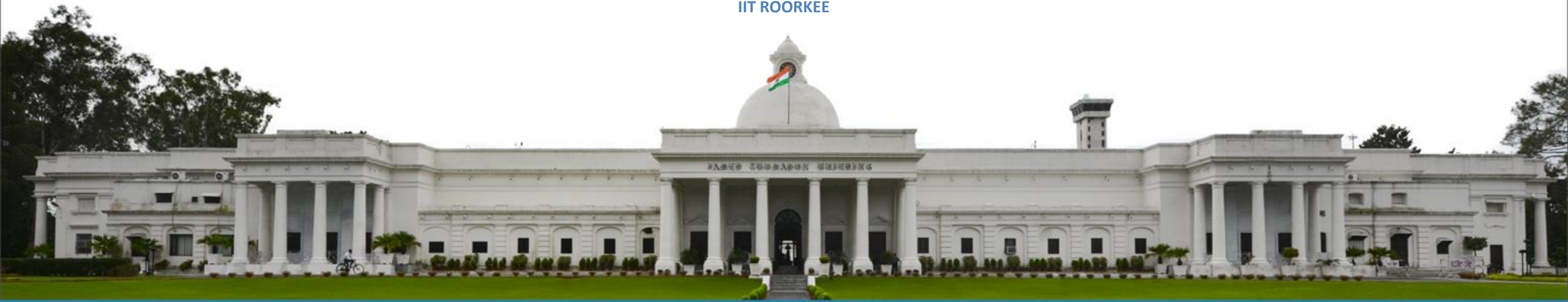
Post Hoc Analysis(Tukey’s test)

Dr. A. Ramesh

DEPARTMENT OF MANAGEMENT STUDIES

IIT ROORKEE



Lecture Objectives

After completing this lecture, you should be able to:

- Use Tukey's test and LSD Test to identify specific differences between means

Designing engineering experiments

- Experimental design methods are also useful in engineering design activities, where new products are developed and existing ones are improved
- By using designed experiments, engineers can determine which subset of the process variables has the greatest influence on process performance

Designing engineering experiments

- The results of an experiment can lead to
  1. Improved process yield
  2. Reduced variability in the process and closer conformance to nominal or target requirements
  3. Reduced design and development time
  4. Reduced cost of operation

Designing engineering experiments

- Every experiment involves a sequence of activities:
    1. **Conjecture**—the original hypothesis that motivates the experiment
    2. **Experiment**—the test performed to investigate the conjecture
    3. **Analysis**—the statistical analysis of the data from the experiment
    4. **Conclusion**—what has been learned about the original conjecture from the experiment. Often the experiment will lead to a revised conjecture, and a new experiment, and so forth

The completely randomized single-factor experiment example

- A manufacturer of paper that is used for making grocery bags is interested in improving the tensile strength of the product
- Product engineer thinks that tensile strength is a function of the hardwood concentration in the pulp and that the range of hardwood concentrations of practical interest is between 5 and 20%.

The completely randomized single-factor experiment example

- A team of engineers responsible for the study decides to investigate four levels of hardwood concentration: 5%, 10%, 15%, and 20%.
- They decide to make up six test specimens at each concentration level, using a pilot plant.
- All 24 specimens are tested on a laboratory tensile tester, in random order. The data from this experiment are shown in Table

The completely randomized single-factor experiment example

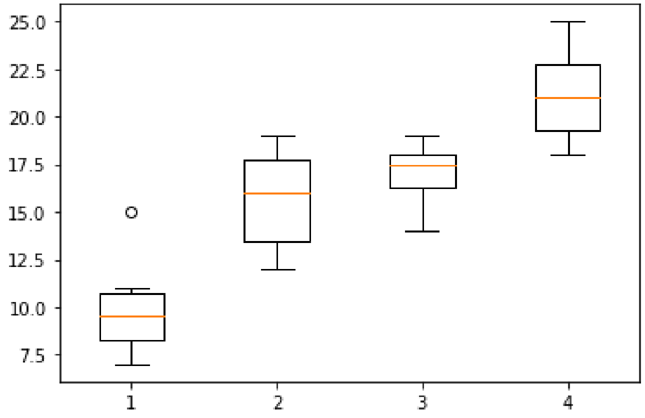
- Tensile Strength of Paper (psi)

