The manufacturer of certain make of electric bulbs claims that his bulbs have a mean life of 25 months with a standard deviation of 5 months. A random sample of six such bulbs gave the following values

Life of months :	$x-\bar{x}$	$(x-\bar{x})^2$
24	1	1
26	3	9
30	7	49
20	-3	9
20	-3	9
18	-5	25

~~Σ 102~~ $\Sigma(x-\bar{x})^2 = 102$

can you Regard the Producer's claim to be valid at 1% level of significance.

$\bar{x} = 23$

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{102}{5}} = \sqrt{20.4} = 4.516$$

$$t = \frac{|\bar{x}-\mu|\sqrt{n}}{s}$$

$$t = \frac{|23-25|\sqrt{6}}{4.516}$$

$$t = \frac{2\sqrt{6}}{4.516}$$

$$t = \frac{4.8989}{4.516}$$

$$t = 1.084$$

Calculated $t = 1.084$

Degree of freedom $= n-1 = 6-1 = 5$

Page No.
Date

Degree of freedom = 5 (Column wise)
level of significance 1%. (Row wise)

4.032

t cal < t table
they we Accept the hypothesis

Let us take the hypothesis that their is no significant difference in the mean life of bulbs in the sample and that of the population

Conclusion: calculated t value: less than table t value
∴ Hypothesis is accepted
Therefore the producer's claim is not valid at 1% level of significance

Page No.
Date

$\bar{x} =$

2. Random Sample Size 16 has a mean of 53 the sum of the squares of the deviations taken from mean is 135. $\Sigma(x-\bar{x})^2 = 135$
Can this sample be regarded as taken from the Population having a mean of 56 $\mu = 56$.
Obtained 95% and 99% confidence limits of the mean of the population

5% LOS   1% LOS

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{135}{16-1}} = \sqrt{\frac{9\,135}{15}} = 3$$

$S = 3$

$$t = \frac{|\bar{x} - \mu|}{S}\sqrt{n}$$

$$t = \frac{|53 - 56|}{3}\sqrt{16}$$

$$t = \frac{3}{3}\sqrt{16}$$

$t_{cal} = 4$ DOF = 15
LOS = 5% or 0.05

t tab = 2.131

DOF, LOS = 1% or 0.01
∴ t tab = 2.947

t cal > t tab
∴ Reject the Hypothesis
(Not accept)

Page No.
Date

For 95%

$53 + \frac{3}{\sqrt{16}} 2.131 \qquad 53 - \frac{3}{\sqrt{16}} 2.131$

$53 + \frac{3}{4} 2.131 \qquad 53 - \frac{3}{4} 2.131$

$54.59825 \qquad 51.40175$

$\boxed{51.5 \text{ to } 52}$

For 99%

$53 + \frac{3}{\sqrt{16}} 2.947 \qquad 53 - \frac{3}{\sqrt{16}} 2.947$

$53 + \frac{3}{4} 2.947 \qquad 53 - \frac{3}{4} 2.947$

$55.21025 \qquad 50.78975$

$\boxed{50.8 \text{ to } 55.2}$

Chi-Square Test. (χ^2 test)
Web Testing often goes beyond A B Testing and tests multiple treatment at once
The Chi Square Test is used with count data to test how well it fits some expected distribution

In an anti malarial campaign in a certain area Quinine was administered to 1624 persons out of total population of 6496 The number of ~~fever~~ fever cases is shown below:

Treatment	Fever	No Fever	Total
Quinine	40	1584	1624
No Quinine	440	4432	4872
Total	480	6016	6496

Discuss the usefulness of Quinine in checking malaria

Let us take the hypothesis that Quinine is not effective in checking malaria.

