NNSE 784
Advanced Analytics Methods

Instructor: F Doyle (CESTM L210)

MW 4:30 – 5:50, NFN 203

Slide Set #13
Inferential Statistics:
Analysis of Variance (ANOVA)

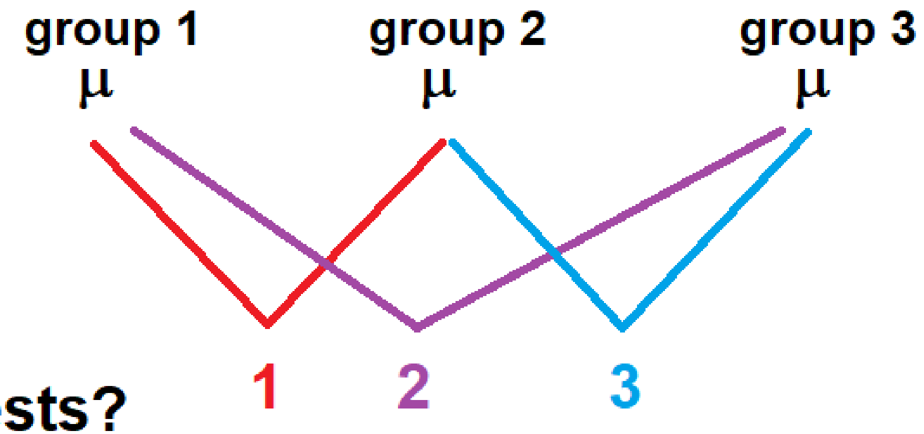
Lecture Outline

- Variance review
- One-way Analysis of variance
- Two-way Analysis of variance
- Multiple testing correction

Note: some material in these slides was adapted from “zedstatistics” https://www.youtube.com/watch?v=9cnSWads6oo

Analysis of Variance (ANOVA)

- Uses variances within groups compared to variances between groups to determine statistical differences
- Particularly useful for comparisons across more than two groups (populations)



Why not just do combinations of t-tests?

Remember Type I error and α? Each t-test contributes toward the possibility of a type I error (rejecting a correct null hypotheses). If each t-test has an alpha of .05, it is not as simple as .05 * 3, but it is close (about .143) which is a much higher probability of a Type I error than acceptable.

Remember the Definition of Sample Variance

A key concept in ANOVA is the “Sum of Squares” (SS). This basically equates to the numerator of the variance formula.

Sum of squares

$$s^2 = \frac{\Sigma\boxed{(X - \bar{X})^2}}{n - 1}$$

Sample variance

Total Sum of Squares (SST, SS_{Tot} , etc.)

$SST = \sum (X - \bar{X})^2$

Find SST for:

A: {2,2,3,5}

B: {4,10,13}

C: {4,8,12,16,20}

Total Sum of Squares (SST, SS_{Tot} , etc.)

$SST = \sum (X - \bar{X})^2$

Find SST for:

A: {2,2,3,5} $\bar{A} = 3$

$(2-3)^2 + (2-3)^2 + (3-3)^2 + (5-3)^2 = (-1)^2 + (-1)^2 + (0)^2 + (2)^2 = 6$

B: {4,10,13} $\bar{B} = 9$

$(4-9)^2 + (10-9)^2 + (13-9)^2 = (-5)^2 + (1)^2 + (4)^2 = 42$

C: {4,8,12,16,20}

Total Sum of Squares (SST, SS_{Tot} , etc.)

$$SST = \sum (X - \bar{X})^2$$

Find SST for:

A: {2,2,3,5} $\bar{A} = 3$

$(2-3)^2 + (2-3)^2 + (3-3)^2 + (5-3)^2 = (-1)^2 + (-1)^2 + (0)^2 + (2)^2 = 6$

B: {4,10,13} $\bar{B} = 9$

$(4-9)^2 + (10-9)^2 + (13-9)^2 = (-5)^2 + (1)^2 + (4)^2 = 42$

C: {4,8,12,16,20} $\bar{C} = 12$

$(4-12)^2 + (8-12)^2 + (12-12)^2 + (16-12)^2 + (20-12)^2 = (-8)^2 + (-4)^2 + (0)^2 + (4)^2 + (8)^2 = 160$

One-way ANOVA

Scores from a test (9 students):

{1,3,4,5,5,5,6,7,9}

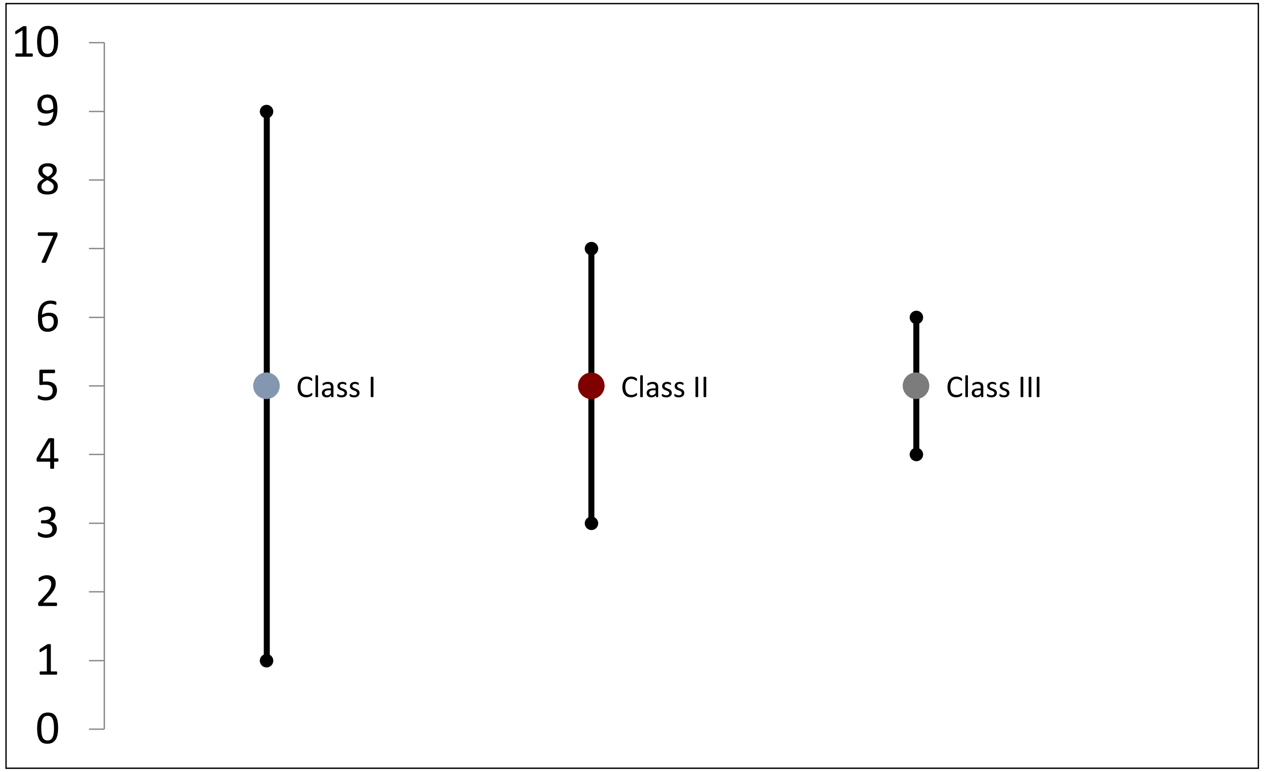
SST = 42

Class I	Class II	Class III
{1,5,9}	{4,5,6}	{3,5,7}

One-way ANOVA

SST = 42

Class I	Class II	Class III
{1,5,9}	{3,5,7}	{4,5,6}
$\overline{X}_I = 5$	$\overline{X}_{II} = 5$	$\overline{X}_{III} = 5$



One-way ANOVA

SST = 42

Class I	Class II	Class III
{1,5,9}	{3,5,7}	{4,5,6}
$\overline{X}_{I} = 5$	$\overline{X}_{II} = 5$	$\overline{X}_{III} = 5$

