| | | |
|---|---|---|
| **CMR<br>INSTITUTE OF<br>TECHNOLOGY** | USN | |

# Internal Assessment Test – III January 2024

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sub: | Mathematics for Computer Science | | | | | | | Code: | BCS301 |
| Date: | 04/03/2024 | Duration: | 90 mins | Max Marks: | 50 | Sem: | III | Branch: | **AIML/AIDS/<br>CSE/ISE/<br>CS DS/CS ML** |

**Question 1 is compulsory and Answer any 6 from the remaining questions.**

<table>
<tr><th rowspan="2"></th><th rowspan="2"></th><th rowspan="2">Marks</th><th colspan="2">OBE</th></tr>
<tr><th>CO</th><th>RBT</th></tr>
<tr><td>1</td><td>Analyze and interpret the following statistics concerning output of wheat per field obtained as a result of experiment conducted to test four varieties of wheat viz. A, B, C and D under a Latin- square design
<table>
<tr><td>C<br>25</td><td>B<br>23</td><td>A<br>20</td><td>D<br>20</td></tr>
<tr><td>A<br>19</td><td>D<br>19</td><td>C<br>21</td><td>B<br>18</td></tr>
<tr><td>B<br>19</td><td>A<br>14</td><td>D<br>17</td><td>C<br>20</td></tr>
<tr><td>D<br>17</td><td>C<br>20</td><td>B<br>21</td><td>A<br>15</td></tr>
</table>
Given that $F_{3,6} = 4.76$</td><td>[8]</td><td>CO6</td><td>L3</td></tr>
<tr><td>2</td><td>In a recent study reported on the Flurry Blog, the mean age of tablet users is 34 years. Suppose the standard deviation is 15 years. Take a sample of size $n = 100$. Using central limit theorem, find the probability that the sample mean age is more than 30 years.<br>Consider $\Phi(2.66) = 0.4961$</td><td>[7]</td><td>CO5</td><td>L2</td></tr>
<tr><td>3</td><td>Suppose that 10, 12, 16, 19 is a sample taken from a normal population with variance 6.25. Find a 95 percent confidence interval for the population mean.</td><td>[7]</td><td>CO5</td><td>L3</td></tr>
<tr><td>4</td><td>A random sample of 10 boys had the following I.Q.:<br>70, 120, 110, 101, 88, 83, 95, 98, 107, 100.<br>Do these data support the assumption of a population mean I.Q. of 100 (at 5% level of significance)?<br>($t_{0.05} = 2.262$)</td><td>[7]</td><td>CO5</td><td>L3</td></tr>
</table>

5

A die was thrown 60 times and the following frequency distribution was observed:

| Faces | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| frequency | 15 | 6 | 4 | 7 | 11 | 17 |

Test whether the die is unbiased at 5% significance level.

$\chi^2_{0.05} = 11.07$

[7] CO5 L3

6

Three different kinds of food are tested on three groups of rats for 5 weeks. The objective is to check the difference in mean weight (in grams) of the rats per week. Apply one-way ANOVA using a 0.05 significance level to the following data:

| | | | | | | |
|---|---|---|---|---|---|---|
| **Food 1** | **8** | **12** | **19** | **8** | **6** | **11** |
| **Food 2** | **4** | **5** | **4** | **6** | **9** | **7** |
| **Food 3** | **11** | **8** | **7** | **13** | **7** | **9** |

$F_{2,15} = 3.68$

[7] CO6 L3

7

The following data show the number of worms quarantined from the GI areas of four groups of muskrats in a carbon tetrachloride anthelmintic study. Conduct a two-way ANOVA test.

| I | II | III | IV |
|---|---|---|---|
| 33 | 41 | 12 | 38 |
| 32 | 38 | 35 | 43 |
| 26 | 40 | 46 | 25 |
| 14 | 23 | 22 | 13 |
| 30 | 21 | 11 | 26 |