To determine the value of χ^2 Steps:

1. Calculated the Expected Frequencies

$$E = \frac{RT \times CT}{N}$$

RT is Row Total for Row containing the cell
CT is Column Total for Column containing the cell
N is the total no. of observations

2. Take the difference between the observed and Expected frequencies and obtained the square of differences

3. Divide the values of observed - Expected squares obtained in step 2 by the respected expected frequency and obtain the total. This gives the values of chi square which can range from 0 to ∞

If $\chi^2 = 0$ it means that observed and Expected frequencies completely coincides

The greater the discrepancy b/w observed and expected values the greater will be value of χ^2

Conclusion :

The calculated value of χ^2 is compared with the table value. for the given degrees of freedom at a certain Level of Significance. if at a stated level (5%) the calculated value of the χ^2 is more than the table value of χ^2 the difference b/w expected and observation is considered to be significant. on the other hand, calculated value of χ^2 is less than the table value the difference b/w expected and observation is not considered as significant.

$$E = \frac{1624 \times 480}{6496}$$

$E = 120$

$$\chi^2 \text{ test } \quad \Sigma\left[\frac{(O-E)^2}{E}\right]$$

$$E = \frac{6016 \times 4872}{6496}$$

$E = 4512$

$DOF = (r-1)(c-1)$ don't include total

$(2-1)(2-1)$

$\boxed{DOF = 1}$

Page No.			
Date			

$$E = \frac{480 \times 4872}{6496}$$

$E = 360$

$$E = \frac{6016 \times 1624}{6496}$$

$E = 1504$,

Expected Frequency

120	1504	1624
360	4512	4872
480	6016	6496

O	E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
40	120	6400	53.33
440	360	6400	17.77
1584	1504	6400	4.255
4432	4512	6400	1.4184
			76.7734

LOS = 5% DOF = 1

$\chi^2 = 3.841$

$\chi^2 cal > \chi^2 tab$

Hypothesis is rejected.

It is significant.

From the data given below about the treatment of 250 patients suffering from the disease State whether the new treatment is superior to the conventional treatment

Treatment	Favourable	Non Favourable	Total
New Treatment	140	30	170
Conventional Treatment	60	20	80
Total	200	50	250

Hypothesis: Their is no significant difference b/w the new and conventional treatment.

$$E = \frac{170 \times 200}{250} = 136$$

$$E = \frac{170 \times 50}{250} = 34$$

$$E = \frac{200 \times 80}{250} = 64$$

$$E = \frac{50 \times 80}{250} = 16$$

O	E	$(O-E)$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
140	136	4	16	0.1176
30	34	-4	16	0.4705
60	64	-4	16	0.25
20	16	4	16	1

$$\Sigma \frac{(O-E)^2}{E} = 1.8381$$

$$DOF = (r-1)(c-1)$$
$$(2-1)(2-1)$$
$$DOF = 1$$

$LOS = 5\%$

$\chi^2_{tab} = 3.841$

$\chi^2_{cal} < \chi^2_{tab}$

Hypothesis is Accepted

2m IA-2

on the other end If the calculated value is less than the table value the difference is not significant

One Way Classification Model

Sources of variation	Sum of squares	Degree of freedom (v)	mean square	Variance Ratio F
Between Samples	SSC	$v_1 = c-1$	$MSC = \frac{SSC}{(c-1)}$	$F = \frac{MSC}{MSE}$
within samples	SSE	$v_2 = n-c$	$MS = \frac{SSE}{(n-c)}$	
Total	SST	$n-1$		

Total

SST = Sum of squares of variations

SSC = Sum of Squares between Samples (columns)

SSE = Sum of squares within samples (rows)

MSC = mean sum of squares between samples

MSE = mean sum of squares within samples

To assess the significance of possible variation in performance in a certain test between the convent schools of a city, a common test was given to a number of students taken at random from the senior 5^{th} class of each of the 4 schools concerned. The test results are given below make an analysis of variance of data

A	B	C	D
8	12	18	13
10	11	12	9
12	9	16	12
8	14	6	16
7	4	8	15
45	50	60	65
Avg = 9	Avg = 10	Avg = 12	Avg = 13

$$\frac{45 + 50 + 60 + 65}{4}$$

55

$$\frac{9+10+12+13}{4} = \frac{44}{4} = 11$$

A		B		C		D	
$(\bar{x}_i - \bar{x})^2$							
$(9-11)^2$	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
	20		5		5		20

$= 20+5+5+20$

SSC = 50.