"Global mean"
Average of averages

SSW = Sum of squares within groups = $\sum(X - \bar{X}i)^2$

SSB = Sum of squares between groups = $\sum(\bar{X}i - \bar{\bar{X}})^2$

$= n_i(\bar{X}i - \bar{\bar{X}})^2$

#obs

One-way ANOVA

SST = 42

Class I	Class II	Class III
{1,5,9}	{3,5,7}	{4,5,6}
$\overline{X}_I = 5$	$\overline{X}_{II} = 5$	$\overline{X}_{III} = 5$

$$SSW = \quad (-4)^2+0^2+4^2 \quad (-2)^2+0^2+2^2 \quad (-1)^2+0^2+1^2 \quad = 42$$

$$32 \qquad 8 \qquad 2$$

$$SSB = \quad 3(0^2) \quad 3(0^2) \quad 3(0^2) \quad = 0$$

$$0 \qquad 0 \qquad 0$$

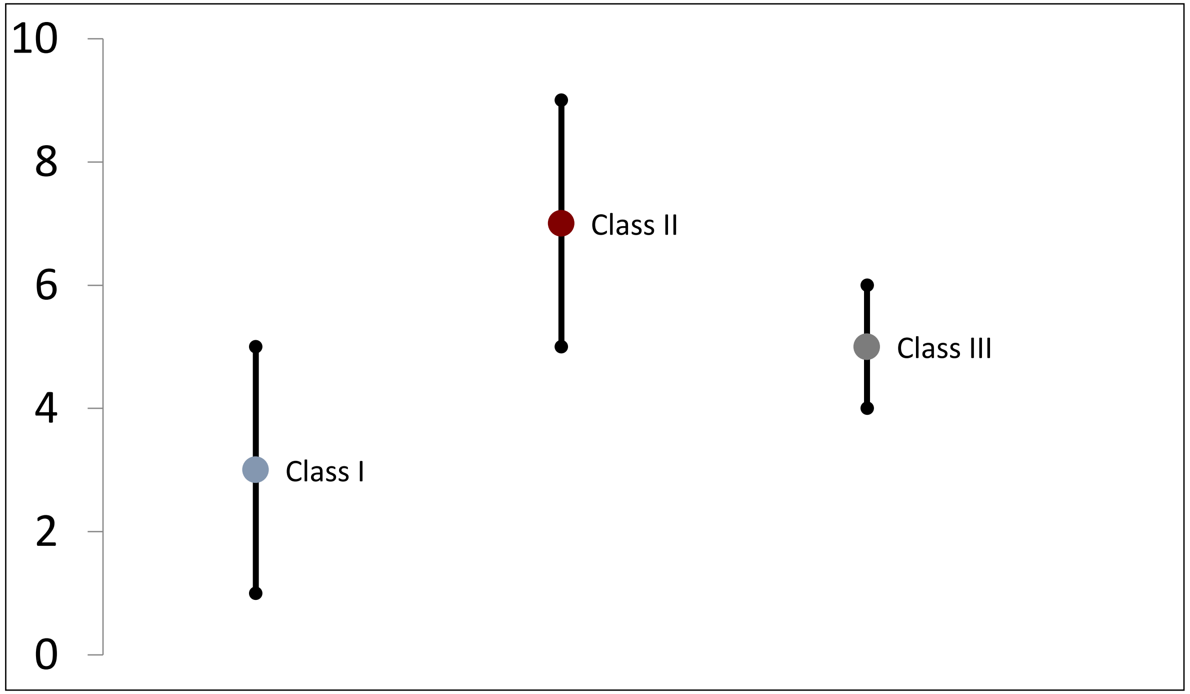
#obs

One-way ANOVA

SST = 42

What if the group members were changed around? →

Class I	Class II	Class III
{1,3,5}	{5,7,9}	{4,5,6}
$\bar{X}_{I} = 3$	$\bar{X}_{II} = 7$	$\bar{X}_{III} = 5$



One-way ANOVA

$SST = 42$

SST = SSW + SSB

	Class I	Class II	Class III	
	{1,3,5}	{5,7,9}	{4,5,6}	
	$\overline{X}_I = 3$	$\overline{X}_{II} = 7$	$\overline{X}_{III} = 5$	
SSW =	$(-2)^2+0^2+2^2$	$(-2)^2+0^2+2^2$	$(-1)^2+0^2+1^2$	$= \boxed{18}$
	8	8	2	
SSB =	$3(3-5)^2$	$3(7-5)^2$	$3(5-5)^2$	$= \boxed{24}$
	12	12	0	

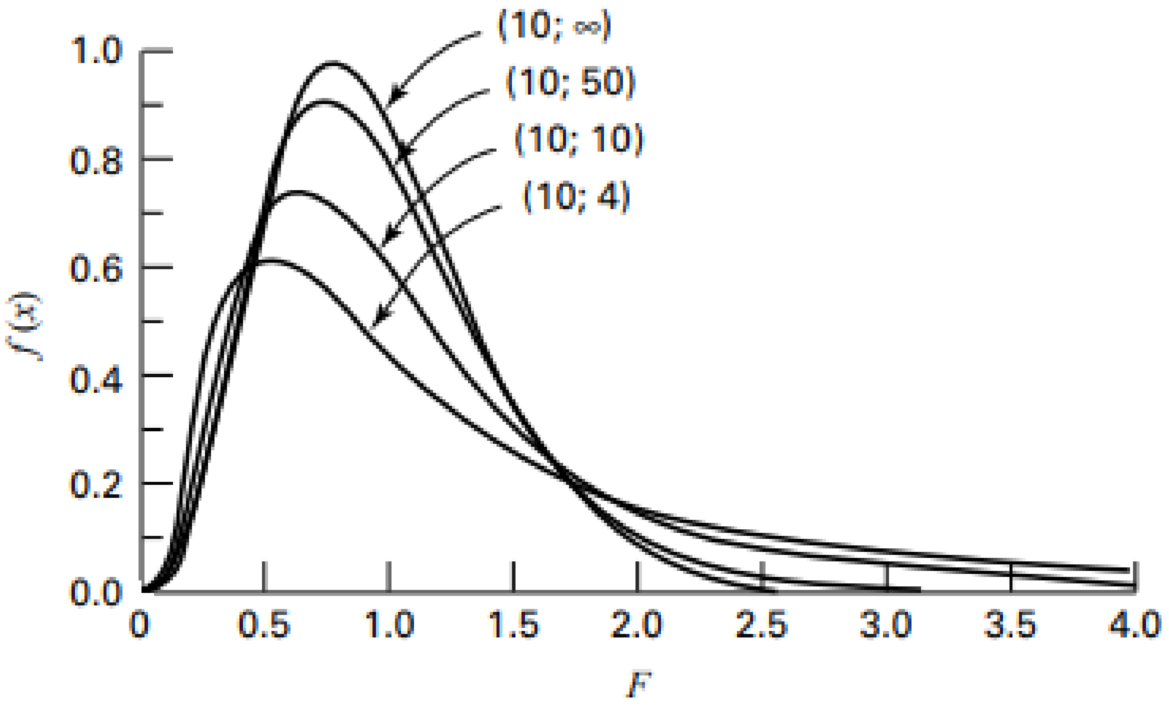
The F Distribution

$$(s_1^2/\sigma_1^2)/(s_2^2/\sigma_2^2)$$

When population variances are equal, we can reduce to:

$$s_1^2 \;/\; s_2^2$$

Which is also distributed as *F*



F Distribution is actually a family of distributions dependent on numerator and denominator degrees of freedom.

