Total No. of Questions: 6 *Total No. of Printed Pages:3*

Enrollment No.....

Faculty of Engineering

End Sem (Even) Examination May-2022

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.

Q.1 i. The standard deviation of the sampling distribution of any statistic is called: **1**

(a) Sampling Error (b) Type -I Error
(c) Non-Sampling Error (d) Standard Error

ii. The process of drawing a sample from a population is known as_____. **1**

(a) Sampling (b) Census
(c) Survey research (d) None of these

iii. The range of Karl Pearson coefficient of correlation is: **1**

(a) 0 to ∞ (b) -∞ to ∞ (c) 0 to 1 (d) -1 to 1

iv. The two lines of regression pass through- **1**

(a) Mean values of x and y (b) Variance of x
(c) Variance of y (d) None of these

v. Which of the following is a type of estimation? **1**

(a) Interval (b) Biased (c) Unbiased (d) None of these

vi. The process of making estimates about the population on parameter from a sample is called- **1**

(a) Statistical Independence (b) Statistical inferences
(c) Statistical decision (d) None of these

vii. Which non-parametric test relies on the calculation of ranks: **1**

(a) Run test (b) Mann Whitney
(c) Sign test (d) None of these

viii. Which of the following tests would be an example of a non-parametric method? **1**

(a) Z test (b) T- test (c) Sign test (d) None of these

P.T.O.

ix. When calculating ANOVA, F data will always fall within the range- **1**
(a) $(0, \infty)$ (b) (0, 1) (c) $(-\infty, 0)$ (d) None of these

x. Analysis of variance is a statistical method of comparing ..... of the several populations **1**
(a) Means (b) Variances
(c) Standard deviations (d) None of these

Q.2 i. Prove that sample mean is an unbiased estimator of population mean. **2**

ii. Write short note on sampling with replacement and sampling without replacement. **3**

iii. Define stratified random sampling with well define example as well as notations. **5**

OR iv. Define simple random sampling with three examples as well as Notations. **5**

Q.3 i. Define following terms with any one example: **4**
(a) Correlation (b) Regression

ii. Find the value of Correlation Coefficient for the following data: **6**

X:	10	12	15	20	25	30	35	40	45	50
Y:	3	6	9	12	15	18	20	22	24	26

OR iii. Find the Regression Line of Y on X and X on Y for the following data: **6**

X:	3	5	7	9	12	15	18	20
Y:	2	4	8	10	13	17	21	23

Q.4 i. Define the following terms: **4**
(a) Type -I error & Type -II error (b) Critical Region
(c) Null Hypothesis (d) Alternative Hypothesis

ii. State and prove Nyman Pearson Lemma. **6**

OR iii. Explain method of Maximum Likelihood Estimator. Also write any four properties of Maximum Likelihood Estimation. **6**

Q.5 i. Write any four differences between Parametric test and non-Parametric test. **4**

ii. There are 30 male and 20 females are standing in row in the following manner **6**

M F M F M M M F F M F M F M F M M M M F
M F M F M M F F F M F M F M F M M F M M
F M M M M F M F M M.

Test that the males and females are standing in random order. Given that: at 1% level of significance and tabulated value of Z is, $Z_{tab} = 2.58$.

OR iii. Use Wilcoxon signed rank test to test hypothesis that the median length (Θ) of ear head of a variety of wheat is $\theta_0 = 9.9$ cm. against the alternative that $\Theta \neq 9.9$cm, with $\alpha = 0.05$ on the basis of the following 20 ear-head Measurement. **6**
9.3, 8.8, 10.7, 11.5, 8.2, 9.7, 10.3, 8.6, 11.3, 10.7, 11.2, 9.0, 9.8, 9.3, 9.9, 10.3, 10.0, 10.1, 9.6, 10.4.
Given that: tabulated value of $T\alpha = 46$.

