Enrollment No......................................

Faculty of Engineering

End Sem Examination May-2024

EN3ES10 Statistical Methods

Programme: B.Tech. Branch/Specialisation: CSBS

Duration: 3 Hrs. **Maximum Marks: 60**

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d. Assume suitable data if necessary. Notations and symbols have their usual meaning.

Q.1	i.	Homogenous subsets in the stratified random sampling are called- (a) Clusters (b) Samples (c) Population (d) Strata	1
	ii.	Which one of the following is non-probability sampling? (a) Snowball sampling (b) Simple random sampling (c) Stratified sampling (d) None of these	1
	iii.	If the values of two variables move in the same direction, the correlation is said to be- (a) Neutral (b) Non corelated (c) Positive (d) Negative	1
	iv.	The value of correlation coefficient ‘r’ lies between- (a) 0 and 1 (b) -1 and 0 (c) -1 and +1 (d) None of these	1
	v.	A single value used to estimate unknown population parameter is called- (a) Point Estimate (b) Correlation (c) Regression (d) None of these	1
	vi.	The null hypothesis is represented by- (a) H_1 (b) H_0 (c) H_p (d) H_n	1
	vii.	Which one of the following is an example of non-parametric test? (a) t-test (b) F-test (c) Mann-Whitney test (d) All of these	1
	viii.	To test the randomness of a data ______ is used. (a) Sign test (b) Run test (c) Median test (d) All of these	1
	ix.	What do ANOVA calculate- (a) T-ratio (b) Z-ratio (c) F-ratio (d) None of these	1
	x.	ANOVA is actually comparison of- (a) Mean (b) Standard deviation (c) Variance (d) All of these	1

Q.2 Attempt any two:

i. A population consist of four members 3, 7, 11, 15. Consider all possible samples of size two which can be drawn without replacement from this population. Find the following- **5**
(a) The mean of sampling distribution of means
(b) The standard deviation of sampling distribution of mean

ii. Explain the following with example: **5**
(a) Sampling with replacement (b) Sampling without replacement

iii. Explain: **5**
(a) Stratified random sampling
(b) Standard error

Q.3 i. Define multiple correlation and partial correlation. **3**

ii. Calculate Spearman's rank correlation coefficient from the following data: **7**

x	24	29	19	14	30	19	27	30	20	28	11
y	37	35	16	36	23	27	19	20	16	11	21

OR iii. The two regression lines are as: **7**

$$2y - x - 50 = 0$$
$$3y - 2x - 10 = 0$$

By using these find (a) mean value of x and y (b) Correlation coefficient between x and y

Q.4 i. Explain the following: **4**
(a) Unbiasedness (b) Type I and Type II error

ii. State and prove Neyman-Pearson Lemma. **6**

OR iii. A random sample $(x_1, x_2, x_3, x_4, x_5)$ of size 5 drawn from a normal population with unknown mean μ. Consider the following estimators to estimate μ **6**

I. $t_1 = \frac{x_1+x_2+x_3+x_4+x_5}{5}$

II. $t_2 = \frac{x_1+x_2}{2} + x_3$

III. $t_3 = \frac{2x_1+x_2+\lambda x_3}{3}$

(a) Find λ such that t_3 is an unbiased estimator of μ.
(b) Check whether t_1 *and* t_2 are unbiased or not?
(c) Which is the most efficient estimator of μ?

Q.5 i. Write the advantages and disadvantages of non-parametric tests as compared to parametric tests. **4**

ii. Explain one sample Wilcoxon sign rank test. **6**

OR iii. Following table shows the drying time (in hours) of paint A on 12 walls of certain area before and after using the additive: **6**

Wall	1	2	3	4	5	6	7	8	9	10	11	12
Before	5.3	4.7	6.2	4.4	5.2	6.7	7.1	7.5	6.0	3.8	4.8	5.5
After	5.5	4.7	6.7	4.6	5.0	7.3	7.7	7.3	4.7	6.2	4.7	5.4

Use sign test at 5 % level of significance to test the null hypothesis that after using additive paint is no better than before using additive in concern of drying time.

Q.6 i. Define time series and write its components. **2**

ii. Following table gives per acre production data for three varieties of wheat, each grown on four plots- **8**

Plot of Land	Variety of Wheat		
	A	B	C
1	6	5	5
2	7	5	4
3	3	3	3
4	8	7	7

Using One-Way ANOVA test if there is any significant difference among varieties of wheat. (Given $F_{0.05}(2,9) = 4.26$).