One-way ANOVA

Class I, Class II, Class III

$$SST = SSW + SSB$$

Mean Sum of Squares Between Groups

categories (3)

$$F = \frac{MSB}{MSW} = \frac{SSB / \boxed{(c - 1)}}{SSW / \boxed{(n - c)}}$$

Numerator degrees of freedom

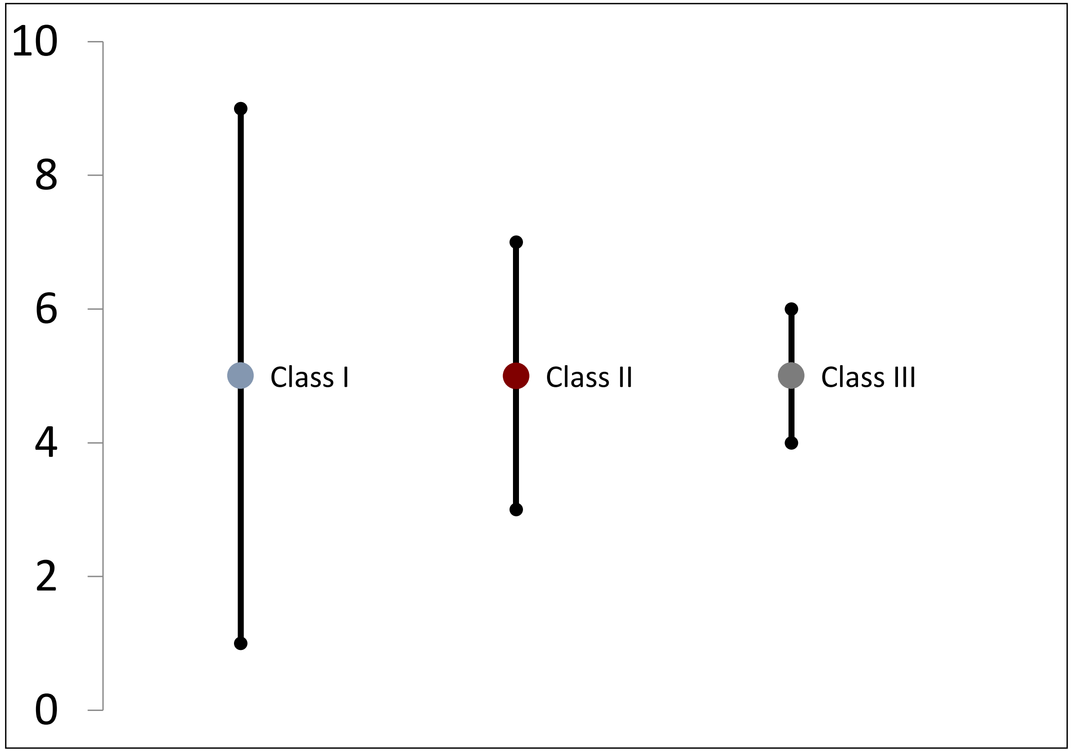
Denominator degrees of freedom

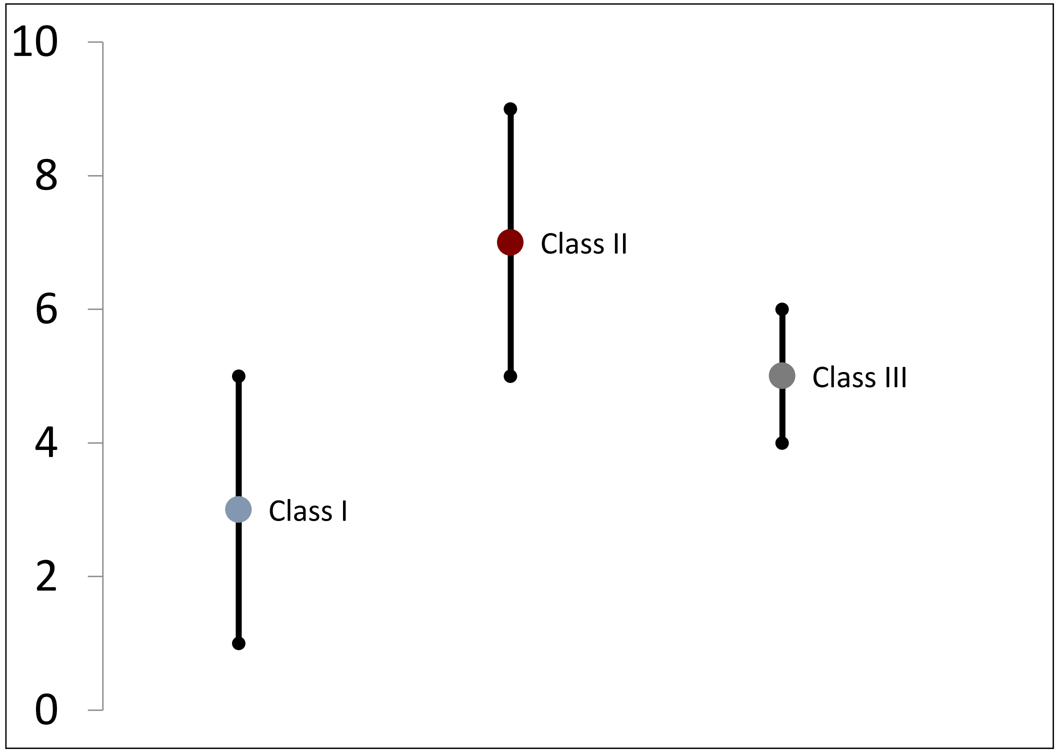
Mean Sum of Squares Within Groups

observations (9)

{1,3,4,5,5,5,6,7,9}

One-way ANOVA





SSW = 42
SSB = 0

F = 0

$$F = \frac{SSB / (c - 1)}{SSW / (n - c)}$$

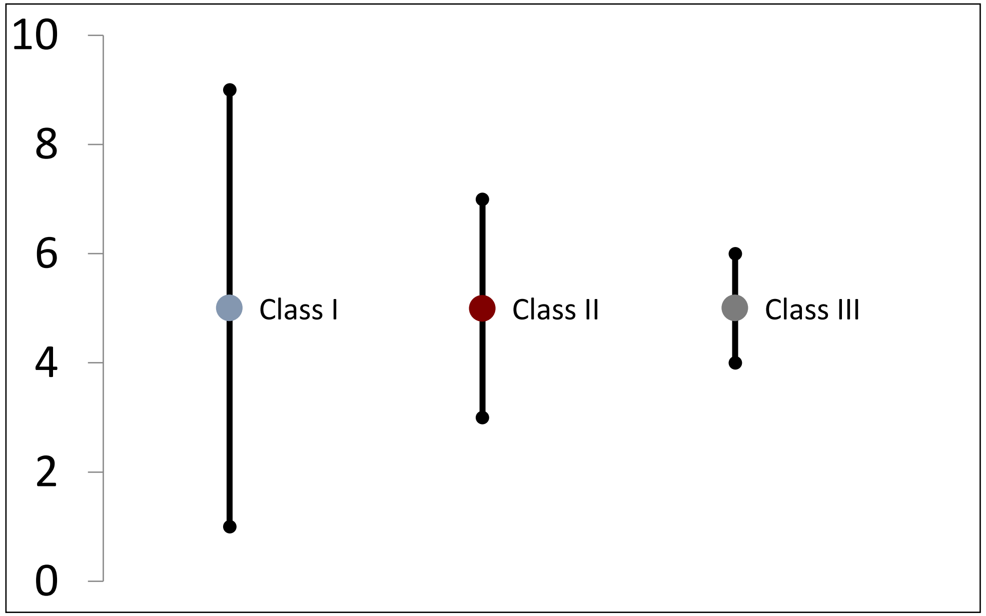
SSW = 18
SSB = 24

F = 4.0

One-way ANOVA

$H_0: \mu_I = \mu_{II} = \mu_{III}$

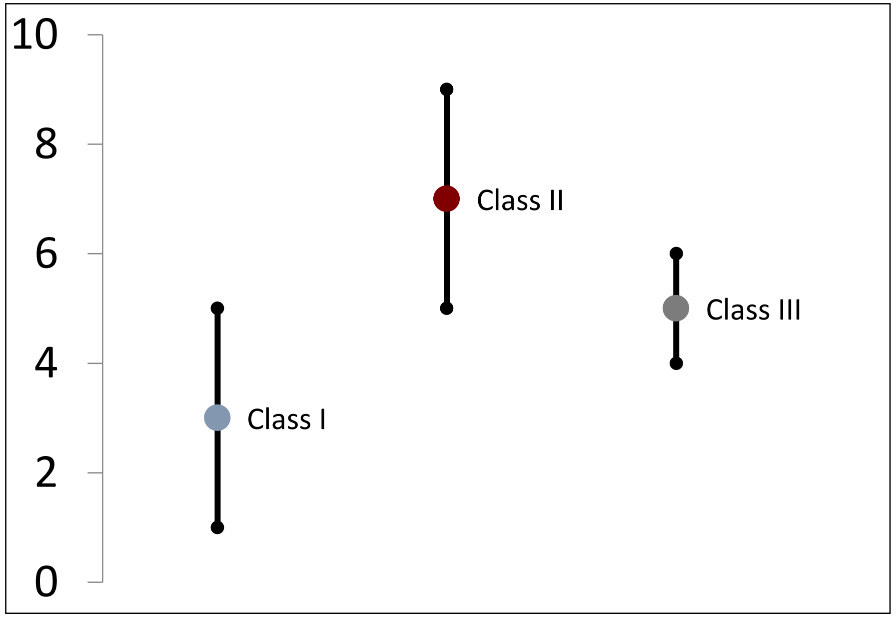
Class I	Class II	Class III
{1,5,9}	{3,5,7}	{4,5,6}