Given $F(4,12) = 3.26$ & $F(3,12) = 3.49$

[7] CO6 L3

8

Set up ANOVA table for the following information relating to three drugs testing to judge the effectiveness in reducing blood pressure for three different groups of people

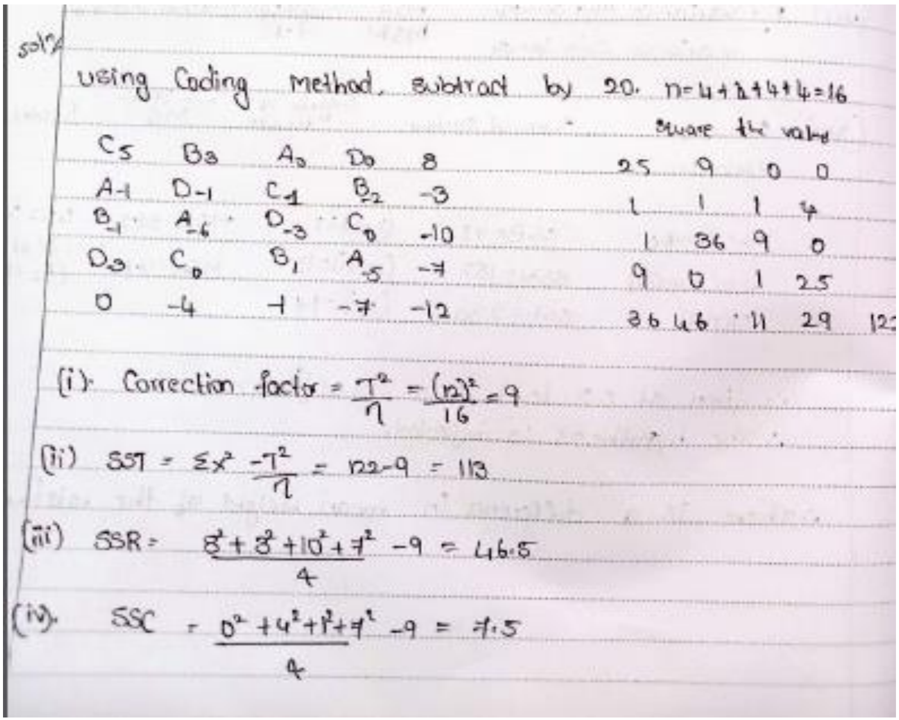
| Group of people | Drug | | |
|---|---|---|---|
| | X | Y | Z |
| A | 14 | 10 | 11 |
| | 15 | 9 | 11 |
| B | 12 | 7 | 10 |
| | 11 | 8 | 11 |
| C | 10 | 11 | 8 |
| | 11 | 11 | 7 |

Do the drugs act differently?
Are the different groups of people affected differently?
Is the interaction term significant?
Answer the above questions taking a significant level of 5%