Between Samples

SSE

within samples

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

16+58+104+30

SSE = 208

One way classification model

Sources of variation	Sum of squares	Degree of freedom (v)	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
within samples	208	$v_2 = 20-4$ = 16	$MSE = \frac{SSE}{(n-c)}$ = 13	
Total.	258	19		

3.2389

n numerator dof = 3 row

m denominator dof = 16. Column

1.28 < 3.24

5% LOS

Conclusion

The calculated value of F is less than the table value of F and hence the difference in the mean value ~~is not~~ of the samples is not significant i.e the samples could have come from the same universe

~~Shortcut~~ method :

The 3 samples below have been obtained from normal populations with equal variances. Test the hypothesis that the sample means are equal

	Sample 1	$(S_1)^2$	Sample 2	$(S_2)^2$	Sample 3	$(S_3)^2$
	8	64	7	49	12	144
	10	100	5	25	9	81
	7	49	10	100	13	169
	14	196	9	81	12	144
	11	121	9	81	14	196
Σ	50	530	40	336	60	734

The table value of F at 5% LOS for $v_1 = 2$ and $v_2 = 12$ is 3.88

Correction Factor $= \frac{T^2}{N}$ $\frac{(150)^2}{15} = \frac{\cancel{22500}^{1500}}{\cancel{15}_{1}} = 1500$

530 + 336 + 734 = 1600

SST = 1600 − 1500

SST = 100

Page No
Date

SSC
between samples

$\frac{(50)^2}{5} + \frac{(40)^2}{5} + \frac{(60)^2}{5} - \frac{(T)^2}{N}$

$\frac{2500}{5} + \frac{1600}{5} + \frac{3600}{5} - 1500$

$500 + 320 + 720 - 1500$

$1540 - 1500$

$= 40$

SSC

SSE $= 100 - 40$

$= 60.$

Sources of variation	Sum of Squares	Degree of freedom (v)	Mean Square	Variance Ratio F
Between sample	40	$v_1 = 3-1 = 2$	$MSC = \frac{SSC}{(c-1)}$ 20	$F = \frac{MSC}{MSE}$
within Samples	60	$v_2 = n-c$ $15-3$ $=12$	5	$\frac{20}{5} = 4$
Total	100			

Page No.
Date

observations	A	$(S_1)^2$	B	$(S_2)^2$	C	$(S_3)^2$
1	25	625	31	961	24	576
2	30	900	39	1521	30	900
3	36	1296	38	1444	28	784
4	38	1444	42	1764	25	625
5	31	961	35	1225	28	784
$\Sigma =$	160	5226	185	6915	135	3669

$160 + 185 + 135$

Correction Factor : $\frac{T^2}{N}$ $\frac{(480)^2}{15}$ $\frac{230400}{15} = 15360$

~~SST = 961 + 1225 + 784~~

$5226 + 6915 + 3669$

SST $= 15810 - CF$

$15810 - 15360$

SST $= 450$

~~SSC~~

~~between samples:~~

~~$\frac{(31)^2}{5} + \frac{(35)^2}{5} + \frac{(28)^2}{5} - \frac{(T)^2}{N}$~~

~~$\frac{961}{5} + \frac{1225}{5} + \frac{784}{5} - 15360$~~

$SSC = \frac{(160)^2}{5} + \frac{(185)^2}{5} + \frac{(135)^2}{5} - 15360$

SSC = 250

SSE = SST − SSC

450 − 250

200

Sources of variation	Sum of squares	Degree of freedom	mean square	Variance Ratio F
Between samples	250	$v_1 = c-1$ $3-1=2$	$MSC = \frac{SSC}{c-1}$ $\frac{250}{2} = 125$	$F = \frac{MSC}{MSE}$
With in Samples	200	$v_2 = n-c$ $= 15-3$ $= 12$	$\frac{200}{12} = 16.67$	$\frac{125}{16.67}$ 7.498
Total	450			

table = 3.88

Calculated value > table value

Sample mean is Significant

Two way ANOVA Table

Source of variation	Sum of Square	Degree of freedom (v)	mean Squares	Variance Ratio F
Between sample	SSC	$(c-1)$	$MSC = \frac{SSC}{(c-1)}$	$\frac{MSC}{MSE}$
Between Rows	SSR	$(r-1)$	$MSR = \frac{SSR}{(r-1)}$	$\frac{MSR}{MSE}$
Residual or Error	SSE	$(c-1)(r-1)$	$MSE = \frac{SSE}{(c-1)(r-1)}$	
Total	SST	$n-1$		