Q.6 i. Explain ONE WAY ANOVA with example and also write it's ANOVA TABLE. **4**

ii. The following data give the yield on 12 plots of land in three samples, each of 4 plots, under three varieties of fertilizers A, B and C. **6**

A	B	C
25	20	24
22	17	26
24	16	30
21	19	20

Is there any significant difference in the average yields of land under the three Varieties of fertilizers?
Given that: at 5% level of significance $F_{tab} = 4.26$.

OR iii. To study the performance of three detergents and three different water Temperatures, the following whiteness readings were obtained with specially designed equipment. **6**

Water temperature	Detergent A	Detergent B	Detergent C
Cold water	57	55	67
Warm water	49	52	68
Hot water	59	46	58

Perform a two-way ANOVA using 5% level of significance.
Given that: F_{tab} for temperature is 6.94 and F_{tab} for water is 6.94.

******

B.Tech. (CSBS) 2nd Sem
SM (EN3ES10)
Paper Solution 2022

Q.1 (i) (d) Standard errors
(ii) (a) Sampling
(iii) (d) (-1 to +1)
(iv) (a) Mean value of x and y
(v) (a) Interval
(vi) (b) Statistical inferences
(vii) (b) Mann whitney
(viii) (c) Sign test
(ix) (a) (0 ∞)
(x) (a) Means

Q.2 (i)

We have to prove that $E(\bar{y}_n) = \bar{y}_N$

We know that

Sample mean $\bar{y}_n = \sum_{i=1}^{n} y_i / n$ also

(2)

We can write

$$\bar{y}_n = \frac{1}{n} \sum_{i=1}^{N} a_i y_i \quad \text{---- (i)}$$

We take expectation both side in eq.(i)

$$E(\bar{y}_n) = \frac{1}{n} \sum_{i=1}^{N} E(a_i) \cdot y_i$$

$$E(a_i) = \frac{n}{N}$$

So

$$E(\bar{y}_n) = \frac{1}{n} \sum_{i=1}^{N} \left(\frac{n}{N}\right) \cdot y_i$$

$$= \frac{1}{N} \sum_{i=1}^{N} y_i = \bar{y}_N$$

---

(3)

(ii) Short Note on :

SRSWR : SRSWR means "Simple random Sampling With replacement". if the selected units are being replaced back in the population before the second draw, it is called SRSWR. (1.5)

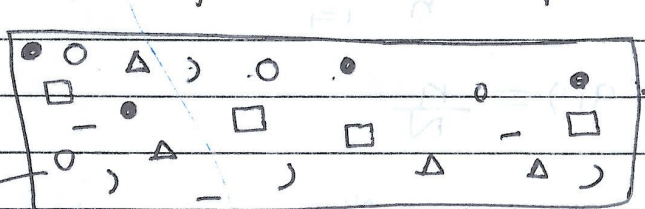
SRSWOR : SRSWOR means "Simple random Sampling without replacement". if the selected units are not being replaced back in the population before the second draw, it is called SRSWOR. (1.5)

(iii) **Stratified random sampling : (5)**

Stratified random sampling is used when your population is devided into strata and you want to include the stratum when taking your sample. The Stratum may be already defined (like census data) or you might make the Stratum yourself to fit the purpose of your research.

In statistics, stratified sampling is a method of Sampling from a population which can be partitioned into Sub-populations. (2)

Example : Vegetables shop



There are many types of vegetables which are mix-up after stratification

Date:
P. No:

S_i^2 = population mean square of the i^{th} Stratum

$$S_i^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (y_{ij} - \bar{y}_{N_i})^2$$

$(i = 1, 2, 3 - k)$

y_{ij} = value of j^{th} sampled unit from i^{th} Stratum.