$F(2,9) = 4.26, F(2,9) = 4.26, F(4,9) = 3.63$

[7] CO6 L3

## 1.



sol^n/

using Coding method. subtract by 20. n=4+4+4+4=16

Square the value

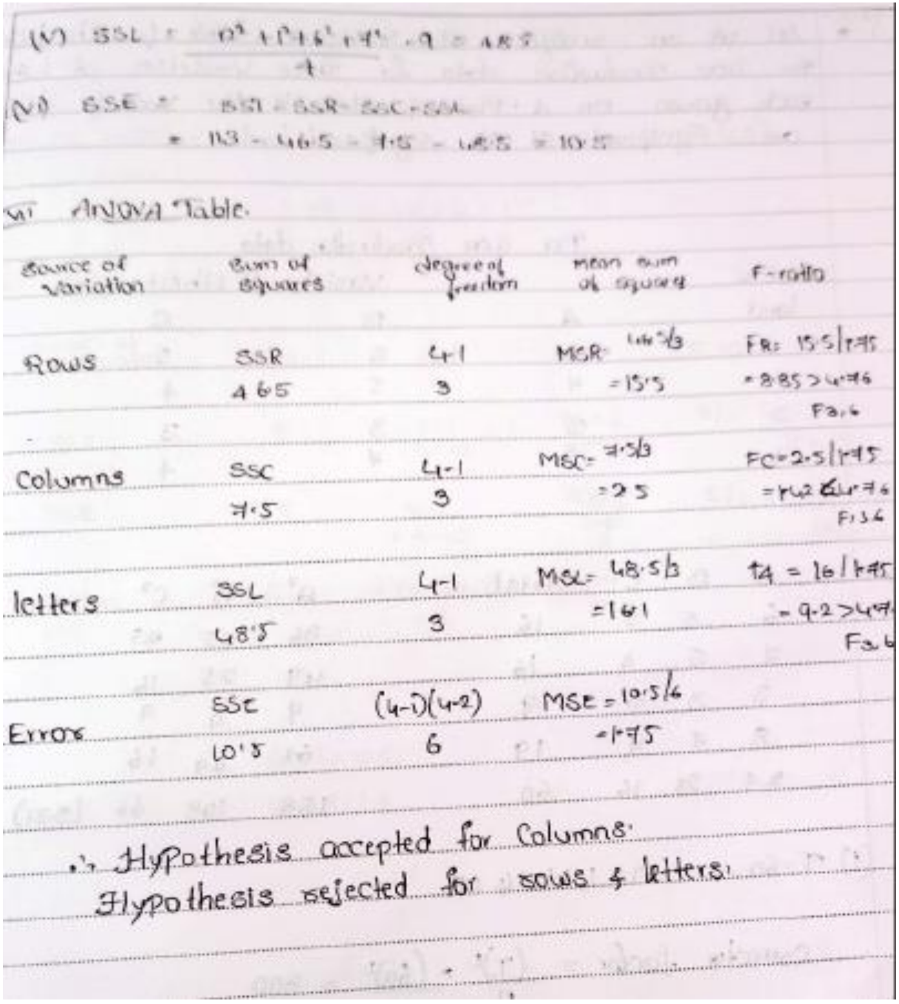
|  |  |  |  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|---|
| $C_5$ | $B_3$ | $A_0$ | $D_0$ | 8 | 25 | 9 | 0 | 0 |  |
| $A_{-1}$ | $D_{-1}$ | $C_1$ | $B_2$ | -3 | 1 | 1 | 1 | 4 |  |
| $B_{-1}$ | $A_{-6}$ | $D_{-3}$ | $C_0$ | -10 | 1 | 36 | 9 | 0 |  |
| $D_3$ | $C_0$ | $B_1$ | $A_{-5}$ | -7 | 9 | 0 | 1 | 25 |  |
| 0 | -4 | -1 | -7 | -12 | 36 | 46 | 11 | 29 | 12 |