Hardwood Concentration (%)	Observations						Total	Avg
	1	2	3	4	5	6		
5	7	8	15	11	9	10	60	10.00
10	12	17	13	18	19	15	94	15.67
15	14	18	19	17	16	18	102	17.00
20	19	25	22	23	18	20	127	21.17
							383	15.96

The completely randomized single-factor experiment example

In [3]:
```
fivepercent= [7,8,15,11,9,10]
tenpercent=[12,17,13,18,19,15]
fifteenpercent=[14,18,19,17,16,18]
twentypercent=[19,25,22,23,18,20]

box_plot_data=[fivepercent,tenpercent,fifteenpercent,twentypercent]
plt.boxplot(box_plot_data)
plt.show()
```



Typical Data for Single Factor Experiment

Treatment	Observations				Totals	Averages
1	y_{11}	y_{12}	…	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{23}	…	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
.	.	.	…	.	.	.
.	.	.	…	.	.	.
.	.	.	…	.	.	.
a	y_{a1}	y_{a2}	…	y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

Sum of Squares

$$\text{Total sum of squares} = SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2$$

$$\text{Treatment sum of squares} = SS_{Treatments} = n\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2$$

$$\text{Error sum of Squares} = SSE = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{j.})^2$$

ANOVA with Equal Sample Sizes

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^{2} - \frac{y^{2}_{..}}{N}$$

$$SSTreatments = \frac{1}{n}\sum_{i=1}^{a} y_{i.}^{2} - \frac{y^{2}_{..}}{N}$$

N = an = No. of Treatments x no. of sample size = Total no. of Sample Size

ANOVA with unequal Sample Sizes

$$SST = \sum_{i=1}^{a} \sum_{j=1}^{n} y_{ij}^{2} - \frac{y^{2}_{..}}{N}$$

$$SSTreatments = \sum_{i=1}^{a} \frac{y_{i.}^{2}}{n_i} - \frac{y^{2}_{..}}{N}$$

N = an = No. of Treatments x no. of sample size = Total no. of Sample Size

Problem: Analysis of variance

- Consider the paper tensile strength experiment described.
- We can use the analysis of variance to test the hypothesis that different hardwood concentrations do not affect the mean tensile strength of the paper.
- The hypotheses are
- H0: $\tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$
- H1: $\tau_i \neq 0$ for at least one i

Problem: Analysis of variance

- We will use $\alpha = 0.01$.
- The sums of squares for the analysis of variance are computed are as follows:

$$SS_T = \sum_{i=1}^{4} \sum_{j=1}^{6} y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$= (7)^2 + (8)^2 + \cdots + (20)^2 - \frac{(383)^2}{24} = 512.96$$

$$SS_{\text{Treatments}} = \sum_{i=1}^{4} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N}$$

$$= \frac{(60)^2 + (94)^2 + (102)^2 + (127)^2}{6} - \frac{(383)^2}{24} = 382.79$$

$$SS_E = SS_T - SS_{\text{Treatments}}$$

$$= 512.96 - 382.79 = 130.17$$

ANOVA Table

Sources of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Treatments	SS Treatments	a-1	MS Treatments	MS Treatments / MSE
Error	SSE	a(n-1)	MSE	
Total	SST	an-1		

Problem: Analysis of variance

- The ANOVA is summarized as follow

Source of Variation	Sum of Squares	Degrees of freedom	Mean Square	F0	P-value
Hardwood concentration	382.79	3	127.6	19.6	3.59 E-6
Error	130.17	20	6.51		
Total	512.96	23			

```
In [9]: from scipy import stats
        1-scipy.stats.f.cdf(19.6, 3, 20)

Out[9]: 3.599599239012541e-06
```