SSC Sum of Squares between Columns

SSR Sum of squares between rows

SSE Sum of squares due to error

SST. Total sum of squares

dof = 4-1

A Tea company appoints 4 salesman A, B, C, D and observes their sales in 3 seasons Summer, winter, monsoon. The Figures (in lakhs) are given in the following table

Salesmen

Seasons	A	B	C	D	Season Total
Summer	36	36	21	36	129
winter	28	29	31	31	119
monsoon	26	28	29	29	112
Σ	90	93	81	96	360

1) Do the Salesman significantly differ in the performance
2) Is their significance between the seasons

	A	$(S_1)^2$	B	$(S_2)^2$	C	$(S_3)^2$	D	$(S_4)^2$	Season Total
Summer	6	36	6	36	-9	81	6	36	9
winter	-2	4	-1	1	1	1	1	1	-1
monsoon	-4	16	-2	4	-1	1	-1	1	-8
Σ	0	56	3	41	-9	83	6	38	0

SST = 56+41+83+38

SST = 218

SSC

Sum of Squares between Samples

A B C

$$\frac{(0)^2}{3} + \frac{(3)^2}{3} + \frac{(-9)^2}{3} + \frac{(6)^2}{3} - 0$$

$$0 + 3 + \frac{81}{3} + \frac{36}{3}$$

$$3 + 27 + 12$$

SSC = 42

~~SSE = 218-42~~

Sum of Squares b/w Seasons rows $= \frac{(9)^2}{4} + \frac{(-1)^2}{4} + \frac{(-8)^2}{4}$

$$\frac{81}{4} + \frac{1}{4} + \frac{64}{4}$$

$$\frac{81+1+64}{4}$$

$$\frac{146}{4}$$

SSR = 36.5

SST = 218 SST − (SSR + SSC)

218 − (36.5 + 42)

SSE = 139.5

Source of variation	Sum of Squares	Degree of Freedom(v)	Mean squares	Variance Ratio F
Between Sample	42	4-1 = 3	MSC = SSC/(C-1); 42/3 = 14	MSC/MSE = 14/23.25 = 0.6021
Within Sample	36.5	3-1 = 2	MSR = SSR/(r-1); 36.5/2 = 18.25	MSR/MSE = 18.25/23.25 = 0.7849
Residual or Error	139.5	(3)(2) = 6	MSE = SSE/((r-1)(c-1)); 139.5/6 = 23.25	
Total	218	12-1 = 11		

Conclusion 1:
For the table value of F for $v_1 = 3$ and $v_2 = 6$ at 5% LOS

Calculated value < Table value
and we conclude that sales of different salesman do not differ significantly.
0.60 < 4.76

Conclusion 2
The table value of F for $v_1 = 2$ and $v_2 = 6$ at 5% LOS is 5.14
calculated value < Table value
Hence their is no significant difference as far as the sales are concerned.
0.78 < 5.14

F test:

F test → two random samples were drawn from two normal population and their values are

A	$(x-\bar{x})$	$(x-\bar{x})^2$	B	$(x-\bar{x}_2)$	$(x-\bar{x}_2)^2$
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
			97	14	196
Σ 720 Avg = 80	0	Σ=734	913 Avg = 83		Σ=1298

Test whether the two populations have the same variance at 5% LOS

F = 3.36 at 5% for $v_1 = 10$ $v_2 = 8$

$$\boxed{F = \frac{S_1^2}{S_2^2}}$$

> Reject Hypothesis

$S_1^2 = \frac{734}{n_1 - 1}$

$\frac{734}{9-1} = \frac{734}{8} = 91.75.$

$S_2^2 = \frac{1298}{10} = 129.8$

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

For $v_1 = 10, v_2 = 8$

$F_{0.05} = 3.36$

The calculated value of F is less than the table value the hypothesis is accepted. Hence it may be calculated that the two populations have same variance

F test

Two random samples were drawn from 2 normal populations and their values are given below. Test whether the two population have the same variance at 5% LOS