(i) Correction factor = $\frac{T^2}{n} = \frac{(12)^2}{16} = 9$

(ii) $SST = \Sigma x^2 - \frac{T^2}{n} = 122 - 9 = 113$

(iii) $SSR = \frac{8^2+3^2+10^2+7^2}{4} - 9 = 46.5$

(iv) $SSC = \frac{0^2+4^2+1^2+7^2}{4} - 9 = 7.5$



(v) $SSL = \frac{0^2+1^2+6^2+4^2}{4} - 9 = 48.5$

(vi) SSE = SST − SSR − SSC + SSL

= 113 − 46.5 − 7.5 − 48.5 = 10.5

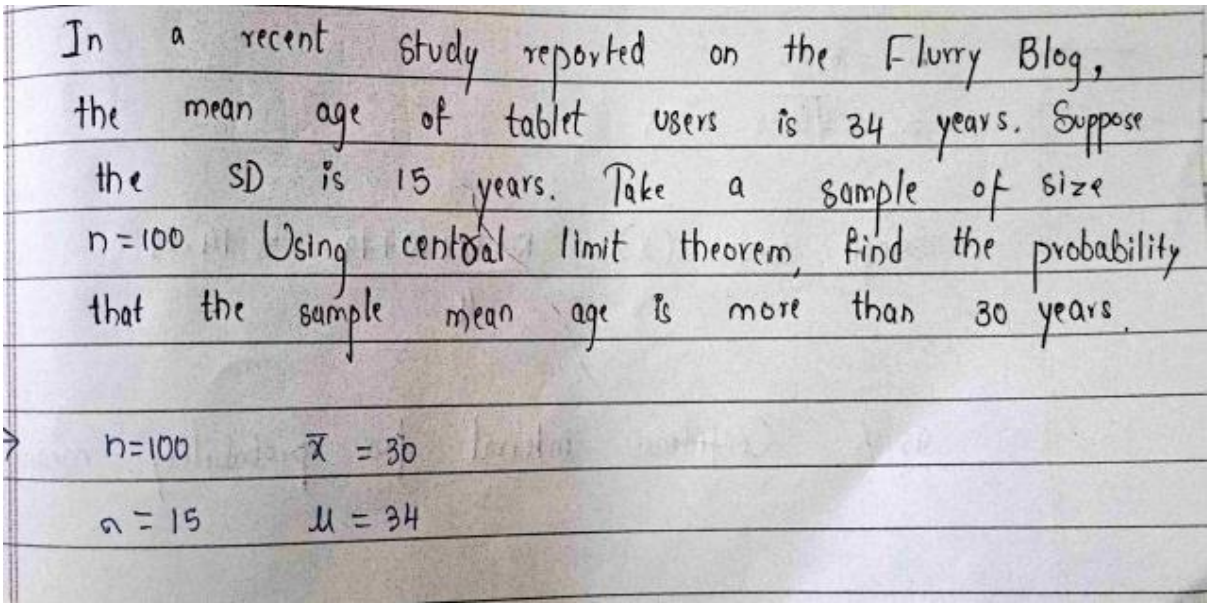
(vii) ANOVA Table.

| Source of variation | Sum of squares | degree of freedom | mean sum of squares | F-ratio |
|---|---|---|---|---|
| Rows | SSR<br>46.5 | 4-1<br>3 | MSR = 46.5/3<br>= 15.5 | FR = 15.5/1.75<br>= 8.85 > 4.76<br>$F_{3,6}$ |
| Columns | SSC<br>7.5 | 4-1<br>3 | MSC = 7.5/3<br>= 2.5 | FC = 2.5/1.75<br>= 1.42 < 4.76<br>$F_{3,6}$ |
| letters | SSL<br>48.5 | 4-1<br>3 | MSL = 48.5/3<br>= 16.1 | FL = 16/1.75<br>= 9.2 > 4.7<br>$F_{3,}$ |
| Error | SSE<br>10.5 | (4-1)(4-2)<br>6 | MSE = 10.5/6<br>= 1.75 |  |

∴ Hypothesis accepted for Columns.

Hypothesis rejected for rows & letters.

