Data 606: Chapter-5 Homework

_____

Student Name : Sachid Deshmukh

Date : 11/16/2018

- GitHub Location for rmd file
- GitHub Location for pdf file
- RPubs location of published file

_____

1. 5.6 Calculate the sample mean, the margin of error, and the sample standard deviation.

```
N = 25
Margin.Error = (77-65)/2
Sample.Mean = 77 - Margin.Error
Crit.Value = qt(1-(0.05), (N-1))
Standard.Error = Margin.Error/Crit.Value
Sample.SD = Standard.Error * sqrt(N)
print(paste("Sample mean =", Sample.Mean))
```

```
## [1] "Sample mean = 71"
```

```
print(paste("Margin of error =", Margin.Error))
```

```
## [1] "Margin of error = 6"
```

```
print(paste("Sample standard deviation =", Sample.SD))
```

```
## [1] "Sample standard deviation = 17.5348145569379"
```

2. 5.48 (A) Write hypotheses for evaluating whether the average number of hours worked varies across the groups.

Answer :

Null Hypothesis : The means of each group are equal Alternate Hypothesis : The means of each group are not equal

3. 5.48 (B) Check conditions and describe any assumptions you must make to proceed with the test.

Answer : Sample sizes are fairly large. Data looks noramlly distributed. This is good for hypothesis testing. We need to assume independence across groups for achieving fair results in hypothesis testing

4. 5.48 (C) Below is part of the output associated with this test. Fill in the empty cells.

```
groups <- 5
samplesize <- 1172
meansq.group <- 501.54
sumsq.res <- 267382
alpha <- 0.0682
```

```
# Find Degree of Freedom
df.group <- groups-1
df.res <- samplesize - groups
df.total <- df.group + df.res

# Find Sum of Squared

sumsq.group <- df.group * meansq.group
sumsq.total <- sumsq.group + sumsq.res

# Find Mean Square
mean.square <- sumsq.res / df.res


# Find F-value
f.stat <- meansq.group / mean.square

print(paste("Degree of freedom for groups = ", df.group))
```

[1] "Degree of freedom for groups = 4"

```
print(paste("Degree of freedom for residual = ", df.res))
```

[1] "Degree of freedom for residual = 1167"

```
print(paste("Degree of freedom for total = ", df.total))
```

[1] "Degree of freedom for total = 1171"

```
print(paste("Sum of squared for groups = ", sumsq.group))
```

[1] "Sum of squared for groups = 2006.16"

```
print(paste("Sum of squared for total = ", sumsq.total))
```

[1] "Sum of squared for total = 269388.16"

```
print(paste("Mean Square residual = ", mean.square))
```

[1] "Mean Square residual = 229.11910882605"

```
print(paste("F Value = ", f.stat))
```

[1] "F Value = 2.18899245274551"

5.   5.48 (D) What is the conclusion of the test

Answer - With 95% confidence interval, level of significance, alpha=0.05. Since p value=0.0682 > alpha, we fail to reject null hypothesis. We need to accept null hypothesis and assume that mean for each groups are equal