F = 0
(p=1.0000)

Do not reject H_0

Class I	Class II	Class III
{1,3,5}	{5,7,9}	{4,5,6}

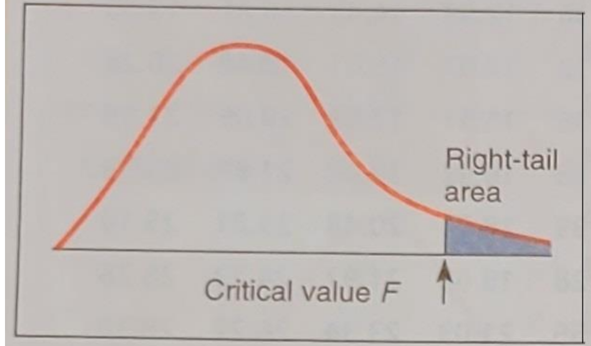


F = 4.0
(p=0.0787)

Do not reject H_0 (at 5% significance)

Example F Statistic Table

TABLE 8 Critical Values For F Distribution



Degrees of freedom denominator, $d.f._D$	Right-tail area	Degrees of freedom numerator, $d.f._N$: 1	2	3	4	5	6	7	8	9
	0.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	0.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
1	0.025	647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66	963.28
	0.010	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5
	0.001	405284	500000	540379	562500	576405	585937	592873	598144	602284
	0.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	0.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
2	0.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
	0.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	0.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36	999.37	999.39
	0.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	0.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
3	0.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
	0.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	0.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58	130.62	129.86

For previous example, with numerator d.f.=2 and denominator d.f. = 7, the .05 cutoff is 4.74 (not shown in this image)

Performing the One-way ANOVA in Python

```
from scipy.stats import f_oneway
from statsmodels.stats.multicomp import pairwise_tukeyhsd

#scores from the three classes
class_1 = [1,3,5]
class_2 = [5,7,9]
class_3 = [4,5,6]

#Conduct the one-way ANOVA
print(f_oneway(class_1, class_2, class_3))
```

```
F_onewayResult(statistic=4.0, pvalue=0.07871720116618075)
```

Another example of One-way ANOVA

- An advanced manufacturing process has three identical tools used interchangeably for a step in the process
- It is suspected that one of these is introducing more defects per-run than the others
- Using sampled data on defect counts, is there a statistical difference?

One-way ANOVA

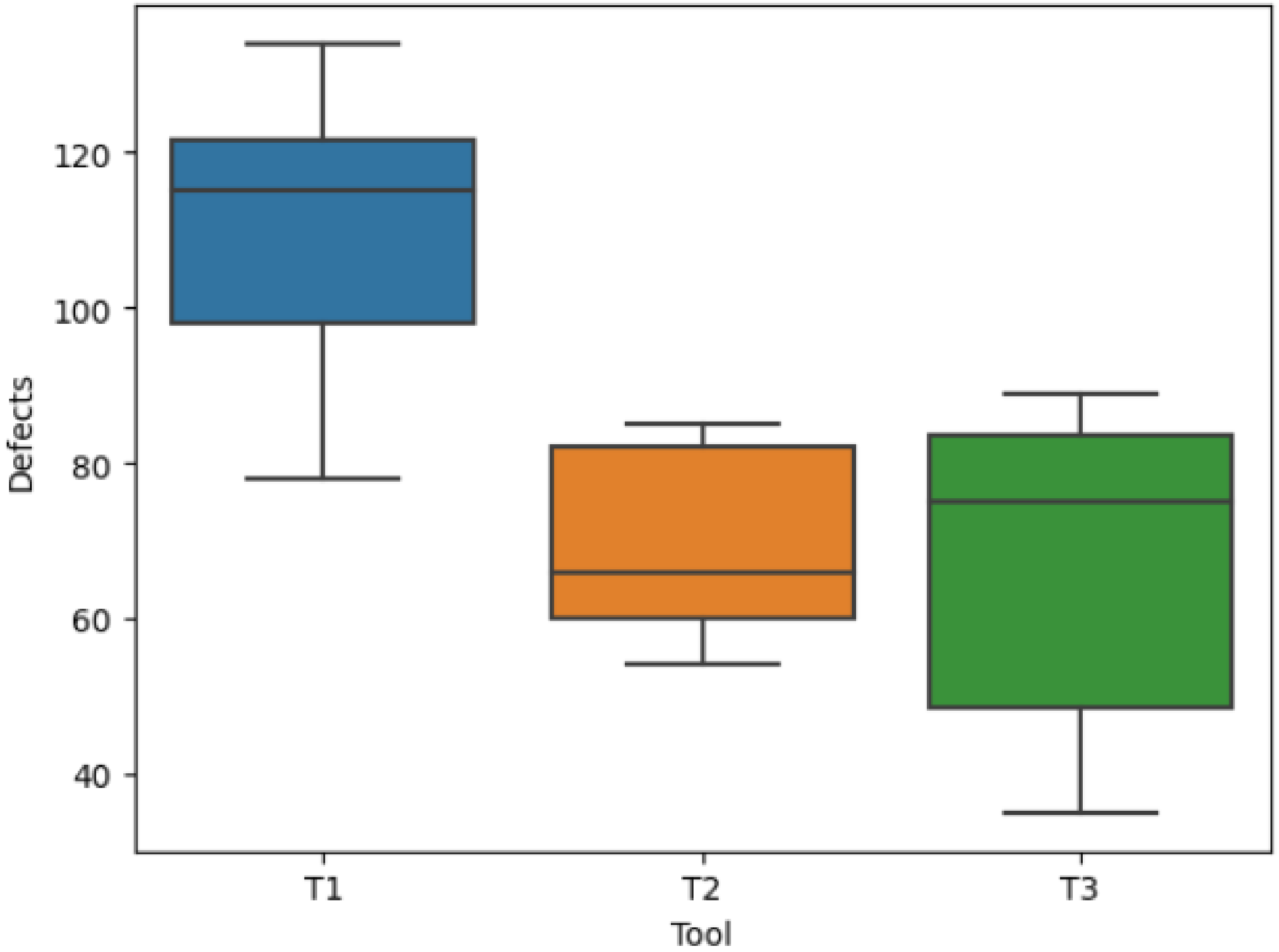
```
tool1_defects = [118, 78, 115, 134, 125, 134, 102, 103, 115, 115, 104, 94, 86, 132, 86]
tool2_defects = [60, 84, 82, 85, 71, 55, 64, 66, 54]
tool3_defects = [85, 82, 35, 85, 76, 42, 64, 85, 79, 62, 43, 84, 69, 75, 89, 75, 44, 42]
#calculate the test statistic
print(f_oneway(tool1_defects, tool2_defects, tool3_defects))
```

```
F_onewayResult(statistic=27.995841285913606, pvalue=2.88995460267374e-08)
```

Post Hoc Analysis & Testing - Visualization

```
import seaborn as sns
fig = sns.boxplot(data = [tool1_defects, tool2_defects, tool3_defects])
fig.set(ylabel='Defects', xlabel='Tool')
fig.set_xticklabels(['T1','T2','T3'])
```

```
[Text(0, 0, 'T1'), Text(1, 0, 'T2'), Text(2, 0, 'T3')]
```