OR iii. Before making a colony, an experiment was conducted to determine the setting time of different cements with different dusts. The following table shows the time for three different types of cements and five different type of dusts: **8**

Cement	Dust				
	I	II	III	IV	V
I	6	5	6	7	8
II	7	6	5	8	6
III	6	4	6	6	7

Using Two-way ANOVA test whether there is any significant difference among cements or among dusts. (Given $F_{0.05}(2,8) = 4.46$, $F_{0.05}(4,8) = 3.84$).

******

Faculty of Engineering
End sem Examination May-2024
EN3ES10 Statistical Methods.

Q-1 MCQs

(i) (d) strata +1

(ii) (a) Snowball sampling +1

(iii) (c) Positive +1

(iv) (c) −1 and +1 +1

(v) (a) Point Estimate +1

(vi) (b) H_0 +1

(vii) (c) Mann-Whiney test +1

(viii) (b) Run test +1

(ix) (c) F-ratio +1

(x) (a) Mean +1

Q.2

(i) SRSWOR $K = {}^{N}C_{n} = {}^{4}C_{2} = 6$ Samples.

Possible samples are :

S. No.	Sample values	Total	Sample mean
1	3, 7	10	5
2	3, 11	14	7
3	3, 15	18	9
4	7, 11	18	9
5	7, 15	22	11
6	11, 15	26	13

+1

Sampling distribution of mean

+1

Sample mean ($\bar{x}$) :	5	7	9	11	13
Probability (P) :	1/6	1/6	1/3	1/6	1/6

(a) Mean of sampling distn of mean :

$$E(\bar{x}) = \tfrac{1}{6}\times 5 + \tfrac{1}{6}\times 7 + \tfrac{1}{3}\times 9 + \tfrac{1}{6}\times 11 + \tfrac{1}{6}\times 13$$

$$= 5/6 + 7/6 + 3 + 11/6 + 13/6$$ +1

$$\boxed{E(\bar{x}) = 9}$$

(b) Variance of Sampling distⁿ of mean.

$$Var(\bar{x}) = E(\bar{x}^2) - [E(\bar{x})]^2$$

$$= \left(\frac{1}{6}\times 5^2 + \frac{1}{6}\times 7^2 + \frac{1}{3}\times 9^2 + \frac{1}{6}\times 9^2 + \frac{1}{6}\times 11^2 + \frac{1}{6}\times 13^2\right) - 9^2$$

$$= \left(\frac{25}{6} + \frac{49}{6} + 27 + \frac{121}{6} + \frac{169}{6}\right) - 81$$

$$= \frac{263}{3} - 81 \qquad +2$$

$$= \frac{20}{3}$$

$$S.D. = \sqrt{Var(\bar{x})}$$

$$= \sqrt{20/3}$$

Q-2

(ii) (a) Sampling with replacement –

Simple random Sampling is said to be "with replacement" when the sampling unit selected for the sample is returned to the populⁿ before the next one is drawn. Thus the populⁿ remains same before each drawing and any of the sampling unit may appear more than once in the sample. +2.5

(b) Sampling without replacement -

Simple random sampling is said to be without replacement when sampling units are drawn one by one in such a manner that after each drawing the selected units is not returned to the popul^n when the next one is drawn. +2.5

Q-2 (iii)

(a) Stratified Random Sampling

2.5

In stratified random sampling the hetrogenous popul^n is divided into no. of homogenous groups called strata, which are differ from one another but each of these groups is homogenous within itself. Then units are sampled at random from each of these stratum. The sample which is aggregate of sampled units of each of the stratum is called stratified sample and the method is known as stratified random sampling.

(b) Standard Error.

2

The standard deviation of the sampling dist^n of a static is known as standard error. The standard error of some of the statistic are as:

Statistic	S.E.
(1) Sample mean	$\sigma/\sqrt{n}$
(2) Sample S.D.	$\sqrt{\sigma^2/2n}$
(3) Sample var.	$\sigma^2\sqrt{2/n}$

1/2

Utility of S.E. —

1

(1) S.E. plays a very important role in the large sample theory.

(2) It is used to find confidence limit within which parameters are expected to lie.

(3) It is used in testing a given statistical hypothesis at different levels of significance

(4) It gives an idea about reliability of sample.

Q-3

(i) Multiple Correlation

1.5

When the value of one variable are influenced by other variable for Ex - Age & height, supply and demand and so on, Karl Pearson's coefficient of correlation can be used as a measure of linear relationship between them. But sometimes there is interrelation between many variables and the value of one variable may be influenced by many otheres for Ex – yeild of crop depends upon seed, fertility of soil, manure, irrigation, temprature etc.

