Page No.
Date

Module 5

Q1 ANOVA (Create One way table) 2m

Sources of Variation	Sum of Squares	Degree of freedom (v)	Mean Square	Variance Ratio
Between Samples	SSC (columns)	$v_1 = c-1$	MSC $= SSC/(c-1)$	$F = \frac{MSC}{MSE}$
Within Samples	SSE (Rows)	$v_2 = n-c$	MSE $= SSE/(n-c)$	
Total	SST	$n-1$		

Total of sum of squares

Q2 ANOVA (create two way table). 2m

Sources of Variation	Sum of square	Degrees of freedom (v)	Mean Squares	Variance Ratio (f)
Between Samples	SSC	$(c-1)$	$MSC = \frac{SSC}{(c-1)}$	MSC / MSE
Between Rows	SSR	$(r-1)$	$MSR = \frac{SSR}{(r-1)}$	MSR / MSE
Residual or Error	SSE	$(c-1)(r-1)$	$MSE = \frac{SSE}{(c-1)(r-1)}$	
Total	SST	$(n-1)$		

Page No.

Date

Q3 F - Test (sum) 5M

→

A	$x-\bar{x}$ ($\bar{x}=80$)	$(x-\bar{x})^2$	B	$(x-\bar{x})$ ($\bar{x}=83$)	$(x-\bar{x})^2$
66	-14	196	64	-19	361
67	-13	169	66	-17	289
75	-5	25	74	-9	81
76	-4	16	78	-5	25
82	2	4	82	-1	1
84	4	16	85	2	4
88	8	64	87	4	16
90	10	100	92	9	81
92	12	144	93	10	100
			95	12	144
	$\Sigma(x-\bar{x})^2 = 734$		97	14	196
					$\Sigma(x-\bar{x})^2 = 1298$

$$S_1 = \frac{(x-\bar{x})^2}{n-1} = \frac{734}{9-1} = \underline{\underline{91.75}}$$

$$S_2 = \frac{(x-\bar{x})^2}{n-1} = \frac{1298}{11-1} = \underline{\underline{129.8}}$$

$$F = \frac{S_1^2}{S_2^2} = \frac{91.75}{129.8} = \underline{\underline{0.7067}}$$

Page No.
Date

Q Two Random sample where drawn from two normal population test wheather two population have the same varience at 5% level of signifiance.

A	$x-\bar{x}$ ($\bar{x}=30$)	$(x-\bar{x})^2$	B	$x-\bar{x}$ ($\bar{x}=34.9$)	$(x-\bar{x})^2$
16	-14	196	16	-18.9	357.21
17	-13	169	24	-10.9	118.81
25	-5	25	28	-6.9	47.61
26	-4	16	32	-2.9	8.41
32	2	4	35	0.1	0.01
34	4	16	37	2.1	4.41
38	8	64	42	7.1	50.41
40	10	100	43	8.1	65.61
42	12	144	45	10.1	102.01
		$\Sigma(x-\bar{x})^2 = 734$	47	12.1	146.41
					$\Sigma(x-\bar{x})^2 = 900.9$

$S_1^2 = \frac{734}{9-1} = 91.75$ $\qquad$ $S_2^2 = \frac{900.9}{10-1} = 100.1$

Let us take the hypothesis that the two population are the same varience ∴ Applying F

$F = \frac{S_1^2}{S_2^2} = \frac{91.75}{100.1} = 0.9165$

Change which is larger than

v_1 = degree of freedom for sample having larger varience $\rightarrow 11-1 = 10$
v_2 = degree of freedom for sample having smaller varience $\rightarrow 9-1 = 8$

$0.7065 < 3.36$

the applied hypothesis is less than the hypothesis

∴ the hypothesis is accepted.

Page No.
Date

Q Suppose you are working in research company and want to the level of carbon oxide emission happening from 2 different brands of cigarettes and wheather they are significantly different or not. In your analysis, you have collected the following informations

	XYZ	ABC
Sample size	11	10
Mean	16.4	15.6
Standard deviation	1.2	1.1

level of signifiance (alpha)	0.05

we know $S.D = \sqrt{var}$

$var = S.D^2$

$\therefore var\ s_1^2 = (1.2)^2$ and $var\ s_2^2 = (1.1)^2$

$s_1^2 = 1.44$ $\quad$ $s_2^2 = 1.21$

$F = \frac{1.44}{1.21} = 1.19 < 3.137$

↑ check in table at 5% LOS

The hypothesis is accepted.