Post Hoc Analysis & Testing – Tukey's HSD

```
defects = []
defects.extend(tool1_defects)
defect_groups = []
for obs in tool1_defects:
    defect_groups.append('tool_1')

defects.extend(tool2_defects)
for obs in tool2_defects:
    defect_groups.append('tool_2')

defects.extend(tool3_defects)
for obs in tool3_defects:
    defect_groups.append('tool_3')

print(defects)
print(defect_groups)
```

```
[118, 78, 115, 134, 125, 134, 102, 103, 115, 115, 104, 94, 86, 132, 86, 60, 84, 82, 85, 71, 55, 64, 6
6, 54, 85, 82, 35, 85, 76, 42, 64, 85, 79, 62, 43, 84, 69, 75, 89, 75, 44, 42]
['tool_1', 'tool_1', 'tool_1', 'tool_1', 'tool_1', 'tool_1', 'tool_1', 'tool_1', 'tool_1', 'tool_1',
'tool_1', 'tool_1', 'tool_1', 'tool_1', 'tool_1', 'tool_2', 'tool_2', 'tool_2', 'tool_2', 'tool_2',
'tool_2', 'tool_2', 'tool_2', 'tool_2', 'tool_3', 'tool_3', 'tool_3', 'tool_3', 'tool_3', 'tool_3',
'tool_3', 'tool_3', 'tool_3', 'tool_3', 'tool_3', 'tool_3', 'tool_3', 'tool_3', 'tool_3', 'tool_3',
'tool_3', 'tool_3']
```

Post Hoc Analysis & Testing – Tukey's HSD

```python
tukey = pairwise_tukeyhsd(endog=defects, groups=defect_groups, alpha=0.05)
print(tukey)
```

```
 Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====================================================
group1 group2 meandiff p-adj   lower    upper   reject
-----------------------------------------------------
tool_1 tool_2    -40.4    0.0 -58.0396 -22.7604   True
tool_1 tool_3 -41.8444    0.0 -56.4704 -27.2185   True
tool_2 tool_3  -1.4444 0.9769 -18.5239   15.635  False
-----------------------------------------------------
```

Two-way ANOVA

Factors

Manufacturer

Age

Mfr	Defects	Tool Age Group
M1	4	New
M1	6	New
M1	8	New
M2	4	New
M2	8	New
M2	9	New
M1	6	Moderate
M1	6	Moderate
M1	9	Moderate
M2	7	Moderate
M2	10	Moderate
M2	13	Moderate
M1	8	Old
M1	9	Old
M1	13	Old
M2	12	Old
M2	14	Old
M2	16	Old

Two-way ANOVA

	New	Moderate	Old	
M1	4	6	8	
	6	6	9	
	8	9	13	
means →	6	7	10	**7.7**
M2	4	7	12	
	8	10	14	
	9	13	16	
means →	7	10	14	**10.3**
				Total Average
	6.5	**8.5**	**12**	9

Two-way ANOVA

2[nd] factor (Age)

1[st] Factor (Manufacturer)

	New	Moderate	Old	Average
M1	6	7	10	**7.7**
M2	7	10	14	**10.3**
Average	**6.5**	**8.5**	**12**	**9**

Vs One-way ANOVA (by age)

	New	Moderate	Old	Total Avg
Average	**6.5**	**8.5**	**12**	**9**

Two-way ANOVA

Sum of Squares 1st Factor (Manufacturer)

+

Sum of Squares 2nd Factor (Age)

+

Sum of Squares Error (Within)

+

Sum of Squares Both Factors

=

Sum of Squares Total

Note: The “error” variation of two-way ANOVA corresponds to the “within-group” variation of one-way ANOVA

Sum of Squares

Remember, we are using a concept related to sample variance. We are first concerned with the numerator portion and then use the relevant degrees of freedom for the denominator.

$$s^2 = \frac{\boxed{\Sigma(X - \bar{X})^2}}{n - 1}$$

Calculating SS_{F1}

Sum of Squares 1st Factor (**Manufacturer**)

Score	M1 Mean		Grand Mean	*
4	7.7	-	9	$= (-1.3)^2 = 1.8$
6	7.7	-	9	$= (-1.3)^2 = 1.8$
8	7.7	-	9	$= (-1.3)^2 = 1.8$
6	7.7	-	9	$= (-1.3)^2 = 1.8$
6	7.7	-	9	$= (-1.3)^2 = 1.8$
9	7.7	-	9	$= (-1.3)^2 = 1.8$
8	7.7	-	9	$= (-1.3)^2 = 1.8$
9	7.7	-	9	$= (-1.3)^2 = 1.8$
13	7.7	-	9	$= (-1.3)^2 = 1.8$

sum of squares = 16

Score	M2 Mean		Grand Mean	
4	10.3	-	9	$= (1.3)^2 = 1.8$
8	10.3	-	9	$= (1.3)^2 = 1.8$
9	10.3	-	9	$= (1.3)^2 = 1.8$
7	10.3	-	9	$= (1.3)^2 = 1.8$
10	10.3	-	9	$= (1.3)^2 = 1.8$
13	10.3	-	9	$= (1.3)^2 = 1.8$
12	10.3	-	9	$= (1.3)^2 = 1.8$
14	10.3	-	9	$= (1.3)^2 = 1.8$
16	10.3	-	9	$= (1.3)^2 = 1.8$

sum of squares = 16

sum of squares for 1st Factor = 16 + 16 = 32
(Manufacturer)

* - discrepancy here is due to rounding of M1 mean to 7.7 (i.e., $(-1.3)^2 = 1.69$)

Calculating SS_{F2}

Sum of Squares 2nd Factor (Age)

M1

4	$6.5 - 9 = (-2.5)^2 = 6.3$
6	$6.5 - 9 = (-2.5)^2 = 6.3$
8	$6.5 - 9 = (-2.5)^2 = 6.3$
6	$8.5 - 9 = (-.5)^2 = .25$
6	$8.5 - 9 = (-.5)^2 = .25$
9	$8.5 - 9 = (-.5)^2 = .25$
8	$12 - 9 = (3)^2 = 9.0$
9	$12 - 9 = (3)^2 = 9.0$
13	$12 - 9 = (3)^2 = 9.0$

sum of squares = 46.5

M2

4	$6.5 - 9 = (-2.5)^2 = 6.3$
8	$6.5 - 9 = (-2.5)^2 = 6.3$
9	$6.5 - 9 = (-2.5)^2 = 6.3$
7	$8.5 - 9 = (-.5)^2 = .25$
10	$8.5 - 9 = (-.5)^2 = .25$
13	$8.5 - 9 = (-.5)^2 = .25$
12	$12 - 9 = (3)^2 = 9.0$
14	$12 - 9 = (3)^2 = 9.0$
16	$12 - 9 = (3)^2 = 9.0$

sum of squares = 46.5

sum of squares for 2nd Factor = 93.0
Age

Calculating SS_E

Sum of Squares Within (Error)

M1

4	-	6	$= (-2.0)^2$	$= 4.0$
6	-	6	$= (\ 0\ )^2$	$= 0.0$
8	-	6	$= (2.0)^2$	$= 4.0$
6	-	7	$= (-2.0)^2$	$= 4.0$
6	-	7	$= (-1.0)^2$	$= 1.0$
9	-	7	$= (2.0)^2$	$= 4.0$
8	-	10	$= (-2.0)^2$	$= 4.0$
9	-	10	$= (-1.0)^2$	$= 1.0$
13	-	10	$= (3.0)^2$	$= 9.0$

sum of squares = 28.0

M2

4	-	7	$= (-3.0)^2$	$= 9.0$
8	-	7	$= (1.0)^2$	$= 1.0$
9	-	7	$= (2.0)^2$	$= 4.0$
7	-	10	$= (-3.0)^2$	$= 9.0$
10	-	10	$= (\ 0\ )^2$	$= 0.0$
13	-	10	$= (3.0)^2$	$= 9.0$
12	-	14	$= (-2.0)^2$	$= 4.0$
14	-	14	$= (\ 0\ )^2$	$= 0.0$
16	-	14	$= (2.0)^2$	$= 4.0$

sum of squares = 40.0

total sum of squares within = 68

Calculating SS_{Tot}

Score	Grand Mean	(Score - Grand Mean)^2
4	- 9	$=(-5)^2 = 25.0$
6	- 9	$=(-3)^2 = 9.0$
8	- 9	$=(-1)^2 = 1.0$
6	- 9	$=(-3)^2 = 9.0$
6	- 9	$=(-3)^2 = 9.0$
9	- 9	$=(0)^2 = 0.0$
8	- 9	$=(-1)^2 = 1.0$
9	- 9	$=(0)^2 = 0.0$
13	- 9	$=(4)^2 = 16.0$
4	- 9	$=(-5)^2 = 25.0$
8	- 9	$=(1)^2 = 1.0$
9	- 9	$=(0)^2 = 0.0$
7	- 9	$=(-2)^2 = 4.0$
10	- 9	$=(1)^2 = 1.0$
13	- 9	$=(4)^2 = 16.0$
12	- 9	$=(-3)^2 = 9.0$
14	- 9	$=(5)^2 = 25.0$
16	- 9	$=(7)^2 = 49.0$
		200

Sum of Squares 1st Factor (Manufacturer)	32	
+	+	
Sum of Squares 2nd Factor (Age)	93	= 193
+	+	
Sum of Squares Error (Within)	68	
+	+	
Sum of Squares Both Factors	?	= 200 – 193 = 7
=	=	
Sum of Squares Total	200	

Hypotheses

H_0: Manufacturer will have no significant effect on defect count

H_0: Age will have no significant effect on defect count

H_0: Manufacturer and age interaction will have no significant effect on defect count

Two-way ANOVA – Degrees of Freedom (d.f.)