So, whenever we are interested in studying the joint effect of a group of variables upon a variable not included in that group our study is that of multiple correlation.

If x_1 be the dep. var. and $x_2, x_3, x_4 \ldots x_n$ be the indip. var. the multiple correlation will be given as

$$R_{1(23\ldots n)} \text{ or } R_{1.23\ldots n}$$

(b) Partial correlation – 1.5

When we measure the correlation between two variables where the effect of all other variable have been eliminated then it is called partial correlation.

For Ex – Assuming rainfall (x_3) as constant, the correlation between the yield of wheat (x_1) and the quantity of fertilizer (x_2) is the partial correlation between x_1 and x_2 when the effect of x_3 has been eliminated.

It is denoted by

$$r_{12.34\ldots n}$$

Q-3

(ii)

x	R_x	y	R_y	$D = R_x - R_y$	D^2
24	6	37	1	5	25
29	3	35	3	0	0
19	8.5	16	9.5	-1	1
14	10	36	2	8	64
30	1.5	23	5	-3.5	12.25
19	8.5	27	4	4.5	20.25
27	5	19	8	-3	9
30	1.5	20	7	-5.5	30.25
20	7	16	9.5	-2.5	6.25
28	4	11	11	-7	49
11	11	21	6	5	25

+4

$\Sigma D^2 = 242$

$$r = 1 - \frac{6\left[\Sigma D^2 + \frac{1}{12}(m^3-m) + \frac{1}{12}(m^3-m) + \frac{1}{12}(m^3-m)\right]}{n(n^2-1)}$$

+2

$$= 1 - \frac{6\left[242 + 0.5 + 0.5 + 0.5\right]}{1320}$$

$$= 1 - \frac{6 \times 243.5}{1320}$$

+1

$$= 1 - \frac{1461}{1320}$$

$$= 1 - 1.11$$

$$r = \boxed{-0.11}$$

Q-3

(iii) (a) The mean values of x and y i.e. $(\bar{x}, \bar{y})$ lies on both the lines.

$\therefore \quad 2\bar{y} - \bar{x} = 50$ +2

$3\bar{y} - 2\bar{x} = 10$

Solving these equations

$\bar{x} = 130$ and $\bar{y} = 90$

(b) Let us assume that the line $2y - x = 50$ be regression eqn of y on x. Thus arrange in the form. +2

$y = a + bx$ we get

$2y = x + 50$

$y = 0.5x + 25$

$\therefore \quad b_{yx} = 0.5$

Now, the regression eqn of x on y is the line +2

$3y - 2x = 10$

Arrange it in form $x = a + by$

$2x = 3y - 10$

$x = 1.5y - 5$

$\therefore \quad b_{xy} = 1.5$

Correlation coff $(r) = \sqrt{b_{yx} \cdot b_{xy}}$ +1

$= \sqrt{0.5 \times 1.5}$

$\boxed{r = 0.87}$

Q-4

(i) (a) Unbiasedness.

An estimator $\hat{\theta}$ is said to be an unbiased estimator of the parameter θ if its expected value is equal to θ i.e

$$E(\hat{\theta}) = \theta$$

$$E(\text{Statistic}) = \text{Parameter}$$

This property of an estimator is known as unbiasdness.

If $E(\hat{\theta}) \neq \theta$ then $\hat{\theta}$ is said to be baised estimator of θ.

+2

(b) Type I and Type II Error

Type I error – The decision of rejecting a null hypothesis H_0 when it is true is called type I error. The probability of type I error is called the size of type I error. and it is denoted by α.

+1

$$\begin{aligned}\alpha &= \text{Prob. of type I error} \\ &= \text{Prob. of rejecting } H_0 \text{ when } H_0 \text{ true} \\ &= P(\text{reject } H_0 \mid H_0 \text{ true}) \\ &= P[x \in W \mid H_0] = \int_W L_0\, dx\end{aligned}$$

Type II error – The decision of accepting the null hypothesis H_0. When it is false called type II error. The prob. of type II error is called the size of type II error and it is denoted by β.

+1

β = Prob. of type II error
= Prob. of accepting Ho when Ho is false
= P[Accept Ho / Ho false]
= P[Accept Ho / H_1 true]
= $P[x \in w' / H_1]$
= $\int_{w'} L_1 \, dx$

Q-4 (ii) Neyman - Pearson Lemma.