Problem: Analysis of variance

- Since $f_{0.01,3,20}$ = 4.94, we reject H_0 and conclude that hardwood concentration in the pulp significantly affects the mean strength of the paper

```
In [32]: scipy.stats.f.ppf(1-0.01, dfn=3, dfd=20)

Out[32]: 4.938193382310539
```

Problem: Analysis of variance

In [23]:
```
scipy.stats.f_oneway(fivepercent,tenpercent,fifteenpercent,twentypercent)
```

Out[23]:
```
F_onewayResult(statistic=19.605206999573184, pvalue=3.5925782584743027e-06)
```

Jupyter code

In [1]:
```
import pandas as pd
import numpy as np
import scipy
import statsmodels.api as sm
from statsmodels.formula.api import ols
```

In [3]:
```
df = pd.read_excel('concentration.xlsx')
df
```

Out[3]:

	concentration5	concentration10	concentration15	concentration20
0	7	12	14	19
1	8	17	18	25
2	15	13	19	22
3	11	18	17	23
4	9	19	16	18
5	10	15	18	20

Jupyter code

In [5]:
```
data_r1 = pd.melt(df.reset_index(), id_vars=['index'], value_vars=['concentration5','concentration10','concentration15','concentration20'])
data_r1.columns = ['index', 'treatments', 'value']
```

In [6]:
```
model = ols('value ~ C(treatments)', data=data_r1).fit()
```

In [7]:
```
model.summary()
```

Jupyter code

Out[7]:

OLS Regression Results

Dep. Variable:	value	**R-squared:**	0.746
Model:	OLS	**Adj. R-squared:**	0.708
Method:	Least Squares	**F-statistic:**	19.61
Date:	Tue, 27 Aug 2019	**Prob (F-statistic):**	3.59e-06
Time:	15:03:38	**Log-Likelihood:**	-54.344
No. Observations:	24	**AIC:**	116.7
Df Residuals:	20	**BIC:**	121.4
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	15.6667	1.041	15.042	0.000	13.494	17.839
C(treatments)[T.concentration15]	1.3333	1.473	0.905	0.376	-1.739	4.406
C(treatments)[T.concentration20]	5.5000	1.473	3.734	0.001	2.428	8.572
C(treatments)[T.concentration5]	-5.6667	1.473	-3.847	0.001	-8.739	-2.594

Omnibus:	0.929	**Durbin-Watson:**	2.181
Prob(Omnibus):	0.628	**Jarque-Bera (JB):**	0.861
Skew:	0.248	**Prob(JB):**	0.650
Kurtosis:	2.215	**Cond. No.**	4.79

Jupyter code

In [8]:
```
aov_table = sm.stats.anova_lm(model, typ=1)
aov_table
```

Out[8]:

	df	sum_sq	mean_sq	F	PR(>F)
C(treatments)	3.0	382.791667	127.597222	19.605207	0.000004
Residual	20.0	130.166667	6.508333	NaN	NaN

Multiple Comparisons Following the ANOVA

- When the null hypothesis is rejected in the ANOVA, we know that some of the treatment or factor level means are different
- ANOVA doesn’t identify which means are different
- Methods for investigating this issue are called multiple comparisons methods

Fisher's least significant difference (LSD) method

- The Fisher LSD method compares all pairs of means with the null hypotheses H0:$\mu_i = \mu_j$ (for all i ≠ j) using the t-statistic

$$t_0 = \frac{\overline{y}_{i^*} - \overline{y}_{j^*}}{\sqrt{\frac{2MS_E}{n}}}$$

Fisher's least significant difference (LSD) method

- Assuming a two-sided alternative hypothesis, the pair of means i and j would be declared significantly different if

$$\left|\overline{y}_{i^*} - \overline{y}_{j^*}\right| > LSD$$

where LSD, the least significant difference, is

$$LSD = t_{\alpha/2, a(n-1)} \sqrt{\frac{2MS_E}{n}}$$

Fisher's least significant difference (LSD) method

- If the sample sizes are different in each treatment, the LSD is defined as

$$LSD = t_{\alpha/2, N-a} \sqrt{MS_E \left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