2nd factor SS d.f. = number of categories – 1 = 3 – 1 = 2

	New	Moderate	Old	Average
M1	6	7	10	**7.7**
M2	7	10	14	**10.3**
Average	**6.5**	**8.5**	**12**	9

1st Factor SS
d.f. = number of categories – 1
= 2 – 1
=1

***Sum of Squares Both Factors:**
d.f. = (1st Factor SS d.f) * (2nd Factor SS d.f) = 1 * 2 = **2**

Vs One-way ANOVA (by age)

	New	Moderate	Old	Total Avg
Average	**6.5**	**8.5**	**12**	9

Two-way ANOVA – Degrees of Freedom (d.f.) (continued)

Sum of Squares Within (Error)

degrees of freedom

M1

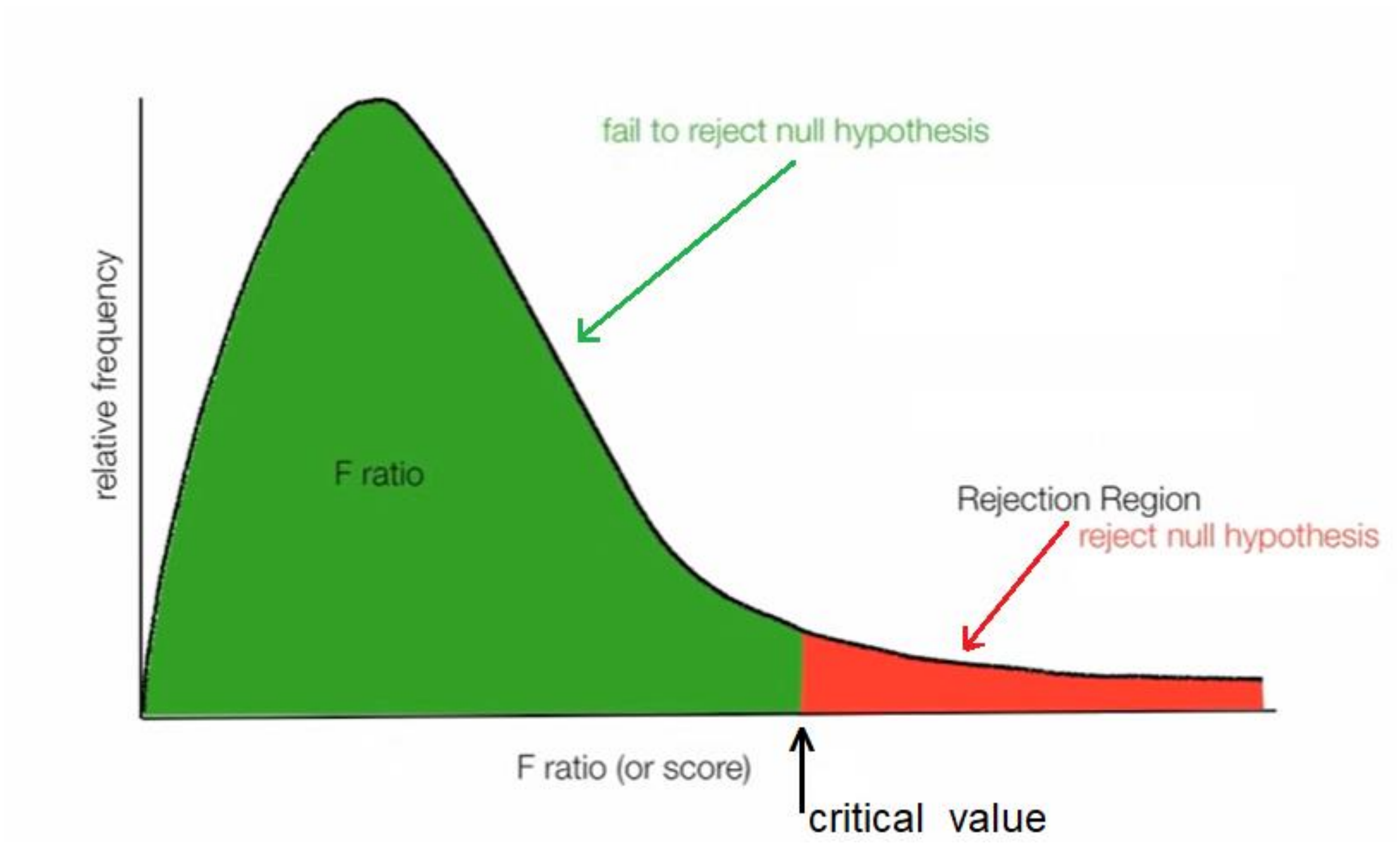
new	moderate	old
4	6	8
6	6	9
8	9	13
n - 1	n - 1	n - 1
3 - 1	3 - 1	3 - 1
2	2	2

M2

new	moderate	old
4	7	12
8	10	14
9	13	16
n - 1	n - 1	n - 1
3 - 1	3 - 1	3 - 1
2	2	2

$2 + 2 + 2 + 2 + 2 + 2 = 12$

Interpreting the 2-way ANOVA Results



Calculating the F-Statistic (aka F-ratio, F-Score) For 1st Factor Analysis

Degrees of Freedom

	Sum of Squares	d.f.	Mean Square	F Score
Sum of Squares 1st Factor (Manufacturer)	32	1	$\frac{32}{\boxed{1}} = 32$	$\frac{32}{5.67} = 5.64$
Sum of Squares Within (Error)	68	12	$\frac{68}{\boxed{12}} = 5.67$	