Statement - Let $x_1, x_2, -- x_n$ be a random sample from the population with p.d.f or p.m.f $f(x, \theta)$. Let $H_0: \theta = \theta_0$ be a simple null hypothesis against a simple alternative hypothesis $H_1: \theta = \theta_1$.

+2

Let L be the liklihood function and L_0 and L_1 be the liklihood function of sample observations under H_1 and H_0.

Let k be the positive constant and w is the critical region of size α such that

$$\alpha = P[x \in w / H_0]$$

(i) $\frac{L_1}{L_0} \geq K$ when $x \in w$ —— (1)

(ii) $\frac{L_1}{L_0} \leq K$ when $x \in w'$ —— (2)

then the critical region w is said to be the best critical region.

Proof :-

We are given

$$P[x \in w / H_0] = \alpha = \int_w L_0 dx$$

$$P[x \in w / H_1] = 1-\beta = \int_w L_1 dx$$

Case - 1

If there is only one such critical region w then obviously it is the best critical region of size α.

We have to prove that there exist no other critical region of size less than or equal to α which is powerful then w.

Case - 2

Let w_1 be the another critical region of size α and power $1-\beta_1$ so that we have

$$P[x \in w_1 / H_0] = \alpha = \int_{w_1} L_0 dx$$

+2

and $$P[x \in w_1 / H_1] = 1-\beta_1 = \int_{w_1} L_1 dx$$

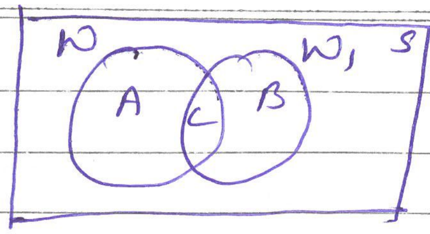
Now, we have to prove that

$$1-\beta \geq 1-\beta_1$$

i.e,

$$P[x \in w / H_1] \geq P[x \in w_1 / H_1]$$

Let $w = A \cup C$ and $w_1 = B \cup C$



$$P[x\in w/H_1] - P[x\in w_1/H_1] = \int_w L_1 dx - \int_{w_1} L_1 dx$$

$$= \int_{A\cup C} L_1 dx - \int_{B\cup C} L_1 dx$$

$$= \int_A L_1 dx + \cancel{\int_C L_1 dx} - \int_B L_1 dx - \cancel{\int_C L_1 dx}$$

$$= \int_A L_1 dx - \int_B L_1 dx \quad \text{--- } (1)$$

Now using condition (1) given in Statement

$$\frac{L_1}{L_0} \geq k.$$

$$\Rightarrow L_1 \geq kL_0 \quad \text{--- } (2)$$

Now using eq^n (2) in eq^n (1) we get

$$P[x\in w/H_1] - P[x\in w_1/H_1] \geq k\int_A L_0 dx - \int_B k L_0 dx$$

Now on adding and subtracting $\int_C k L_0 dx$ on right hand side of the above eq^n

13

$$P[x \in w | H_1] - P[x \in w_1 | H_1] \geq K\left[\int_A L_0 dx + \int_C L_0 dx - \int_B L_0 dx - \int_C L_0 dx\right]$$

$$\geq K\left[\int_{A \cup C} L_0 dx - \int_{B \cup C} L_0 dx\right]$$

$$\geq K\left[\int_{w} L_0 dx - \int_{w_1} L_0 dx\right] \quad +2$$

$$\geq K\left[P[x \in w | H_1] - P[x \in w_1 | H_1]\right]$$

$$\geq K(\alpha - \alpha)$$

$$P[x \in w | H_1] - P[x \in w_1 | H_1] \geq 0$$

$$P[x \in w | H_1] \geq P[x \in w_1 | H_1]$$

$$\boxed{1 - \beta \geq 1 - \beta_1}$$

Q-4
(iii)

$x_i \sim N(\mu, \sigma^2)$

$\therefore E(x_i) = \mu \quad , \quad Var(x_i) = \sigma^2$

(a) (i) $E(t_1) = E\left(\dfrac{x_1 + x_2 + x_3 + x_4 + x_5}{5}\right)$ 1

$$= \frac{1}{5} E\left(\sum_{i=1}^{5} x_i\right)$$

$$= \frac{1}{\cancel{5}} \times \cancel{5}\mu$$

$$E(t_1) = \mu$$

(ii) $$E(t_2) = E\left[\frac{x_1+x_2}{2} + x_3\right]$$

$$= E\left(\frac{x_1+x_2}{2}\right) + E(x_3)$$ 1

$$= \frac{1}{2}\left[E(x_1 + E(x_2)\right] + E(x_3)$$

$$= \frac{1}{2}(\mu+\mu) + \mu$$

$$E(t_2) = 2\mu$$

$$E(t_2) \neq \mu$$

$\therefore$ t_2 is biased estimater of μ.