$\bar{y}_{ni}$ = mean of sample selected from i^{th} Stratum

s_i^2 = Sample mean square of the i^{th} stratum

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{ni})^2 \quad (2)$$

(OR)

(iv) Simple random Sampling: (5)

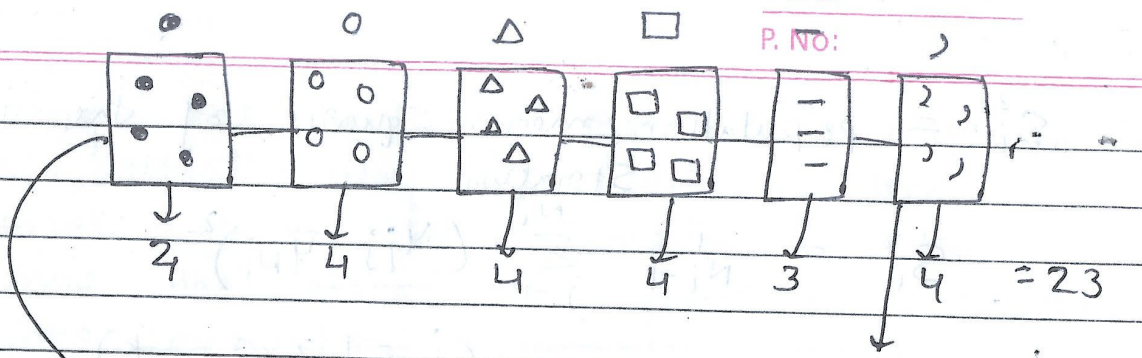
Simple random Sampling is a type of probabili-ty sampling in which the researcher randomly selects a subset of participants from a population. each member of the population has an equal chance of being selected. This method is the most straightforward of all the probability Sampling methods, Since it only involves a single random selection and requires little advance knowledge about the population. Simple random sampling Works best if you have a lot of time and resources to conduct your study, or if you are studying a limited population that can easily be sampled. (2)

example: Method of Lottery

Homogeneous

Date:

P. No:



Strata → 6

Stratum (Individual)

Now Sampling:

Sample Size 8

$n = \{$ ● ○ △ □ — , ● ○ $\}$

(1)

Notation:

K be the number of strata

N_i = The Number of Sampling units in the ith Stratum

$N = \sum_{i=1}^{K} N_i$, Total number of Sampling units in the population.

n_i = The number of sampling units selected with SRSWOR

$n = \sum_{i=1}^{K} n_i$, total Sample Size from all the strata

y_{ij} be the value of the jth unit in the ith Stratum.

$\bar{y}_{N_i}$ = population mean of ith Stratum

$$\bar{y}_{N_i} = \frac{1}{N_i} \sum_{j=1}^{N_i} y_{ij}$$

$\bar{y}_N$ = Population mean = $\sum_i \sum_j y_{ij} / N$

Date:
P. No:

Using the lottery method is one of the oldest ways and is a mechanical example of random sampling. In this method, the researcher gives each member of the population a number. researchers draw numbers from the box randomly to choose samples. (1)

Notations :

N = Total Number of observations in population

y_i = Sample observation

Y_i = population observation

population mean $\bar{Y}_N = \frac{1}{N} \sum_{i=1}^{N} Y_i$

Sample mean $\bar{y}_n = \frac{1}{n} \sum_{i=1}^{n} y_i$

$\bar{y}_n = \frac{1}{n} \sum_{i=1}^{N} a_i Y_i$

S^2 = Mean Square for the population

$S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y}_N)^2$

s^2 = mean square for the sample

$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y}_n)^2$ (2)

Q.3 (i) → (a)

Correlation : (2)

Correlation means association more precisely it is a measure of the extent to which two variables are related. There are three possible results of a correlation study : a positive correlation, a negative correlation and

Date:
P. No:

Correlation.

A positive correlation is a relationship b/w two variables in which both variables move in the same direction. Therefore, when one variable increases as one other variables increases, or one variable decreases while the other decreases. An example of positive correlation would be height and weight. Taller people tend to be heavier. (1)

A Negative correlation is a relationship b/w two variables in which an increase in one variable is associated with a decrease in the other.

An example of negative correlation would be height above sea level and temperature. As you climb the mountain (increase in height) it gets colder (decrease in temperature).