Problem : LSD method

- We will apply the Fisher LSD method to the hardwood concentration experiment. There are a = 4 means, n = 6, MSE = 6.51, and $t_{0.025,20}$ = 2.086. The treatment means are

$$\begin{aligned} \bar{y}_{1\cdot} &= 10.00 \text{ psi} \\ \bar{y}_{2\cdot} &= 15.67 \text{ psi} \\ \bar{y}_{3\cdot} &= 17.00 \text{ psi} \\ \bar{y}_{4\cdot} &= 21.17 \text{ psi} \end{aligned}$$

Problem : LSD method

- The value of LSD is:

$$LSD = t_{0.025,20}\sqrt{\frac{2MS_E}{n}} = 2.086\sqrt{\frac{2(6.51)}{6}} = 3.07$$

- Therefore, any pair of treatment averages that differs by more than 3.07 implies that the corresponding pair of treatment means are different.

Jupyter code

In [8]:
```
aov_table = sm.stats.anova_lm(model, typ=1)
aov_table
```

Out[8]:

	df	sum_sq	mean_sq	F	PR(>F)
C(treatments)	3.0	382.791667	127.597222	19.605207	0.000004
Residual	20.0	130.166667	6.508333	NaN	NaN

LSD

In [39]:
```
import math
t = -1*scipy.stats.t.ppf(0.025,20)
n= 6
MSE = 6.508333
lsd = t*math.sqrt(2*MSE/n)
lsd
```

Out[39]: 3.072422588325206

Problem : LSD method

- The comparisons among the observed treatment averages are as follows:

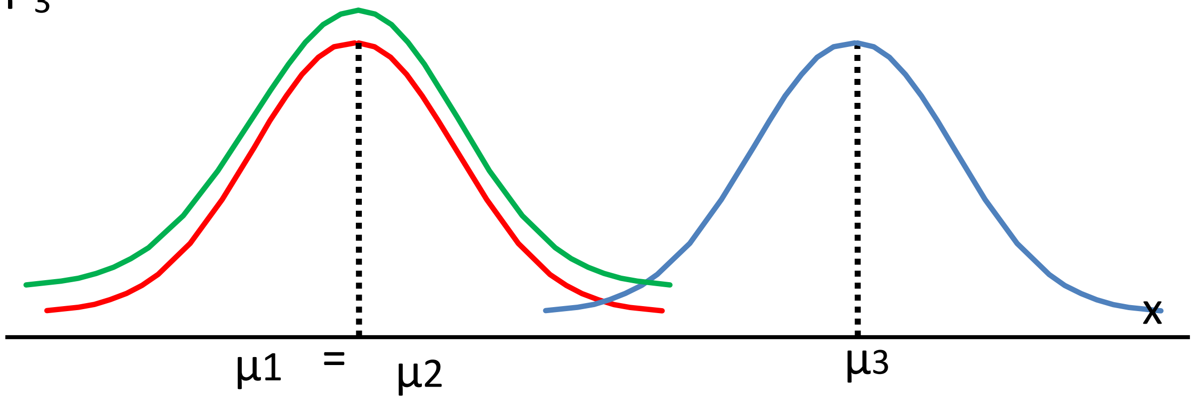
$$
\begin{aligned}
4 \text{ vs. } 1 &= 21.17 - 10.00 = 11.17 > 3.07 \\
4 \text{ vs. } 2 &= 21.17 - 15.67 = 5.50 > 3.07 \\
4 \text{ vs. } 3 &= 21.17 - 17.00 = 4.17 > 3.07 \\
3 \text{ vs. } 1 &= 17.00 - 10.00 = 7.00 > 3.07 \\
3 \text{ vs. } 2 &= 17.00 - 15.67 = 1.33 < 3.07 \\
2 \text{ vs. } 1 &= 15.67 - 10.00 = 5.67 > 3.07
\end{aligned}
$$