(b)

(iii) $$E(t_3) = \mu$$

$$E\left(\frac{2x_1+x_2+\lambda x_3}{3}\right) = \mu$$

$$\frac{1}{3}\left[2E(x_1) + E(x_2) + \lambda E(x_3)\right] = \mu$$ 1

$$\frac{1}{3}\left[2\mu + \mu + \lambda\mu\right) = \mu$$

$$\frac{1}{3}(3\mu + \lambda\mu) = \mu$$

$$\frac{\mu}{3}(3+\lambda) = \mu$$

$$3+\lambda = 3$$

$$\boxed{\lambda = 0}$$

(c) $$t_1 = \frac{x_1+x_2+x_3+x_4+x_5}{5}$$

$$Var(t_1) = Var\left(\frac{x_1+x_2+x_3+x_4+x_5}{5}\right)$$

$$= \frac{1}{25}\left[Var(x_1)+Var(x_2)+Var(x_3) + Var(x_4)+Var(x_5)\right]$$

$$= \frac{1}{25}\left[\sigma^2+\sigma^2+\sigma^2+\sigma^2+\sigma^2\right]$$

+1

$$= \frac{5\sigma^2}{25}$$

$$\boxed{Var(t_1) = \frac{\sigma^2}{5}}$$

$$t_2 = \frac{x_1+x_2}{2} + x_3$$

$$Var(t_2) = Var\left[\frac{x_1+x_2}{2} + x_3\right]$$

$$= Var\left[\frac{x_1+x_2}{2}\right] + Var(x_3)$$

$$= \frac{1}{4}\left[Var(x_1)+Var(x_2)\right] + Var(x_3)$$

$$= \frac{1}{4}(\sigma^2+\sigma^2)+\sigma^2$$

+1

$$= \frac{2\sigma^2}{4}+\sigma^2$$

$$= \frac{\sigma^2}{2}+\sigma^2$$

$$\boxed{Var(t_2) = \frac{3\sigma^2}{2}}$$

$$Var(t_3) = Var\left(\frac{2x_1 + x_2}{3}\right)$$

$$= \frac{1}{9}\left[Var(2x_1) + Var(x_2)\right]$$

+1

$$= \frac{1}{9}\left[4\sigma^2 + \sigma^2\right]$$

$$\boxed{Var(t_3) = \frac{5}{9}\sigma^2}$$

Variance of t_1 is least so t_1 is the best estimator.

Q-5

(i) Advantages of non parametric test

(i) Non parametric test are very simple and easy to apply. and do not require complicated sample theory.

+2

(ii) No assumption is made about the form of probability distribution of the parent populn

(iii) Non-parametric test may be very powerful even if the sample size is small.

(iv) They can deal with data which are given in ranks.

Disadvantage of non-parametric test

(i) Non-parametric test can be apply if the measurment are nominal or ordinal. The use of non parametric test with data that can be handled with a parametric test is a wastage of time

+2

(ii) It is not possible to determine the actual power of non parametric test.

(iii) These are designed to test statistical hypothesis only and not for estimation of the parameters.

Q-5

(ii) Wilcoxon Signed Rank test

Let $x_1, x_2, \ldots x_n$ be a random sample of size n from a continuous popu^n which is symetric about unknown median m. Here our problem is to test

$$H_0 : m = m_0$$

+1

against $H_1 : m \neq m_0$

$H_1 : m > m_0$

$H_1 : m < m_0$

Procedure –

One sample wilcoxon signed rank test proceeds as -

(i) Subtract m_0 from each observation and discard all differences equal to zero.

(ii) Find the absolute difference for each of the differences. +1

(iii) Rank the differences as rank 1 to the smallest diff rank 2 to the next smallest and so on. If two or more diff. are same then average of their ranks will be assigned to each

(iv) Compute $Z = \dfrac{R - E(R)}{\sqrt{Var(R)}}$ (+2)

R = Sum of ranks corresponding to positive differences.