A zero correlation exists when there is no relationship b/w the amount of tea drunk and level of intelligence. (1)

x x

(b) Regression: (2)

Regression is a statistical Method used in finance, investing and other disciplines that attempts to determine the strength and character of the relationship b/w one dependent variable (usually denoted by Y) and a series of other variables (known as independent variables).

The two basic types of regression are simple linear regression and multiple linear regression, although there are non-linear regression methods for more complicated data and analysis.

Simple linear regression uses one independent

Date:
P. No:

variable to explain or predict the outcome of the dependent variable Y, while multiple Linear regression uses two or more independent variables to predict the outcome. (1)

Example:

many businesses use linear regression to forecast how much cash they will have on hand in the future. (1)

(ii) (6) Correlation coefficient = ?

S.N	X	Y	XY	X^2	Y^2
1	10	3	30	100	9
2	12	6	72	144	36
3	15	9	135	225	81
4	20	12	240	400	144
5	25	15	375	625	225
6	30	18	540	900	324
7	35	20	700	1225	400
8	40	22	880	1600	484
9	45	24	1080	2025	576
10	50	26	1300	2500	676
Total	$\Sigma X =$ 282	$\Sigma Y =$ 155	$\Sigma XY =$ 5352	$\Sigma X^2 =$ 9744	$\Sigma Y^2 =$ 2955

(2)

formula

$$r = \frac{cov(X,Y)}{\sigma_X \cdot \sigma_Y} \quad (\frac{1}{2})$$

Date:
P. No:

$$Cov(X,Y) = \frac{\Sigma XY}{n} - \bar{X}\cdot\bar{Y} \quad (\tfrac{1}{2})$$

$$\sigma_X^2 = \frac{1}{n}\Sigma X^2 - \left(\frac{\Sigma X}{n}\right)^2 \quad (\tfrac{1}{2})$$

$$\sigma_Y^2 = \frac{1}{n}\Sigma Y^2 - \left(\frac{\Sigma Y}{n}\right)^2 \quad (\tfrac{1}{2})$$

here

$$n = 10$$

$$\boxed{\bar{X} = \Sigma X/n \Rightarrow 282/10 \Rightarrow 28.2}$$

$$\boxed{\bar{Y} = \Sigma Y/n \Rightarrow 155/10 \Rightarrow 15.5}$$

$$Cov(X,Y) = \frac{5352}{10} - (28.2)(15.5)$$

$$= 535.2 - 437.1$$

$$\boxed{Cov(X,Y) = 98.1} \quad (\tfrac{1}{2})$$

$$\sigma_X^2 = \frac{9744}{10} - (28.2)^2$$

$$= 974.4 - 795.24$$

$$\sigma_X^2 = 179.16$$

$$\boxed{\sigma_X = 13.38} \quad (\tfrac{1}{2})$$

$$\sigma_Y^2 = \frac{2955}{10} - (15.5)^2$$

$$= 295.5 - 240.25$$

$$\sigma_Y^2 = 55.25$$

$$\boxed{\sigma_Y = 7.43} \quad (\tfrac{1}{2})$$

Date:
P. No:

$$\gamma = \frac{98.1}{(13.38)(7.43)}$$

$$\gamma = \frac{98.1}{99.41}$$

$$\boxed{\gamma = 0.98} \quad (\tfrac{1}{2})$$

OR (6)

(iii) Regression Line x on y and y on x = ?

S.N	X	Y	XY	X^2	Y^2
1	3	2	6	9	4
2	5	4	20	25	16
3	7	8	56	49	64
4	9	10	90	81	100
5	12	13	156	144	169
6	15	17	255	225	289
7	18	21	378	324	441
8	20	23	460	400	529
Total	ΣX = 89	ΣY = 98	ΣXY = 1421	ΣX^2 = 1257	ΣY^2 = 1612

(2)

Regression Line Y on X is

$$Y - \bar{Y} = \gamma \cdot \frac{\sigma_Y}{\sigma_X}(X - \bar{X}) \quad (\tfrac{1}{2})$$

