Chi Square Test

Shashank Semwal

19 March 2018

A movie producer is bringing out a new movie. In order to map out a advertising campaign, he wants to determin wether the movie will appeal most to a particular age group or whether it will appeal equally to all age group. The producer takes a random sample from persons attending preview of the new movie and obtains the follwoing results.

```
movie.df<-data.frame('Under 20'=c(146,54,20),'20-39'=c(78,22,10),'40-59'=c(48,42,10),'60 and
 Over'=c(28,22,20))
row.names(movie.df)<-c('Liked the Movie','Disliked the Movie','Indifferent')
movie.df
```

```
##                    Under.20 X20.39 X40.59 X60.and.Over
## Liked the Movie         146     78     48           28
## Disliked the Movie       54     22     42           22
## Indifferent              20     10     10           20
```

Our Hypothesis

H0 : Movies will not appeal any age group

H1: Movies will appeal the age groups

For X2 (Chi-Square) at 0.05 significance level i.e. 3.84

```
# Chi square contigency Table
movies.chi<-chisq.test(as.matrix(movie.df))
```

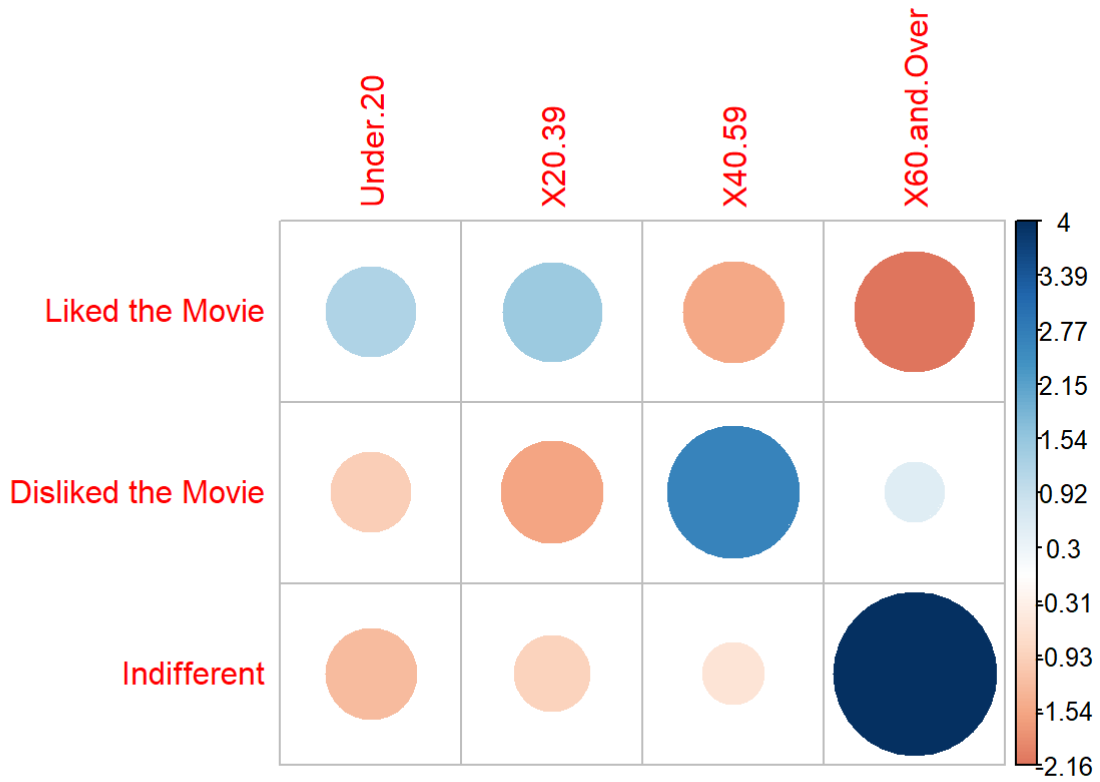
```
# Observed Values
movies.chi$observed
```

```
##                    Under.20 X20.39 X40.59 X60.and.Over
## Liked the Movie         146     78     48           28
## Disliked the Movie       54     22     42           22
## Indifferent              20     10     10           20
```

```
# Expected Values
movies.chi$expected
```

```
##                    Under.20 X20.39 X40.59 X60.and.Over
## Liked the Movie       132.0   66.0     60         42.0
## Disliked the Movie     61.6   30.8     28         19.6
## Indifferent            26.4   13.2     12          8.4
```

```
# Age most constributing To likes, dislike aned indifferent about movies
corrplot(movies.chi$residuals,is.corr = F)
```




```
# we can see from the contribution table that
# 1) People above age 60 are most indifferent about movies
# 2) People age b/w 40-50 mostly disliked the movie
# 3) Out of many people above age 60 dislked the movie
# 4) More people disliked the movie rather liked
```

```
# For n=(3-1)(4-1)=6 for alpha=0.05, Chi-sq = 12.59

movies.chi
```

```
##
##  Pearson's Chi-squared test
##
## data:  as.matrix(movie.df)
## X-squared = 40.159, df = 6, p-value = 4.239e-07
```

```
#we can see that p-value= 40.589, hence we reject the null hypothsis, Age group do affect the
 likes and dislikes about movies.
```