$E(R) = \dfrac{n(n+1)}{4}$, $Var(R) = \dfrac{n(n+1)(2n+1)}{24}$ (+2)

(v) If $Z \in CR$ reject H_0 otherwise accept H_0.

Q-5
(iii)

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 < \mu_2$

Wall	Before	After	Sign.
1	5.3	5.5	+
2	4.7	4.7	ignore
3	6.2	6.7	+
4	4.4	4.6	+
5	5.2	5.0	−
6	6.7	7.3	+
7	7.1	7.7	+
8	7.5	7.3	−
9	6.0	4.7	−
10	3.8	6.2	+
11	4.8	4.7	−
12	5.5	5.4	−

(+2)

Total no. of + sign = 6 (R) (+1)
Total no. of − sign = 5
$n = 11$

$$Z = \frac{R - E(R)}{\sqrt{Var(R)}} = \cancel{6}$$

+2

$E(R) = \frac{n}{2} = \frac{11}{2} = 5.5$

$Var(R) = \frac{n}{4} = \frac{11}{4} = 2.75$

$$Z = \frac{6 - 5.5}{\sqrt{2.75}} = \frac{0.5}{1.65} = 0.303$$

+1

$0.303 < 1.96$

$Z \notin CR$ Ho is accepted

Q-6

(i) Time series -

Definition - (i) A time series consists of a data arranged chronologically. +1

(ii) A time series is a set of statistical observations arranged in chronoligeal order.

Components of time series are

1. Secular trend
2. Seasonal Variation +1
3. Cyclical Variation
4. Irregular Variation

Q-6

(ii) $H_0: \mu_1 = \mu_2 = \mu_3$ +1

H_1: At least two of the means are not equal.

(i) Grand total $G = 63$ +1

(ii) Correction factor $= \frac{G^2}{n} = \frac{63^2}{12} = \boxed{330.75}$ +1

(iii) Total Sum of Square TSS = +1

$$(6^2 + 7^2 + 3^2 + \cdots + 7^2) - 330.75$$

$$= \boxed{34.25}$$ +1

(iv) Sum of square between sample (SSC)

$$= \frac{1337}{4} - 330.75 = \boxed{3.50}$$ +1

(v) Sum of square within sample (SSE)

$$= TSS - SSC$$

$$= 34.25 - 3.50 = \boxed{30.75}$$

Anova Table +1

Source of Var.	Sum of Square	d.f	MSS	F-ratio
Between sample	3.50	2	1.75	$F = 0.512$
Within sample	30.75	9	3.417	
Total	34.25	11		

$$0.512 < 4.26$$ +1

∴ Null hypothesis is accepted.

Q-6

(iii) H_{01} : Cements are homogenous (+1)

H_{02} : Dusts are homengous.

H_{11} : Cements are not homogenous.

H_{12} : Dusts are not homogenous

(i) Grand Total $G = 90$

(ii) Correction factor $= \frac{G^2}{n} = \frac{90^2}{15} = 540$ (+1)

(iii) Total Sum of Square (TSS) = (+1)

$$(6^2 + 5^2 + 6^2 + \cdots + 6^2 + 7^2) - 540$$

$$= 560 - 540 = \boxed{20}$$

(iv) Sum of Square between the coloumn (SSC) (+1)

$$= \frac{16^2 + 15^2 + 17^2 + 21^2 + 21^2}{3} - 540$$

$$= \frac{256 + 225 + 289 + 441 + 441}{3} - 540$$

$$\boxed{SSC = 10.67}$$

(v) Sum of Square between the rows (SSR) =

$$= \frac{32^2 + 29^2 + 29^2}{5} - 540$$ (+1)

$$= \frac{1024 + 841 + 841}{5} - 540$$

$$\boxed{SSR = 1.20}$$

Sum of Square within the sample (SSE) =

$SSE = TSS - (SSC + SSR)$

$= 20 - 1.20 - 10.67$

$\boxed{SSE = 8.13}$

ANOVA Table

+2

Sources of Var.	d.f.	Sum of Square	MSS	F-ratio
Between Coloumn	2	1.20	0.60	$F_C = 0.59$
Between Row	4	10.67	2.67	$F_R = 2.62$
within sample	8	8.13	1.02	
Total	14	20		

Since $0.59 < 4.46 = F_{0.05}(2,8)$ $\therefore$ H_{01} is accepted i.e. cements are homogenous. +1

and $2.62 < 3.84 = F_{0.05}(4,8)$ $\therefore$ H_{02} is also accepted i.e. dust are also homogenous.