OR

$$Y - \bar{Y} = b_{YX}(X - \bar{X})$$

$$b_{yx} = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$$ (½)

$$b_{xy} = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{n\Sigma y^2 - (\Sigma y)^2}$$ (½)

$n = 8$

$$b_{yx} = \frac{8[1421] - (89)(98)}{8[1257] - [89]^2}$$

$$b_{yx} = \frac{11368 - 8722}{10056 - 7921}$$

$$b_{yx} = \frac{2646}{2135}$$

$$b_{yx} = 1.23$$ (½)

$$b_{xy} = \frac{8[1421] - (89)(98)}{8[1612] - [98]^2}$$

$$b_{xy} = \frac{11368 - 8722}{12896 - 9604}$$

$$b_{xy} = \frac{2646}{3292}$$

$$b_{xy} = 0.80$$ (½)

Date:
P. No:

(½) $\bar{y} = \Sigma y/n \Rightarrow 98/8 = \boxed{12.25}$

$\bar{x} = \Sigma x/n \Rightarrow 89/8 = \boxed{11.12}$

$y - (12.25) = 1.23(x - 11.12)$

$y - 12.25 = 1.23x - 13.67$

$y - 1.23x = 12.25 - 13.67$

$y - 1.23x = -1.42$

$\boxed{y = 1.23x - 1.42}$

$x - \bar{x} = b_{xy}(y - \bar{y})$ (½)

$x - 11.12 = 0.80(y - 12.25)$

$x - 11.12 = 0.80y - 9.8$

$x - 0.80y = 11.12 - 9.80$

$x - 0.80y = +1.32$

(½) $\boxed{x = 0.80y + 1.32}$

3.4 (i) → (a) Type - I error & Type - II error

or

Type - I error

In statistical hypothesis Testing, a type I error is the mistaken rejection

Date
P No

a an actually true Null Hypothesis. while a type II error is the failure to reject a null Hypothesis that is actually false. A Type-I error is a kind of fault that occurs during the Hypothesis testing process when a Null Hypothesis is rejected. (1/2)

Type II error:

A Type II error is a statistical term within the context of hypothesis testing that describes the error that occurs when one fails to reject a null hypothesis that is actually false. A type-II error produces a false negative, also known as an error of omission. (1/2)

(b) Critical Region:

A critical region, also known as the rejection region, is a set of values for the test statistic for which the null hypothesis is rejected. i.e. if the observed test statistic is in the critical region then we reject the null hypothesis and accept the alternative hypothesis. (1)

(c) Null Hypothesis:

The null hypothesis is a typical statistical theory which suggests that no statistical relationship and significance exists in a set of given single observed variable, b/w two sets of observed data and measured phenomena.

Null hypothesis is denoted by H_0.

Date:
P. No:

(d) Alternative Hypothesis : An alternative hypothesis is one in which a difference (or an effect) b/w two or more variables is anticipated by the researchers. that is the observed pattern of the data is not due to a chance occurance.

its denoted by H_1 (1)

(ii) Neymann pearson Lemma :

Let $K > 0$ be a constant and W be a critical region of size α such that

$$W = \left\{ x \in S ; \frac{f(x, \theta_1)}{f(x, \theta_0)} > K \right\}$$

$$W = \left\{ x \in S ; \frac{L_1}{L_0} > K \right\}$$

and $$\overline{W} = \left\{ x \in S ; \frac{L_1}{L_0} \leq K \right\}$$

where L_0 and L_1 are the Likelihood function of the sample observations under H_0 and H_1 respectively. then W is the most powerful critical region

Date:
P. No:

of the test hypothesis $H_0: \theta = \theta_0$ against $H_1: \theta \neq \theta_0$

