value, we will reject the Null-Hypothesis here as well. Since we rejected the null hypothesis in both the cases, we wil say that the factors are significant from each other. Since the p-value of Interaction is above 5% (0.05), we say that there is no interaction between gender and drug.

4) Considering the factor, ‘gender' and 'drug' as they show significance

```r
library(lsmeans)
```

```
## Warning: package 'lsmeans' was built under R version 3.6.3
```

```
## Loading required package: emmeans
```

```
## The 'lsmeans' package is now basically a front end for 'emmeans'.
## Users are encouraged to switch the rest of the way.
## See help('transition') for more information, including how to
## convert old 'lsmeans' objects and scripts to work with 'emmeans'.
```

```r
PI=dat$`Pain Index`

lm1=lm(PI~dat$Gender+dat$Drug+dat$Interaction)
lsm1=lsmeans(lm1,"Gender")
```

```
## NOTE: A nesting structure was detected in the fitted model:
##     Interaction %in% (Gender*Drug)
```

```r
lsm1
```

```
##  Gender lsmean    SE df lower.CL upper.CL
##  female   6.96 0.768 27     5.39     8.54
##  male    10.47 0.623 27     9.19    11.75
##
## Results are averaged over the levels of: Drug
## Confidence level used: 0.95
```

#This shows the average predictions of the pain levels along with their confidence level.
#Female pain levels has a 95 % confidence level of (5.39, 8.54)
#Male pain levels has a 95 % confidence level of (9.19,11.75)

```r
pairs(lsm1)
```

```
##  contrast      estimate    SE df t.ratio p.value
##  female - male    -3.51 0.989 27  -3.551  0.0014
##
## Results are averaged over the levels of: Drug
```

We find an average pain level difference of 3.51 between female and male.

```r
lsm2=lsmeans(lm1,"Drug")
```

```
## NOTE: A nesting structure was detected in the fitted model:
##     Interaction %in% (Gender*Drug)
```

```r
lsm2
```

```
##  Drug lsmean    SE df lower.CL upper.CL
##  A      7.13 0.635 27     5.83     8.44
##  B      8.70 1.010 27     6.63    10.78
##  C     10.31 0.882 27     8.50    12.12
##
## Results are averaged over the levels of: Gender
## Confidence level used: 0.95
```

This shows the average predictions of the pain levels along with their confidence level.
#Drug A pain levels has a 95 % confidence interval of (5.39, 8.54)
#Drug B pain levels has a 95 % confidence interval of (6.63,10.78)
#Drug C pain levels has a 95 % confidence interval of (8.50,12.12)

```
pairs(lsm2)
```

```
##  contrast estimate   SE df t.ratio p.value
##  A - B       -1.57 1.19 27  -1.316  0.3988
##  A - C       -3.18 1.09 27  -2.924  0.0184
##  B - C       -1.61 1.34 27  -1.199  0.4641
##
## Results are averaged over the levels of: Gender
## P value adjustment: tukey method for comparing a family of 3 estimates
```

We observe that only A and C are statistically significant as their p-value is lower than the significance value (0.05),Drug A and Drug B are not significant with respect to their mean pain index levels and Drug B and Drug C are not significant either with respect to thier mean levels of pain index.

Analysis

1) We observe from the ANOVA table constructed above that, the pain levels of male and female are different. It can be observed that on an average female have lower pain levels or high pain tolerance levels. Men have lower pain tolerance levels as their pain index is high. On an average, women have lower pain levels than men with an estimate of 3.51.

2) From the above ANOVA table, we can observe that the different pain levels of people after taking the different types of drugs are different. It can be observed that Drug A and Drug C are significantly different from each other. On an average, Drug A gives more relief from pain than drug C by providing lesser pain level of 3.19 than Drug C. Drug A and Drug B provide similar type of pain relief, Drug C and Drug B provide similar type of pain relief.

3) There is no interaction between the the two factors Drug and Gender.This means that by increasing or decreasing the levels of any of the factors will not have an impact on the other factor.

Conclusion

1) Female have more pain tolerance than males

2**) Drug A provides the most pain relief from the rest**