

Perform two-way Anova

Aniket Roy and Kruitk Shah

15/10/2020

Example

Moth dataset are stored in csv file and then reading csv file to get data from csv file.

```
data1 <- read.csv('moth.csv')
data1
```

```
##      X28      Top Scent
## 1     19      Top Scent
## 2     32      Top Scent
## 3     15      Top Scent
## 4     13      Top Scent
## 5     35      Top Sugar
## 6     22      Top Sugar
## 7     33      Top Sugar
## 8     21      Top Sugar
## 9     17      Top Sugar
## 10    31 Middle Scent
## 11    21 Middle Scent
## 12    14 Middle Scent
## 13    18 Middle Scent
## 14    24 Middle Scent
## 15    25 Middle Sugar
## 16    12 Middle Sugar
## 17    19 Middle Sugar
## 18    21 Middle Sugar
## 19    27 Middle Sugar
## 20    21 Middle Scent
## 21    34 Middle Scent
## 22    32 Middle Scent
## 23    11 Middle Scent
## 24    10 Bottom Scent
## 25    39 Bottom Sugar
## 26    37 Bottom Sugar
## 27    30 Bottom Sugar
## 28    22 Bottom Sugar
## 29    28 Bottom Sugar
```

```
data2<-data.frame("Moth"=data1[,1], "Position"=data1[,2], "lure"=data1[,3])
data2
```

```
##      Moth Position  lure
## 1      19      Top Scent
## 2      32      Top Scent
## 3      15      Top Scent
## 4      13      Top Scent
## 5      35      Top Sugar
## 6      22      Top Sugar
## 7      33      Top Sugar
## 8      21      Top Sugar
## 9      17      Top Sugar
## 10     31 Middle Scent
## 11     21 Middle Scent
## 12     14 Middle Scent
## 13     18 Middle Scent
## 14     24 Middle Scent
## 15     25 Middle Sugar
## 16     12 Middle Sugar
## 17     19 Middle Sugar
## 18     21 Middle Sugar
## 19     27 Middle Sugar
## 20     21 Middle Scent
## 21     34 Middle Scent
## 22     32 Middle Scent
## 23     11 Middle Scent
## 24     10 Bottom Scent
## 25     39 Bottom Sugar
## 26     37 Bottom Sugar
## 27     30 Bottom Sugar
## 28     22 Bottom Sugar
## 29     28 Bottom Sugar
```

The hypothesis testing problem is:

H0:There is no Significance difference between means of all

H1:There is a significant difference among means of all.

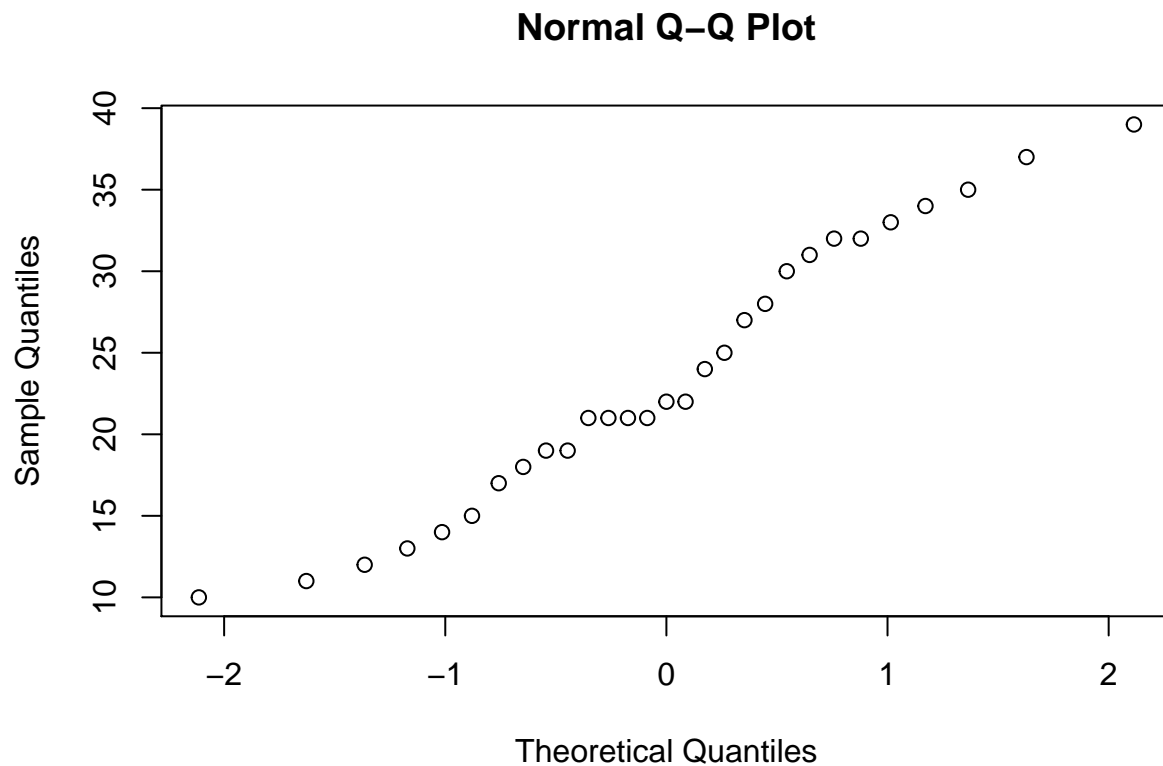
To check the normality of a data we have to do the shapiro test

```
res <- shapiro.test(data2$Moth)
res
```

```
##
##  Shapiro-Wilk normality test
##
## data:  data2$Moth
## W = 0.96219, p-value = 0.3717
```

Testing Normality Graphically

```
qqnorm(data2$Moth)
```



Bartlett's test is used to test if k samples have equal variances. Equal variances across samples is called homogeneity of variances.

```
bartlett.test(Moth~Position,data=data2)
```

```
##
## Bartlett test of homogeneity of variances
##
## data: Moth by Position
## Bartlett's K-squared = 1.1313, df = 2, p-value = 0.568
```

```
bartlett.test(Moth~lure,data=data2)
```

```
##
## Bartlett test of homogeneity of variances
##
## data: Moth by lure
## Bartlett's K-squared = 0.059131, df = 1, p-value = 0.8079
```

Here in moth by position p-value is 0.568 which is greater than 0.05 so we accept NULL Hypothesis. And in moth by lure p-value is 0.8079 which is greater than 0.05 so we accept NULL Hypothesis.

```
twoway <- aov(Moth~Position+lure,data=data2)
summary(twoway)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Position    2  132.1   66.06   0.974  0.391
## lure        1   91.5   91.50   1.349  0.256
## Residuals   25 1695.5   67.82
```

```
twoway
```

```
## Call:
## aov(formula = Moth ~ Position + lure, data = data2)
##
## Terms:
##           Position      lure Residuals
## Sum of Squares  132.1248   91.5042 1695.5435
## Deg. of Freedom      2        1      25
##
## Residual standard error: 8.235395
## Estimated effects may be unbalanced
```

```
TK <- TukeyHSD(twoway,"Position")
TK
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = Moth ~ Position + lure, data = data2)
##
## $Position
##           diff      lwr      upr      p adj
## Middle-Bottom -5.5238095 -15.533115  4.485495  0.3690257
## Top-Bottom    -4.6666667 -15.477952  6.144618  0.5378869
## Top-Middle     0.8571429  -7.906955  9.621241  0.9678568
```

```
plot(TK)
```