Proof: we are given

$$P(x \in W | H_0) = \int_W L_0 dx = \alpha$$

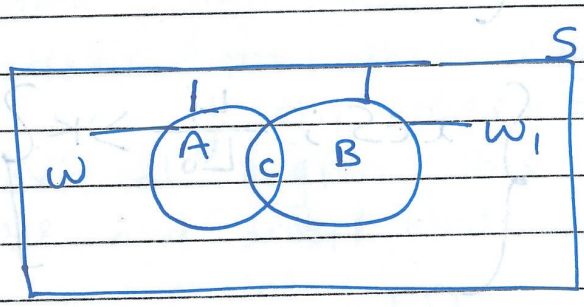
$$P(x \in W | H_1) = \int_W L_1 dx = 1 - \beta \text{ (say)}$$

Let W_1 be another critical region of size $\alpha_1 \leq \alpha$ and power $1-\beta_1$ so that we have

$$P(x \in W_1 | H_0) = \int_{W_1} L_0 dx = \alpha_1$$

and $$P(x \in W_1 | H_1) = \int_{W_1} L_1 dx = 1 - \beta_1$$

Now we have to prove that $1 - \beta \geq 1 - \beta_1$



$W = A \cup C$, $W_1 = B \cup C$

if $\alpha \geq \alpha_1$ we have

$$\int_{W_1} L_0 dx \leq \int_W L_0 dx$$

$$\int_{B \cup C} L_0 dx \leq \int_{A \cup C} L_0 dx$$

$$\int_B L_0\,dx \leq \int_A L_0\,dx$$

$$\boxed{\int_A L_0\,dx \geq \int_B L_0\,dx}$$

ACW

$$\int_A L_1\,dx > k\int_A L_0\,dx \geq k\int_B L_0\,dx$$

$$\frac{L_1}{L_0} \leq k \quad \forall\, x \in W$$

$$\int_W L_1\,dx \leq k\int_W L_0\,dx$$

$$\int_B L_1\,dx \leq k\int_B L_0\,dx \leq \int_A L_1\,dx$$

$$\int_{W_1} L_1\,dx \leq \int_W L_1\,dx$$

$$\boxed{1-\beta \geq 1-\beta_1}$$ (6)

Or

(iii) M·L·E : The method of maximum Likelihood estimation is one of the most important methods to obtain the estimates of the parameters involved in the given probability function of the random variables (2)

properties of M·L·E

Date:
P. No:

(4)

(a) MLE is a consistent estimator in general.
(b) MLE is a Sufficient estimator, if sufficient estimator exists.
(c) MLE may or may not be unbiased estimator. However, it can usually be made unbiased with some modifications.
(d) if MLE exists, it is most efficient in the class of such estimators. (6)

Q.5 (i) four difference

parametric:

(1) Used mainly on interval and ratio scale data.
(2) Tend to need larger samples.
(3) data should fit a particular distribution, the data can be transformed to that distribution.
(4) Samples should be drawn randomly from the population.

(2)

Non-parametric

(1) Can be used on ordinal and nominal scale data.
(2) Can be used on data that are not normally distributed.
(3) Can be used where the samples are not selected randomly.
(4) Have less power than the equivalent parametric test.

(2)

(4)

Date
P No:

(ii) Run test

H_0: males and females are standing in random order

H_1: males and females are not standing in random order.

$r = 35$, $n_1 = 30$, $n_2 = 20$

$$\mu_r = \frac{2n_1n_2}{(n_1+n_2)} + 1$$

$$= \frac{2(30)(20)}{30+20} + 1$$

$$= \frac{1200}{50} + 1$$

$$\boxed{\mu_r = 25}$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$$

$$\sigma_r = \sqrt{\frac{2(30)(20)(2\cdot 30\cdot 20 - 30 - 20)}{(30+20)^2(30+20-1)}}$$

$$\boxed{\sigma_r = 3.356}$$

$$Z_{cal} = \frac{r - \mu_r}{\sigma_r}$$

$$\boxed{Z_{cal} = 2.979}$$

Date:
P. No:

Z_{tab} at 1% = 2.58

Since

$Z_{cal} > Z_{tab}$

so H_0 is rejected

or (iii) Wilcoxon Sign rank test

$H_0: \theta = \theta_0$

$H_1: \theta \neq \theta_0$

x_i	$d_i = x_i - \theta_0$	ranks
9.3	-0.6	9.5
8.8	-1.1	14
10.7	0.8	11.5
11.5	1.6	18
8.2	-1.7	19
9.7	-0.2	3.5
10.3	0.4	6.5
8.6	-1.3	15.5
11.3	1.4	17
10.7	0.8	11.5
11.2	1.3	15.5
9.0	-0.9	13
9.8	-0.1	1.5
9.3	-0.6	9.5
9.9	0	Ignored
10.3	0.4	6.5
10.0	0.1	1.5
10.1	0.2	3.5
9.6	-0.3	5
10.4	0.5	8

T^+ = Sum of ranks of the + di

$T^+ = 99.5$

T^- = Sum of ranks of the - di

$T^- = 90.5$

$n = 19$

$\alpha = 0.05$

$T_\alpha = 46$ [from the table]

$T^+ > T_\alpha$ or $T^- > T_\alpha$

Since T^+ and T^- both greater than T_α there is no significant evidence to reject H_0. it mean at we accept Null hypothesis (6)

Q.6 (i) One way Anova (1)

one way Anova compares the mean of two or more independent groups in order to determine whether there is statistical evidence that the associated population means are significantly different.

(1) example

As a crop researcher, you want to test the effect of three different fertilizer mixtures on crop yield.

Date:
P. No:

find out if there is a difference in crop yields b/w the three groups.

(2) Anova Table

Source of Variation	Sum of Squares	df	mean Sum of squares	Variance ratio
treatment	$TSS = S_t^2$	$k-1$	$\delta_t^2 = S_t^2/k-1$	$F_{cal} = \frac{\delta_t^2}{\delta_E^2}$
error	$ESS = S_E^2$	$n-k$	$\delta_E^2 = S_E^2/n-k$	
Total	$SST = S_T^2$	$n-1$	$\delta_T^2 = S_T^2/n-1$	

(4)

ii) H_0: There is no significant difference in the avg. yield under the three varieties.

H_1: The difference is significant

A		B		C	
x_1	x_1^2	x_2	x_2^2	x_3	x_3^2
25	625	20	400	24	576
22	484	17	289	26	676
24	576	16	256	30	900
21	441	19	361	20	400
$\Sigma x_1 = 92$	$\Sigma x_1^2 =$ 2126	$\Sigma x_2 =$ 72	$\Sigma x_2^2 =$ 1306	$\Sigma x_3 =$ 100	$\Sigma x_3^2 =$ 2552

SST = 176

N = 12

TSS = 104

SSE = 72

t.df = 2

E.df = 9

$s_t^2 = 52$

$s_E^2 = 8$

$F_{cal} = \frac{52}{8} = 6.5$

$F_{tab} = 4.26$

$F_{cal} > F_{tab}$

H_0 is rejected.

G = 264

N = 12

$CF = \frac{G^2}{N}$

$= \frac{(264)^2}{12}$

$CF = \frac{69696}{12}$

CF = 5808

6

Date:
P. No:

Or

Detergent

(iii) H_0: There is no Significant difference in whitness due to three varieties of detergent.

H_1: There is significant difference

water temperature

H_0: There is significant diff b/w water

H_1: There is significant diff

D→	A	B	C		
W↓				Total	mean
C·W	57	55	67	179	59.6
W·W	49	52	68	169	56.3
H·W	59	46	58	163	54.3
Total	165	153	193	G=511	
mean	55	51	64.3		

SST = 439.56

SSR = 43.5

SSC = 280.9

cf = 29013.44

E df = 4

N = 9

N.df = 8

R.df = 2

C.df = 2.

MSR = 21.75

MSC = 140.45

MSE = 28.8

$F_{R\,cal} = 0.75$

$F_{C\,cal} = 4.87$

$F_{tab} = 6.94$

$F_{tab} > F_{R\,cal}$

$F_{tab} > F_{C\,cal}$

H_0 is accepted

6