The Tukey-Kramer Test for Post Hoc analysis

- Tells which population means are significantly different
- Done after rejection of equal means in ANOVA
- Allows pair-wise comparisons
- Compare absolute mean differences with critical range

The Tukey-Kramer Test for Post Hoc analysis

- Determine is there any significant difference between the means
- is $\mu_1 = \mu_2 \neq \mu_3$



Tukey-Kramer Critical Range

$$\text{Critical Range} = Q_U \sqrt{\frac{\text{MSW}}{2}\left(\frac{1}{n_j} + \frac{1}{n_{j'}}\right)}$$

where:

Q_U = Value from Studentized Range Distribution with c and n - c degrees of freedom for the desired level of α

MSW = Mean Square Within

n_j and $n_{j'}$ = Sample sizes from groups j and j'

Problem: Tukey- Kramer test

- Tensile Strength of Paper (psi)

Hardwood Concentration (%)	Observations						Total	Avg
	1	2	3	4	5	6		
5	7	8	15	11	9	10	60	10.00
10	12	17	13	18	19	15	94	15.67
15	14	18	19	17	16	18	102	17.00
20	19	25	22	23	18	20	127	21.17
							383	15.96

The Tukey-Kramer Procedure

1. Compute absolute mean differences:

$$|\bar{x}_1 - \bar{x}_2| = |10.00 - 15.67| = 5.67$$

$$|\bar{x}_1 - \bar{x}_3| = |10.00 - 17.00| = 7$$

$$|\bar{x}_2 - \bar{x}_3| = |15.67 - 17.00| = 1.33$$

$$|\bar{x}_1 - \bar{x}_4| = |10.00 - 21.17| = 11.17$$

$$|\bar{x}_2 - \bar{x}_4| = |15.67 - 21.17| = 5.5$$

$$|\bar{x}_3 - \bar{x}_4| = |17.00 - 21.17| = 4.17$$

The Tukey-Kramer Procedure

2. Find the Q_U value from the table with $c = 4$ and $(n – c) = (24 – 4) = 20$ degrees of freedom for the desired level of α (α = .05 used here):

$$Q_U = 3.96$$

Error Term	2	3	4	5	6	7	3	9	10
5	3.64 5.70	4.60 6.98	5.22 7.80	5.67 8.42	6.03 8.91	6.33 9.32	6.58 9.67	6.80 9.97	6.99 10.24
6	3.46 5.24	4.34 6.33	4.90 7.03	5.30 7.56	5.63 7.97	5.90 8.32	6.12 8.61	6.32 8.87	6.49 9.10
7	3.34 4.95	4.16 5.92	4.68 6.54	5.06 7.01	5.36 7.37	5.61 7.68	5.82 7.94	6.00 8.17	6.16 8.37
8	3.26 4.75	4.04 5.64	4.53 6.20	4.89 6.62	5.17 6.96	5.40 7.24	5.60 7.47	5.77 7.68	5.92 7.86
9	3.20 4.60	3.95 5.43	4.41 5.96	4.76 6.35	5.02 6.66	5.24 6.91	5.43 7.13	5.59 7.33	5.74 7.49
10	3.15 4.48	3.88 5.27	4.33 5.77	4.65 6.14	4.91 6.43	5.12 6.67	5.30 6.87	5.46 7.05	5.60 7.21
11	3.11 4.39	3.82 5.15	4.26 5.62	4.57 5.97	4.82 6.25	5.03 6.48	5.20 6.67	5.35 6.84	5.49 6.99
12	3.08 4.32	3.77 5.05	4.20 5.50	4.51 5.84	4.75 6.10	4.95 6.32	5.12 6.51	5.27 6.67	5.39 6.81
13	3.06 4.26	3.73 4.96	4.15 5.40	4.45 5.73	4.69 5.98	4.88 6.19	5.05 6.37	5.19 6.53	5.32 6.67
14	3.03 4.21	3.70 4.89	4.11 5.32	4.41 5.63	4.64 5.88	4.83 6.08	4.99 6.26	5.13 6.41	5.25 6.54
15	3.01 4.17	3.67 4.84	4.08 5.25	4.37 5.56	4.59 5.80	4.78 5.99	4.94 6.16	5.08 6.31	5.20 6.44
16	3.00 4.13	3.65 4.79	4.05 5.19	4.33 5.49	4.56 5.72	4.74 5.92	4.90 6.08	5.03 6.22	5.15 6.35
17	2.98 4.10	3.63 4.74	4.02 5.14	4.30 5.43	4.52 5.66	4.70 5.85	4.86 6.01	4.99 6.15	5.11 6.27
18	2.97 4.07	3.61 4.70	4.00 5.09	4.28 5.38	4.49 5.60	4.67 5.79	4.82 5.94	4.96 6.08	5.07 6.20
19	2.96 4.05	3.59 4.67	3.98 [illegible]	4.25 5.33	4.47 5.55	4.65 5.73	4.79 5.89	4.92 6.02	5.04 6.14
20	2.95 4.02	3.58 4.64	3.96 [illegible]	4.23 5.29	4.45 5.51	4.62 5.69	4.77 5.84	4.90 5.97	5.01 6.09