Numerator degrees of freedom

Denominator degrees of freedom

Critical Value for 1st Factor Sum of Squares

F Distribution

degrees of freedom numerator

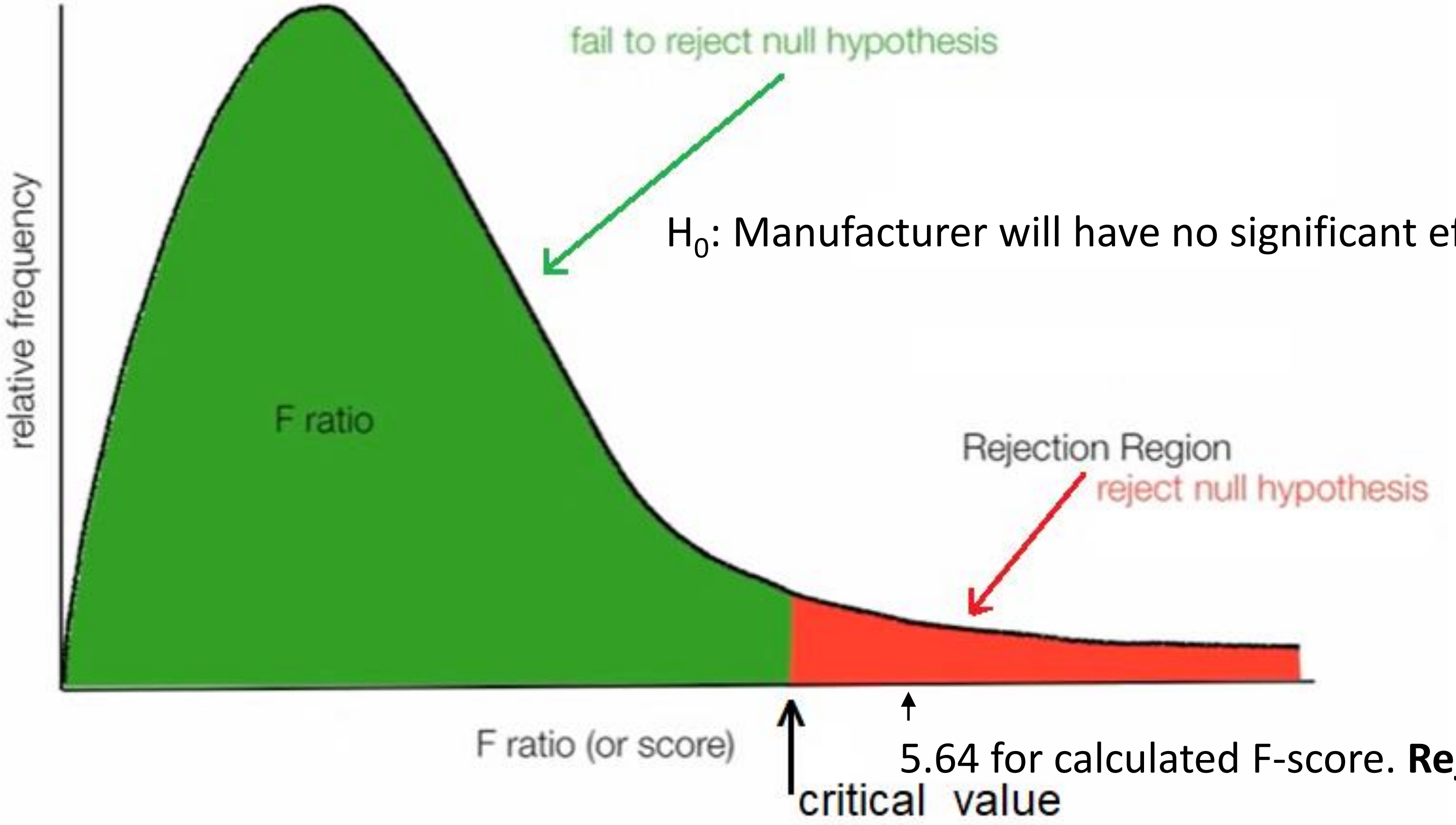
Table for .05 right tail area

degrees of freedom denominator

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	161.5	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	246.0	248.0	249.1	250.1
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25

$F(1,12) = 5.64$, $p < .05$

Interpreting the 2-way ANOVA Results



H_0: Manufacturer will have no significant effect on defect count

5.64 for calculated F-score. **Reject null hypothesis!**

4.75 (for .05 significance threshold)

Calculating the F-Statistic For 2nd Factor Analysis

Degrees of Freedom

	Sum of Squares	d.f.	Mean Square	F Score
Sum of Squares 2nd Factor (Age)	93	2	$\frac{93}{2} = 46.50$	$\frac{46.50}{5.67} = 8.20$
Sum of Squares Within (Error)	68	12	$\frac{68}{12} = 5.67$	

Critical Value for 2nd Factor Sum of Squares

F Distribution

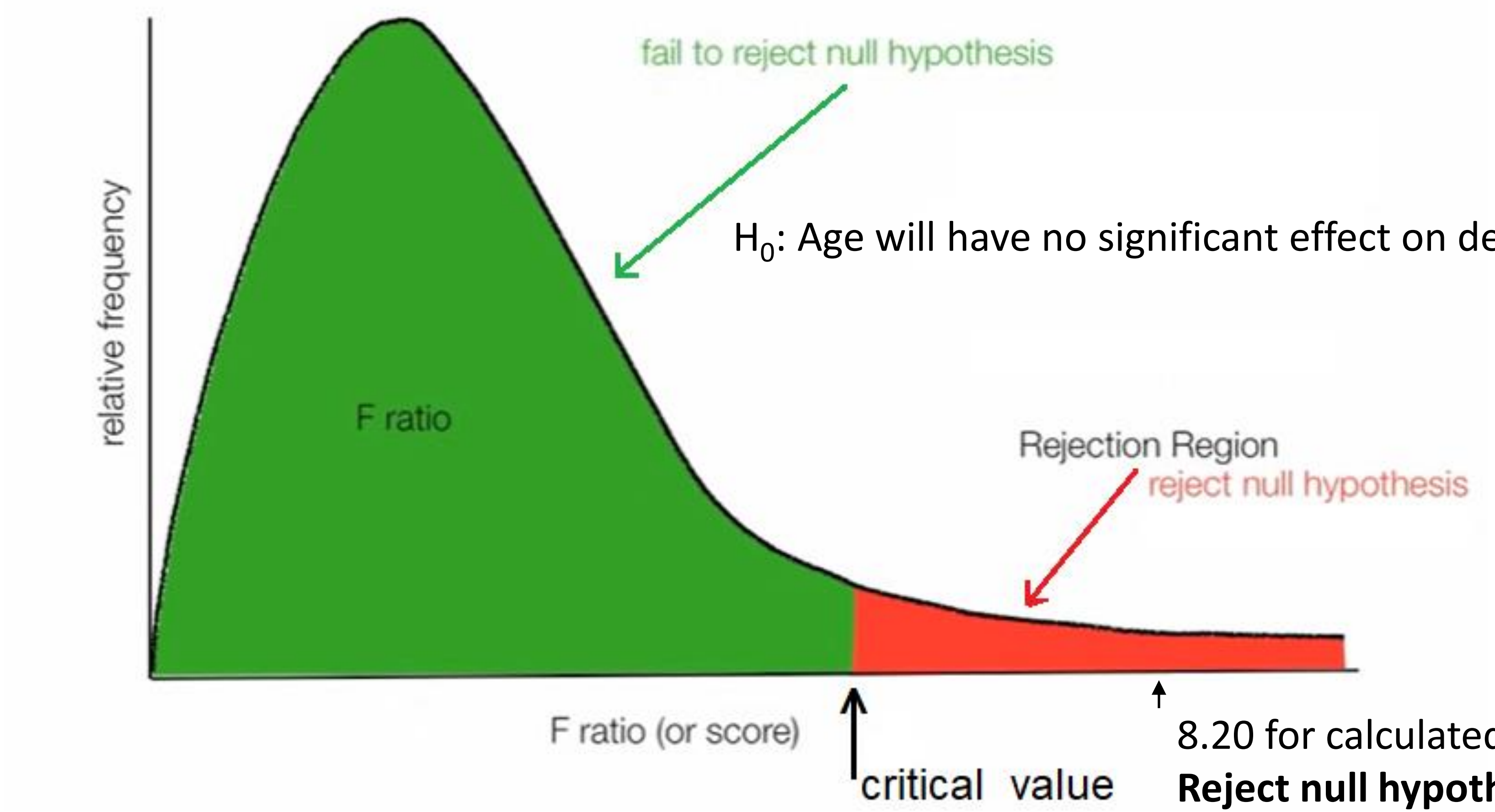
Table for .05 right tail area

degrees of freedom numerator

degrees of freedom denominator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	161.5	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	246.0	248.0	249.1	250.1
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25

$F(2,12) = 8.20 \quad p<.05$

Interpreting the 2-way ANOVA Results



H_0: Age will have no significant effect on defect count

8.20 for calculated F-score.
Reject null hypothesis!