Page No.
Date

Q4 Kruskal-Wallis Test (H-Test)

$$H = \frac{12}{N(N+1)} \left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} \dots \frac{R_k^2}{n_k}\right] - 3(N+1)$$

Q Company ....

method A : ~~80~~ ~~83~~ ~~79~~ ~~85~~ ~~90~~ ~~68~~

method B : ~~82~~ ~~84~~ ~~60~~ ~~72~~ ~~86~~ ~~67~~ ~~91~~

method C : ~~93~~ ~~65~~ ~~77~~ ~~78~~ ~~88~~

Values	Ranks	Rank of sample A	Rank of Sample B	Rank of Sample C
60	1	9	10	18
65	2	11	12	2
67	3	8	1	6
68	4	13	5	7
72	5	16	14	15
77	6	4	3	
78	7		17	
79	8	$\Sigma 61$	$\Sigma 62$	$\Sigma 48$
80	9	$R_1 =$	$R_2 =$	$R_3 =$
82	10			
83	11			
84	12			
85	13			
86	14			
88	15			
90	16			
91	17			
93	18			

Page No.
Date

$$H = \frac{12}{N(N+1)}\left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} + \cdots + \frac{R_k^2}{n_k}\right] - 3(N+1)$$

$$= \frac{12}{18\times(18+1)}\left[\frac{(61)^2}{6} + \frac{(62)^2}{7} + \frac{(48)^2}{5}\right] - 3(18+1)$$

$$= \frac{12}{342}\left[\frac{3721}{6} + \frac{3844}{7} + \frac{2304}{5}\right] - 57$$

$$= 0.03508\left[620.166 + 549.142 + 460.8\right] - 57$$

$$= 57.18418 - 57$$

$$= \underline{\underline{0.18418}}$$

$v = n-1$
$v = 3-1$
$= 2$

$\chi^2_{0.05}$

$\nearrow$ Chi-square

$$\underline{\chi^2 = 5.991}$$

$0.18418 < 5.991$

∴ The null hypothesis is accepted and conclude that three method equally affected

$$\therefore \mu_A = \mu_B = \mu_C$$

Page No.
Date

Q In a manfacturing unit, four teams of operators were randomly selected and sent to four different facilities for machining technique. After the training, the supervisor conducted the exam and recorded the test scores. At 95% confidence level does

Facility 1	Facility 2	Facility 3	Facility 4
88	77	71	52
82	76	56	65
86	84	64	68
87	59	51	81

Value	Rank	Rank of Facility 1	Rank of F 2	Rank of F 3	Rank of F 4
51	1	16	10	8	2
52	2	12	9	3	6
56	3	14	13	5	7
59	4	15	4	1	11
64	5				
65	6	$\Sigma R_1 = 57$	$\Sigma R_2 = 36$	$\Sigma R_3 = 17$	$\Sigma R_4 = 26$
68	7				
71	8				
76	9				
77	10				
81	11				
82	12				
84	13				
86	14				
87	15				
88	16				

$$H = \frac{12}{N(N+1)}\left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} + \cdots\right] - 3(N+1)$$

$$= \frac{12}{16(17)}\left[\frac{(57)^2}{4} + \frac{(36)^2}{4} + \frac{(17)^2}{4} + \frac{(26)^2}{4}\right] - 3(17)$$

$$= 0.0441\left[\frac{(57)^2}{4} + \frac{36^2}{4} + \frac{17^2}{4} + \frac{26^2}{4}\right] - 3\times 17$$

$$= 60.77205 - 51$$

$$= 9.77205$$

$v = 4-1$

$v = 3$

$\chi^2_{0.05}$

Chi square

$\chi^2 = 9.488$

$9.77205 > 9.488$

$\therefore$ The hypothesis is rejected.

Q5 Friedman's Test :

Q Department of public health and safety monitors the measures taken to cleanup drinking water were effective. Trihalomethanes (THMs) at 12 countries drinking water compared before cleanup, 1 week later and 2 weeks after cleanup.

County	Trihalomethanes (THMs)		
	Before cleanup	week 1	week 2
1	21.1 (3)	19.2 (2)	18.4 (1)
2	24.1 (3)	22.3 (2)	21.2 (1)
3	14.1 (3)	12.9 (1.5)	12.9 (1.5)
4	18.1 (3)	17.8 (2)	17.3 (1)
5	15.4 (3)	15.1 (2)	14.9 (1)
6	16.2 (3)	15.1 (1.5)	15.1 (1.5)
7	7.4 (3)	7.2 (2)	6.8 (1)
8	7.5 (3)	6.7 (2)	6.1 (1)
9	14.2 (3)	13.6 (2)	13.1 (1)
10	21.3 (3)	20.9 (2)	20.4 (1)
11	9.5 (2)	9.8 (3)	9.2 (1)
12	11.9 (3)	10.5 (2)	10.1 (1)

$\Sigma R_1 = 35$ $\quad \Sigma R_2 = 24$ $\quad \Sigma R_3 = 13$

$$Q = \frac{12}{n k(k+1)} \sum_{j=1}^{k} R_j^2 - 3n(k+1)$$

$$= \frac{12}{12 \times 3(3+1)} \sum_{j=1}^{3} (35^2 + 24^2 + 13^2) - 3 \times 12(3+1)$$

$$= \frac{1}{12} \times (1970) - 36 \times 4$$

$$= \frac{1970}{12} - 144$$

$$= 164.166 - 144$$

$$= \underline{\underline{20.1666}}$$

Friedman table

$\xrightarrow{k}$ $K = 3$

$\downarrow$ $n-1$

$\therefore$ $\xrightarrow{3}$

$\downarrow$ 11 $= \underline{\underline{6.545}}$

$\therefore 20.16 > 6.545$

$Q_{cal} > Q_{table}$

The calculated Q value is greater than critical value of Q

$\therefore$ hence rejected the null hypothesis.