Q table: The critical values for q corresponding to alpha = .05 (top) and alpha = .01 (bottom)

The Tukey-Kramer Procedure

Source of Variation	Sum of Squares	Degrees of freedom	Mean Square	F_0	P-value
Hardwood concentration	382.79	3	127.6	19.6	3.59 E-6
Error	130.17	20	6.51		
Total	512.96	23			

The Tukey-Kramer Procedure

3. Compute Critical Range:

$$\text{Critical Range} = Q_U \sqrt{\frac{MSW}{2}\left(\frac{1}{n_j} + \frac{1}{n_{j'}}\right)} = 3.96\sqrt{\frac{6.51}{2}\left(\frac{1}{6} + \frac{1}{6}\right)} = 4.124$$

4. Compare:

$$|\bar{x}_1 - \bar{x}_2| = |10.00 - 15.67| = 5.67$$

$$|\bar{x}_1 - \bar{x}_3| = |10.00 - 17.00| = 7$$

$$|\bar{x}_2 - \bar{x}_3| = |15.67 - 17.00| = 1.33$$

$$|\bar{x}_1 - \bar{x}_4| = |10.00 - 21.17| = 11.17$$

$$|\bar{x}_2 - \bar{x}_4| = |15.67 - 21.17| = 5.5$$

$$|\bar{x}_3 - \bar{x}_4| = |17.00 - 21.17| = 4.17$$

The Tukey-Kramer Procedure

5. Other then $\left|\bar{x}_2 - \bar{x}_3\right|$, all of the absolute **mean differences are greater than critical range.** Therefore there is significant difference between each pair of means, except 10% concentration and 15% concentration at the 5% level of significance.

Jupyter code

In [53]:
```
from statsmodels.stats.multicomp import pairwise_tukeyhsd
from statsmodels.stats.multicomp import MultiComparison
mc = MultiComparison(data_r1['value'], data_r1['treatments'])
mcresult = mc.tukeyhsd(0.05)
mcresult.summary()
```

Out[53]:

Multiple Comparison of Means - Tukey HSD,FWER=0.05

group1	group2	meandiff	lower	upper	reject
concentration10	concentration15	1.3333	-2.7894	5.4561	False
concentration10	concentration20	5.5	1.3773	9.6227	True
concentration10	concentration5	-5.6667	-9.7894	-1.5439	True
concentration15	concentration20	4.1667	0.0439	8.2894	True
concentration15	concentration5	-7.0	-11.1227	-2.8773	True
concentration20	concentration5	-11.1667	-15.2894	-7.0439	True

Problem 2

- Following table shows observed tensile strength (lb/in square) of different clothes having different weight percentage of cotton.
- Check whether having different weight percentage of cotton, plays any role in tensile strength (lb/in square) of clothes.