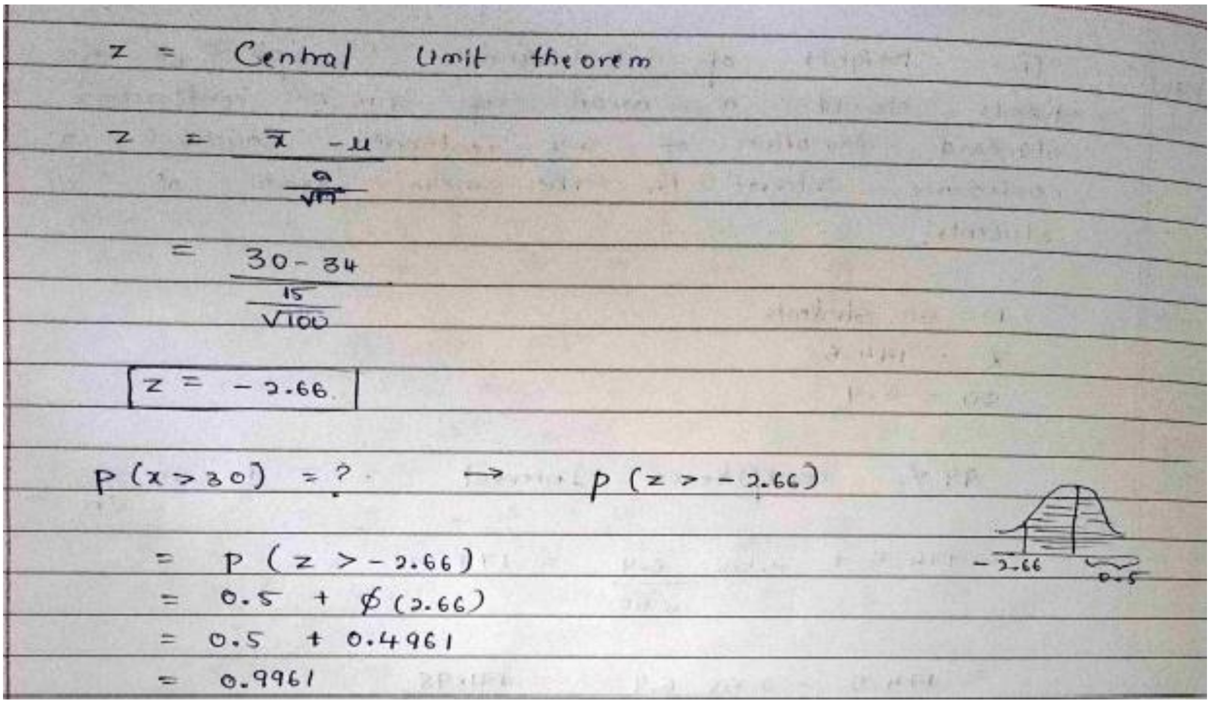
**2.**



In a recent study reported on the Flurry Blog, the mean age of tablet users is 34 years. Suppose the SD is 15 years. Take a sample of size $n=100$ Using central limit theorem, find the probability that the sample mean age is more than 30 years.

$n=100 \quad \bar{x} = 30$

$\sigma = 15 \quad \mu = 34$



z = Central limit theorem

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{30-34}{\frac{15}{\sqrt{100}}}$$