Q 7 random people

	Drug(A)	Drug(B)	Drug(C)
1	1.24 (1)	1.50 (2)	1.62 (3)
2	1.71 (1)	1.85 (2)	2.05 (3)
3	1.37 (1)	2.12 (3)	1.68 (2)
4	2.53 (2)	1.87 (1)	2.62 (3)
5	1.23 (1)	1.34 (2)	1.51 (3)
6	1.94 (1)	2.33 (2)	2.86 (3)
7	1.72 (2)	1.43 (1)	2.56 (3)
	$\Sigma R_1 = 9$	$\Sigma R_2 = 13$	$\Sigma R_3 = 20$

$$Q = \frac{12}{nk(k+1)} \sum_{j=1}^{k} (R_j)^2 - 3n(k+1)$$

$$= \frac{12}{7 \times 3(3+1)} \left[(9)^2 + (13)^2 + (20)^2 \right] - 3 \times 7(3+1)$$

$$= \frac{12}{84} (650) - 84$$

$$= 92.8557 - 84$$

$$= 8.8571$$

Friedman's table

$\longrightarrow k$; $n-1 \downarrow$

$\longrightarrow 3$; $6 \downarrow$

7.000

$$8.8571 > 7.000$$

The calculated Q value is greater than table value

∴ The null is rejected.

Q6 Non parametric methods (theory)

→ A Non-parametric test is a statistical test that is used when the population data not belong to a parametrized distribution. They are also called distribution-free tests.

Non-parametric tests are mathematical methods that are used in statistical hypothesis testing. They are done by:

1. Determining a critical value of U
2. Rejecting H0 in favor of H1 if the observed value of U is less than or equal to the critical value.
3. Not rejecting H0 if the observed value of U exceeds the critical value.

Page No.
Date

Q7 One way ANOVA (sum)
→

A	B	C	D
8	12	18	13
10	11	12	9
12	9	16	12
8	14	6	16
7	4	8	15
Σ 45	Σ 50	Σ 60	Σ 65
$\bar{A} = 9$	$\bar{B} = 10$	$\bar{C} = 12$	$\bar{D} = 13$

Mean of Mean's $= \frac{9+10+12+13}{4} = \frac{44}{4} = 11$

A	$(\bar{x}_1 - \bar{\bar{x}})^2$	B		C		D	
$(9-11)^2$	4	$(10-11)^2$	1	$(12-11)^2$	1	$(13-11)^2$	4
$(9-11)^2$	4	$(10-11)^2$	1	$(12-11)^2$	1	$(13-11)^2$	4
$(9-11)^2$	4	$(10-11)^2$	1	$(12-11)^2$	1	$(13-11)^2$	4
$(9-11)^2$	4	$(10-11)^2$	1	$(12-11)^2$	1	$(13-11)^2$	4
$(9-11)^2$	4	$(10-11)^2$	1	$(12-11)^2$	1	$(13-11)^2$	4
	↓		↓		↓		↓
	$\Sigma = 20$		$\Sigma = 5$		$\Sigma = 5$		$\Sigma = 20$

$= 20 + 5 + 5 + 20$

SSC $= 50$

Between Samples

Page No.
Date

SSE

Within Samples

A	B	C	D
$(x-\bar{x}_1)^2$	$(x-\bar{x}_2)^2$	$(x-\bar{x}_3)^2$	$(x-\bar{x}_4)^2$
1	4	36	0
1	1	0	16
9	1	16	1
1	16	36	9
4	36	16	4
$\Sigma = 16$	$\Sigma = 58$	$\Sigma = 104$	$\Sigma = 30$

$SSE = 16 + 58 + 104 + 30$

$= 208$

$SST = SSC + SSE$

$= 258$

Sources of variation	Sum of squares	Degree of freedom	Mean Square	Variance Ratio F
Between Sample	50	$v_1 = 4-1$ $= 3$	$MSC = \frac{SSC}{c-1}$ $= 16.7$	$F = \frac{MSC}{MSE} = \frac{16.7}{13}$ $= 1.28$
With Sample	208	$v_2 = 20-4$ $= 16$	$MSE = \frac{SSE}{(n-c)}$ $= 13$	
Total	258	19		