Problem 2

Weight Percentage of cotton	Observed tensile strength (lb/in square)					Total	Average
	1	2	3	4	5		
15	7	7	15	11	9	49	9.8
20	12	17	12	18	18	77	15.4
25	14	18	18	19	19	88	17.6
30	19	25	22	19	23	108	21.6
35	7	10	11	15	11	54	10.8
						Grand total=376	Grand mean= 15.004

- SSA = $5\,(9.8 - 15.04)^2 + 5\,(15.4 - 15.04)^2 + 5\,(17.6 - 15.04)^2 + 5(21.6 - 15.04)^2 + 5(10.8 - 15.04)^2 = 475.76$

SST = 636.96

SSE = 636.96 - 475.76=161.20

Sources of variation	Sum of squares	Degrees of freedom	Mean square	F-value
Cotton weight percentage	475.76	4	118.94	14.76
Error	161.20	20	8.06	
Total	639.96	24		

Problem 2

- When alpha =.05, $F_{(0.05,4,20)}$ =2.87
- Reject Ho

```
In [17]: scipy.stats.f.ppf(1-0.05, dfn=4, dfd=20)

Out[17]: 2.8660814020156584
```

Error Term	2	3	4	5	6	7	3	9	10
5	3.64 5.70	4.60 6.98	5.22 7.80	5.67 8.42	6.03 8.91	6.33 9.32	6.58 9.67	6.80 9.97	6.99 10.24
6	3.46 5.24	4.34 6.33	4.90 7.03	5.30 7.56	5.63 7.97	5.90 8.32	6.12 8.61	6.32 8.87	6.49 9.10
7	3.34 4.95	4.16 5.92	4.68 6.54	5.06 7.01	5.36 7.37	5.61 7.68	5.82 7.94	6.00 8.17	6.16 8.37
8	3.26 4.75	4.04 5.64	4.53 6.20	4.89 6.62	5.17 6.96	5.40 7.24	5.60 7.47	5.77 7.68	5.92 7.86
9	3.20 4.60	3.95 5.43	4.41 5.96	4.76 6.35	5.02 6.66	5.24 6.91	5.43 7.13	5.59 7.33	5.74 7.49
10	3.15 4.48	3.88 5.27	4.33 5.77	4.65 6.14	4.91 6.43	5.12 6.67	5.30 6.87	5.46 7.05	5.60 7.21
11	3.11 4.39	3.82 5.15	4.26 5.62	4.57 5.97	4.82 6.25	5.03 6.48	5.20 6.67	5.35 6.84	5.49 6.99
12	3.08 4.32	3.77 5.05	4.20 5.50	4.51 5.84	4.75 6.10	4.95 6.32	5.12 6.51	5.27 6.67	5.39 6.81
13	3.06 4.26	3.73 4.96	4.15 5.40	4.45 5.73	4.69 5.98	4.88 6.19	5.05 6.37	5.19 6.53	5.32 6.67
14	3.03 4.21	3.70 4.89	4.11 5.32	4.41 5.63	4.64 5.88	4.83 6.08	4.99 6.26	5.13 6.41	5.25 6.54
15	3.01 4.17	3.67 4.84	4.08 5.25	4.37 5.56	4.59 5.80	4.78 5.99	4.94 6.16	5.08 6.31	5.20 6.44
16	3.00 4.13	3.65 4.79	4.05 5.19	4.33 5.49	4.56 5.72	4.74 5.92	4.90 6.08	5.03 6.22	5.15 6.35
17	2.98 4.10	3.63 4.74	4.02 5.14	4.30 5.43	4.52 5.66	4.70 5.85	4.86 6.01	4.99 6.15	5.11 6.27
18	2.97 4.07	3.61 4.70	4.00 5.09	4.28 5.38	4.49 5.60	4.67 5.79	4.82 5.94	4.96 6.08	5.07 6.20
19	2.96 4.05	3.59 4.67	3.98 5.05	4.25 [illegible]	4.47 5.55	4.65 5.73	4.79 5.89	4.92 6.02	5.04 6.14
20	2.95 4.02	3.58 4.64	3.96 5.02	4.23 [illegible]	4.45 5.51	4.62 5.69	4.77 5.84	4.90 5.97	5.01 6.09