$$\boxed{z = -2.66}$$

$p(x>30) = ? \longrightarrow p(z > -2.66)$

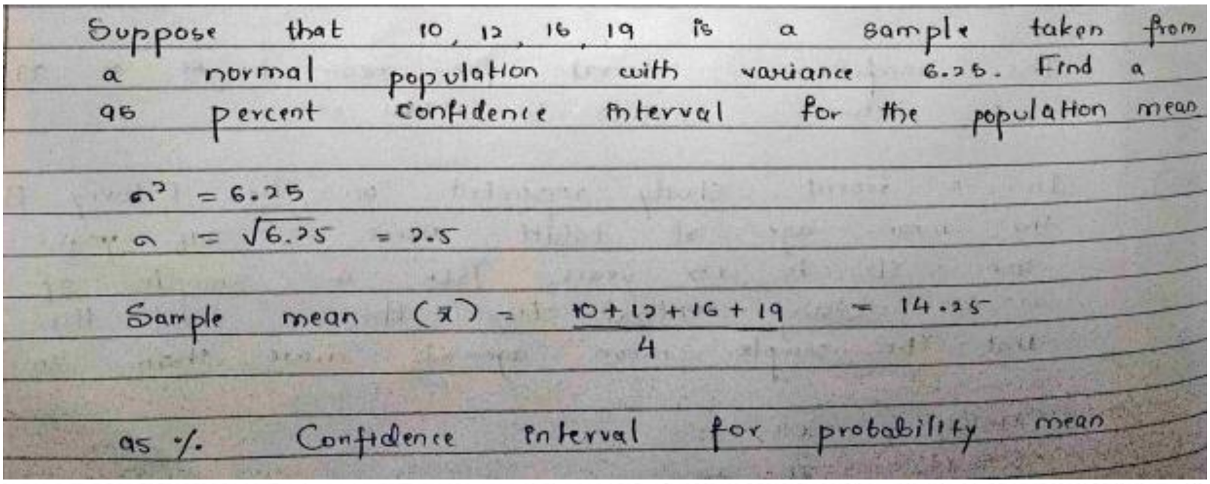
$= p(z > -2.66)$

$= 0.5 + \phi(2.66)$

$= 0.5 + 0.4961$

$= 0.9961$

# 3.



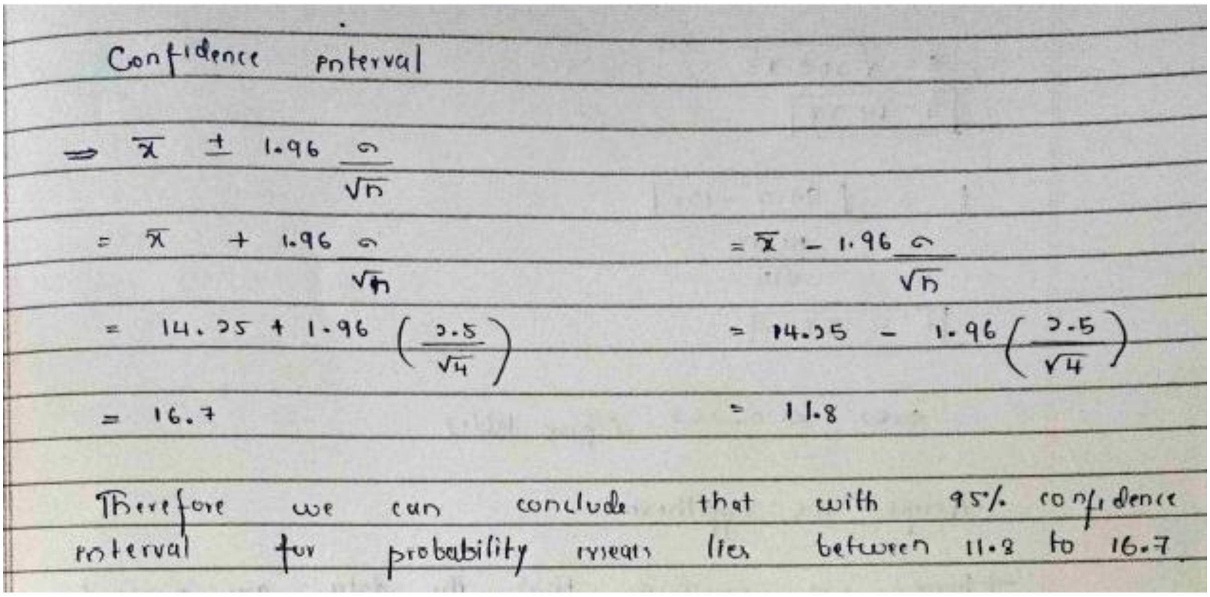
Suppose that 10, 12, 16, 19 is a sample taken from a normal population with variance 6.25. Find a 95 percent confidence interval for the population mean