A	$(x-\bar{x})$	B	$(x-\bar{x})^2$	$(y-\bar{y})$	$(y-\bar{y})^2$
16	-14	14	196	-19	361
17	-13	16	169	-17	289
25	-5	24	25	-9	81
26	-4	28	16	-5	25
32	2	32	4	-1	1
34	4	35	16	2	4
38	8	37	64	4	16
40	10	42	100	9	81
42	12	43	144	10	100
		45		12	144
		47		14	196
Σ 270		Σ 363	734		1298
Avg = 30		Avg = 33			

Let us take the Hypothesis that the 2 populations have the same variance. Applying F test, $F = \frac{S_1^2}{S_2^2}$

$F = \frac{S_1^2}{S_2^2}$

$$s_1^2 = \frac{\Sigma(x-\bar{x})^2}{n-1}$$

$$s_1^2 = \frac{734}{9-1}$$ $$s_2^2 = \frac{1298}{11-1}$$

$$\frac{734}{}$$ $$s_2^2 = 129.8$$

$$s_1^2 = 91.75$$

$$F = \frac{91.75}{129.8}$$

$F = 0.706856$
calculated

$v_1 = 10$ $v_2 = 8$ $F_{0.05} = 3.3$

v_1 = degrees of freedom for sample having larger variance $v_1 = n-1 = 11-1 = 10$

v_2 = degrees of freedom for sample having smaller variance $v_2 = n-1$ $9-1 = 8$.

$0.7068 < 3.34$

$m = v_1$

$n = v_2$

Conclusion:

The calculated value of F is less than the table value therefore the hypothesis is accepted. Hence it may be concluded that the 2 populations have the same variance.

F-Test

Suppose that you are working in a research company and want to know the level of carbon oxide emission happening from two different brands of cigarettes and whether they are significantly different or not. In your analysis, you have collected the following information 5% LOS

	xyz	ABC
sample size	11	10
mean	15.4	15.6
Standard deviation	1.2	1.1

Calculate the variance for xyz and ABC

$$s_1 = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$s_1^2 = (1.2)^2$$
$$s_1^2 = 1.44$$

$$s_2^2 = (1.1)^2$$
$$s_2^2 = 1.21$$

$$F = \frac{s_1^2}{s_2^2} = \frac{1.44}{1.21} \quad 1.190$$

$1.190 < 3.1373$

$$s = \sqrt{variance}$$

Hypothesis is accepted.

LOS 5%. $v_1 = 11-1 = 10$

$v_2 = 10-1 = 9$

3.1373

Both cigarettes companies are producing same carbon emission

S_1^2 depend for dof v_1, v_2
S_2^2.

Page No.
Date

Kruskal - Wallis Test : / H Test :
(2 or more samples)

If several independent samples are involved, analysis of variance is the usual procedure. Failure to meet the assumptions needed for the analysis of variance makes its value doubtful. An alternative technique was developed called the Kruskal-Wallis Test (one way Analysis of Variance) or H Test. This test helps in testing the null hypothesis that k independent random samples come from identical population against the alternative hypothesis that the mean of the samples are not all equal

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

H - statistic
t - statistic
F statistic

N = Total of all the values

Page No.
Date

Company's trainees are randomly assigned to groups which are taught a certain industrial inspection procedures by 3 different methods, at the end of instructing period they are tested for inspection performance quality. The following are there scores

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

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

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

Use the H Test to determine at the 5% LOS whether the 3 methods are equally effective.

Values	Ranks	R_1 Rank of Sample A	R_2 Rank of Sample B	R_3 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	Σ 61	62	48
80	9			
82	10			
83	11			
84	12			
85	13			
86	14			
88	15			
90	16			
91	17			
93	18			

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

$$\frac{12}{18(19)}\left[\frac{3721}{6} + \frac{3844}{7} + \frac{2304}{5}\right] - 57$$

$$\frac{12}{342}\left[1630.1095\right] - 57$$

$$0.0350\left(1630.1095\right) - 57$$

$$57.0538 - 57$$

$$0.0197$$

dof = n-1 = 3-1 = 2

$\chi^2_{0.05}$ V = 2

table 5.991

$\chi^2_{cal} < \chi^2_{table}$

Conclusion:

The calculated value is less than the table value the NULL hypothesis is accepted and we conclude that the 3 methods are equally effective

$\mu_A = \mu_B = \mu_C$

The manufacturing unit 4 teams of operators were randomly selected and send to 4 different facilities for machining techniques training. After the training the supervisor conducted the exam and recorded the test scores. At 95 % confidence level does the scores are same in all 4 facilities.