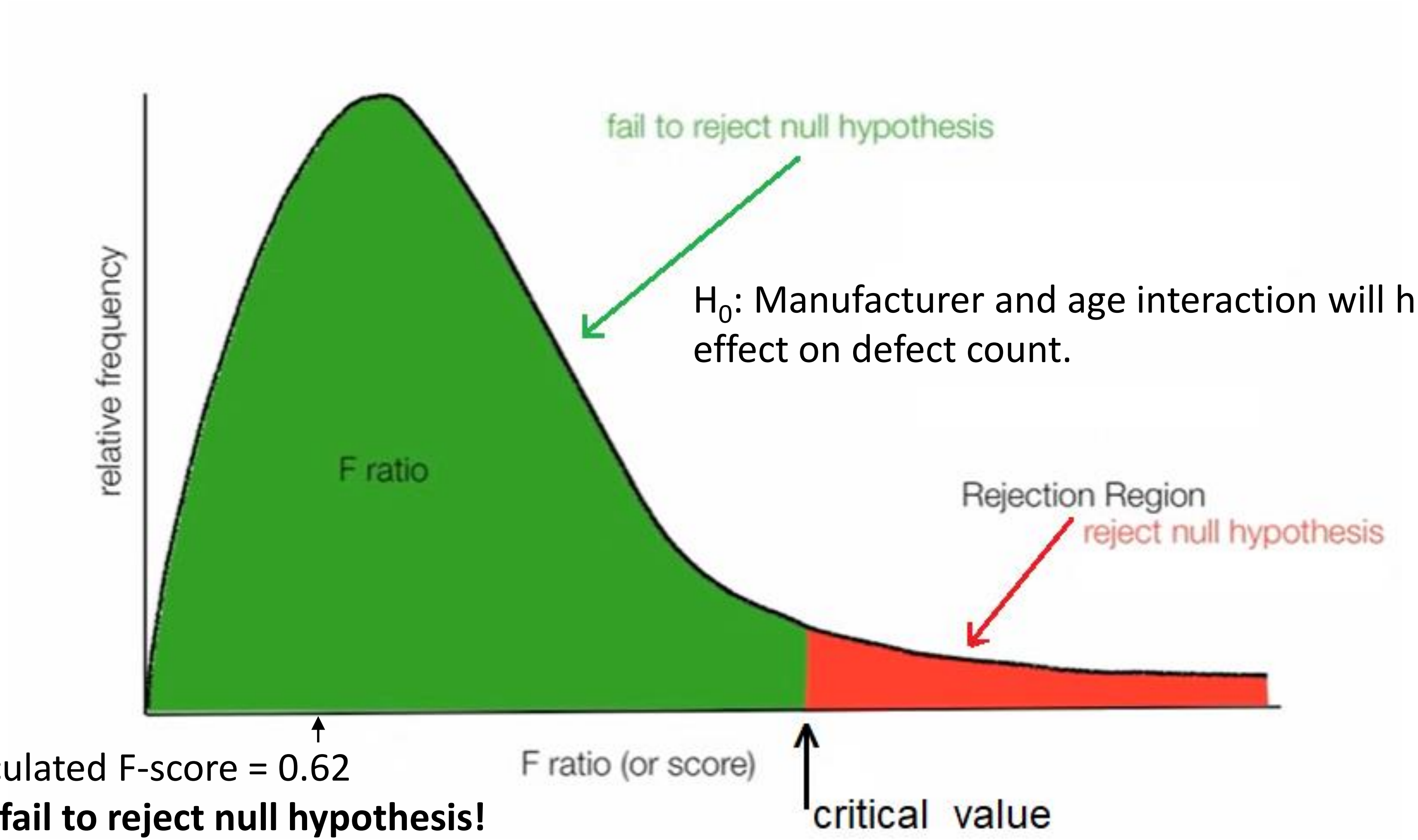
3.89 (for .05 significance threshold)

Calculating the F-Statistic For Both Factors

Degrees of Freedom

	Sum of Squares	d.f.	Mean Square	F Score
Sum of Square Both Factors	7	2	$\frac{7}{2} = 3.5$	$\frac{3.5}{5.67} = .62$
Sum of Squares Within (Error)	68	12	$\frac{68}{12} = 5.67$	

Interpreting the 2-way ANOVA Results



H_0: Manufacturer and age interaction will have no significant effect on defect count.

Calculated F-score = 0.62

We fail to reject null hypothesis!

3.89 (for .05 significance threshold)

In Python – create data

```
Mfr,Defects,Age
M1,4,New
M1,6,New
M1,8,New
M2,4,New
M2,8,New
M2,9,New
M1,6,Moderate
M1,6,Moderate
M1,9,Moderate
M2,7,Moderate
M2,10,Moderate
M2,13,Moderate
M1,8,Old
M1,9,Old
M1,13,Old
M2,12,Old
M2,14,Old
M2,16,Old
```

Create a .csv file named:
“2way_data.csv”

In Python – load into a dataframe

```python
import pandas as pd
df = pd.read_csv("./2way_data.csv")
df.head()
```

	Mfr	Defects	Age
0	M1	4	New
1	M1	6	New
2	M1	8	New
3	M2	4	New
4	M2	8	New

Python – perform the 2-way ANOVA

```python
# Importing libraries
import statsmodels.api as sm
from statsmodels.formula.api import ols

# Performing two-way ANOVA
model = ols('Defects ~ C(Mfr) + C(Age) + C(Mfr):C(Age)', data=df).fit()
result = sm.stats.anova_lm(model, type=2)
print(result)
```

```
                 df  sum_sq    mean_sq         F    PR(>F)
C(Mfr)          1.0    32.0  32.000000  5.647059  0.034994
C(Age)          2.0    93.0  46.500000  8.205882  0.005677
C(Mfr):C(Age)   2.0     7.0   3.500000  0.617647  0.555502
Residual       12.0    68.0   5.666667       NaN       NaN
```

Python – perform post-hoc Tukey HSD

```
from statsmodels.stats.multicomp import pairwise_tukeyhsd
df['combination'] = df.Mfr + " / " + df.Age
m_comp = pairwise_tukeyhsd(endog=df['Defects'], groups=df['combination'], alpha=0.05)
```

```
m_comp.summary()
```

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
M1 / Moderate	M1 / New	-1.0	0.9945	-7.5286	5.5286	False
M1 / Moderate	M1 / Old	3.0	0.6459	-3.5286	9.5286	False
M1 / Moderate	M2 / Moderate	3.0	0.6459	-3.5286	9.5286	False
M1 / Moderate	M2 / New	0.0	1.0	-6.5286	6.5286	False
M1 / Moderate	M2 / Old	7.0	0.0332	0.4714	13.5286	True
M1 / New	M1 / Old	4.0	0.3676	-2.5286	10.5286	False
M1 / New	M2 / Moderate	4.0	0.3676	-2.5286	10.5286	False
M1 / New	M2 / New	1.0	0.9945	-5.5286	7.5286	False

Multiple Testing Correction (MTC)

- False Discovery Rate (FDR) – “false positives”
- Adjusted p-values (sometimes referred to as “q-vals”)

For example: In a biological microarray experiment, several thousand genes may be simultaneously tested across different conditions.

When testing for potential differential expression across those conditions, each gene is considered independently from one another.

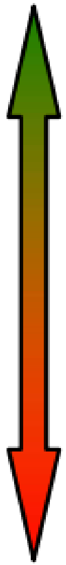
In other words, a t-test or ANOVA is performed on each gene separately.

The incidence of false positives (or genes falsely called differentially expressed when they are not) is proportional to the number of tests performed and the critical significance level (p-value cutoff).

MTC - Approaches

Bonferroni
Bonferroni Step-Down
Westfall and Young Permutation
Benjamini and Hochberg False Discovery Rate
None

More false negatives



More false positives

Bonferroni Correction

The p-value of each gene is multiplied by the number of genes in the gene list. If the corrected p-value is still below the error rate, the gene will be significant:

Corrected P-value= p-value * n (number of genes in test) <0.05

As a consequence, if testing 1000 genes at a time, the highest accepted individual p-value is 0.00005, making the correction very stringent. With a Family-wise error rate of 0.05 (i.e., the probability of at least one error in the family), the expected number of false positives will be 0.05.

Benjamini and Hochberg False Discovery Rate

This correction is the least stringent of all 4 options, and therefore tolerates more false positives. There will be also less false negative genes. Here is how it works:

1) The p-values of each gene are ranked from the smallest to the largest.
2) The largest p-value remains as it is.
3) The second largest p-value is multiplied by the total number of genes in gene list divided by its rank. If less than 0.05, it is significant.
Corrected p-value = p-value*(n/n-1) < 0.05, if so, gene is significant.
4) The third p-value is multiplied as in step 3:
Corrected p-value = p-value*(n/n-2) < 0.05, if so, gene is significant.

And so on.

Benjamini and Hochberg - example

Let n=1000, error rate=0.05

Gene name	p-value (from largest to smallest)	Rank	Correction	Is gene significant after correction?
A	0.1	1000	No correction	0.1 > 0.05 => No
B	0.06	999	1000/999*0.06 = 0.06006	0.06006 > 0.05 => No
C	0.04	998…	1000/998*0.04 = 0.04008	0.04008 < 0.05 => Yes

As you can see from the example above, the correction becomes more stringent as the p-value decreases, similarly as the Bonferroni Step-down correction. This method provides a good alternative to Family-wise error rate methods. The error rate is a proportion of the number of called genes.