Q table: The critical values for q corresponding to alpha = .05 (top) and alpha = .01 (bottom)

Problem 2

$$T_\alpha = q_\alpha(c, n-c)\sqrt{\frac{MS_E}{n}}$$

$$\alpha = 0.05$$

$$q_{0.05}(5, 20) = 4.23$$

$$T_{0.05} = 4.23\sqrt{\frac{8.06}{5}} = 5.37$$

Problem 2

Any pair of treatment averages that differ in absolute value by more than 5.37 would imply that the corresponding pair of population means are significantly different.

Problem 2

$$\bar{y}_{1.} - \bar{y}_{2.} = |9.8 - 15.4| = 5.6^{*}$$

$$\bar{y}_{1.} - \bar{y}_{3.} = |9.8 - 17.6| = 7.8^{*}$$

$$\bar{y}_{1.} - \bar{y}_{4.} = |9.8 - 21.6| = 11.8^{*}$$

$$\bar{y}_{1.} - \bar{y}_{5.} = |9.8 - 10.8| = 1$$

Starred values indicate pairs of means that are significantly different.

$$\bar{y}_{2.} - \bar{y}_{3.} = |15.4 - 17.6| = 2.2$$

$$\bar{y}_{2.} - \bar{y}_{4.} = |15.4 - 21.6| = 6.2^{*}$$

$$\bar{y}_{2.} - \bar{y}_{5.} = |15.4 - 10.8| = 4.6$$

$$\bar{y}_{3.} - \bar{y}_{4.} = |17.6 - 21.6| = 4$$

$$\bar{y}_{3.} - \bar{y}_{5.} = |17.6 - 10.8| = 6.8^{*}$$

$$\bar{y}_{4.} - \bar{y}_{5.} = |21.6 - 10.8| = 10.8*$$

Jupyter code

```
In [2]: df3 = pd.read_excel('C:/Users/Somi/Documents/cotton weight.xlsx')
```

```
In [12]: data1 = pd.melt(df3.reset_index(), id_vars=['index'], value_vars=['cotwt.15','cotwt.20','cotwt.25','cotwt.30','cotwt.35'])
         data1.columns = ['id', 'treatments', 'value']
```

Jupyter Code

In [16]:
```
mc = MultiComparison(data1['value'], data1['treatments'])
mcresults = mc.tukeyhsd(0.05)
mcresults.summary()
```

Out[16]:

Multiple Comparison of Means - Tukey HSD,FWER=0.05

group1	group2	meandiff	lower	upper	reject
cotwt.15	cotwt.20	5.6	0.2266	10.9734	True
cotwt.15	cotwt.25	7.8	2.4266	13.1734	True
cotwt.15	cotwt.30	11.8	6.4266	17.1734	True
cotwt.15	cotwt.35	1.0	-4.3734	6.3734	False
cotwt.20	cotwt.25	2.2	-3.1734	7.5734	False
cotwt.20	cotwt.30	6.2	0.8266	11.5734	True
cotwt.20	cotwt.35	-4.6	-9.9734	0.7734	False
cotwt.25	cotwt.30	4.0	-1.3734	9.3734	False
cotwt.25	cotwt.35	-6.8	-12.1734	-1.4266	True
cotwt.30	cotwt.35	-10.8	-16.1734	-5.4266	True

Thank you