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

Values	Ranks	Rank of Facility1	Rank of Facility2	Rank of Facility3	Rank of Facility4
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	57	36	17	26
68	7				
71	8				
76	9				
77	10				
81	11				
82	12				
84	13				
86	14				
87	15				
88	16				

$$H = \frac{12}{16(17)}\left[\frac{(52)^2}{4} + \frac{(36)^2}{4} + \frac{(17)^2}{4} + \frac{(26)^2}{4}\right] - 3(17)$$

$H_{cal} = 9.7720$

dof $= n-1$
$4-1 = 3$

$\chi^2 tab = 7.815$

$\chi^2 cal > \chi^2 table$

Reject the null hypothesis

Friedman's Test:

Department of public health and safety monitors the measures taken to clean up drinking water were effective ~~for~~ Trihalomethane (THS) at 12 counties. Drinking water was compared before clean up one week later and 2 weeks after clean up.

County	Before Clean up	1 week Later	2 weeks later
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)
	35	24	13

Let the null hypothesis be h_0: The clean up system had no effect on the THM's

Alternate h_a: the clean up had effected the THMS

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

Kruskal wallis Test.

$$Q = \frac{12}{nk(k+1)}\left[R_1^2 + R_2^2 + \cdots R_k^2\right] - 3n(k+1)$$

Friedman's Test.

t = co...

n = no. of countries

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

$$Q = \frac{12}{12\cdot 3(3+1)} \left(35^2+24^2+13^2\right) - 3\times 12(4)$$

$$Q = 20.16$$

Friedman table:

6.500

$20.16 > 6.5$

$Q_{cal} > Q_{table}$

Calculated Q value is greater than critical value of Q for 5% LOS

Hence Reject the Null Hypothesis.

So it is concluded the clean up system effected the TIIMS of drinking water.

18. 7 random people were given 3 different drugs and for each person, the reaction time corresponding to the drugs were noted. Test the claim at the 5% significance level that all the 3 drugs have the same probability distribution.

~~Drug A~~ ~~Drug B~~ ~~Drug C~~

~~1~~
~~2~~
~~3~~

	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)	2.86 (3)
Σ	Σ = 9	Σ = 13	Σ = 20

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

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

$$= \frac{12}{21(4)} (81+169+400) - 21(4)$$

$$= \frac{12}{21}$$

$$\frac{3}{21}(450) - 84$$

$$\frac{1350}{21} - 84$$

Q_{cal} : 8.857

$Q_{tab} = 7.143$

Conclusion:

Calculated value of Q is greater than table critical value of Q. Hence Reject the hypothesis. They do not have the same probability distribution.

Page No
Date

2m. Non Parametric Tests

(Created)

Statisticians have devised alternated procedures which can be used to test hypotheses about data which are non normal or for which meaningful sample statistic cannot be calculated since these tests do not depend on the shape of the distribution they are called distribution free tests. these tests do not depend upon the population parameter such as mean and the variance and thus they are called as non parametric test.

Some of the non-parametric tests are

1) Sign Test
2) Kruskal wallis Test. / H test
3) Spearman Rank correlation Test
4) Friedman's Test

Advantages of non parametric tests:

① Non parametric test are distribution free i.e they do not require any assumptions to be made about population following normal or any other distribution

② They are simple to understand and easy to apply when the sample size is small.

③ They do not require lengthy calculations. ~~they are applical~~

4 They are applicable to all types of data Qualitative data in the Rank form as well as in the interval or ratio form.

Page No.
Date

Limitations of non parametric test:

① It should be noted that the statistical method which require no assumptions about the populations from which we are sampling are usually less effecient than the corresponding standard techniques. It is generally true that the more we assume the more we can infer from an sample however at the same time the more we assume the more we have to be cautious about the applicability of the methods and the inferences based on the them. It is disadvantageous to use non-parametric tests when all the assumptions of the classical procedures (t test, z test) can be me. and the data are measured either on interval or ratio scale.

Stem and leaf

weights

77
62
46
55
67
56
89
69
58
76

4	6		
5	5	6	8
6	2	7	9
7	6	7	
8	9		

leaf