$\sigma^2 = 6.25$

$\therefore \sigma = \sqrt{6.25} = 2.5$

Sample mean $(\bar{x}) = \frac{10+12+16+19}{4} = 14.25$

95 % Confidence interval for probability mean



Confidence interval

$\Rightarrow \bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}$

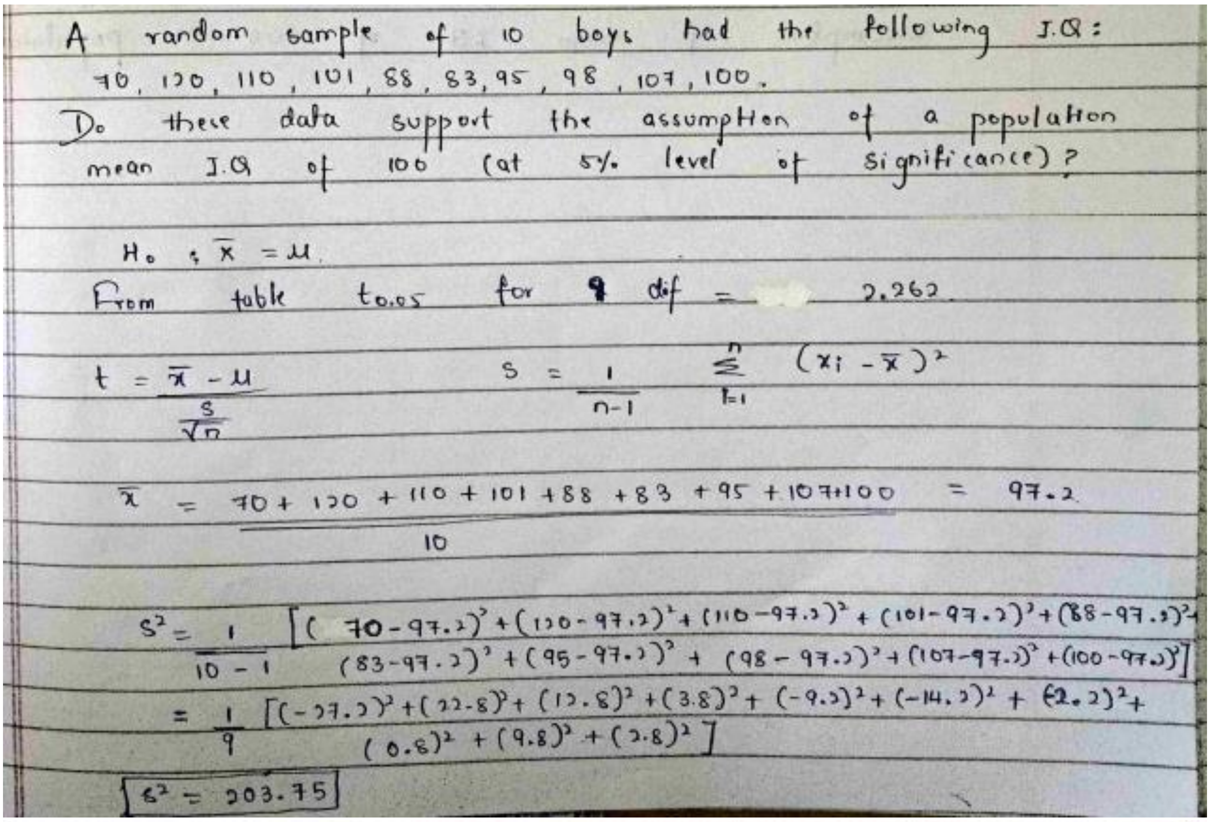
$= \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}$ $\qquad = \bar{x} - 1.96 \frac{\sigma}{\sqrt{n}}$

$= 14.25 + 1.96 \left(\frac{2.5}{\sqrt{4}}\right)$ $\qquad = 14.25 - 1.96 \left(\frac{2.5}{\sqrt{4}}\right)$

$= 16.7$ $\qquad = 11.8$

Therefore we can conclude that with 95% confidence interval for probability means lies between 11.8 to 16.7.

# 4.



A random sample of 10 boys had the following I.Q :
70, 120, 110, 101, 88, 83, 95, 98, 107, 100.
Do these data support the assumption of a population mean I.Q of 100 (at 5% level of significance) ?

$H_0 : \bar{x} = \mu$

From table $t_{0.05}$ for 9 d.f = 2.262

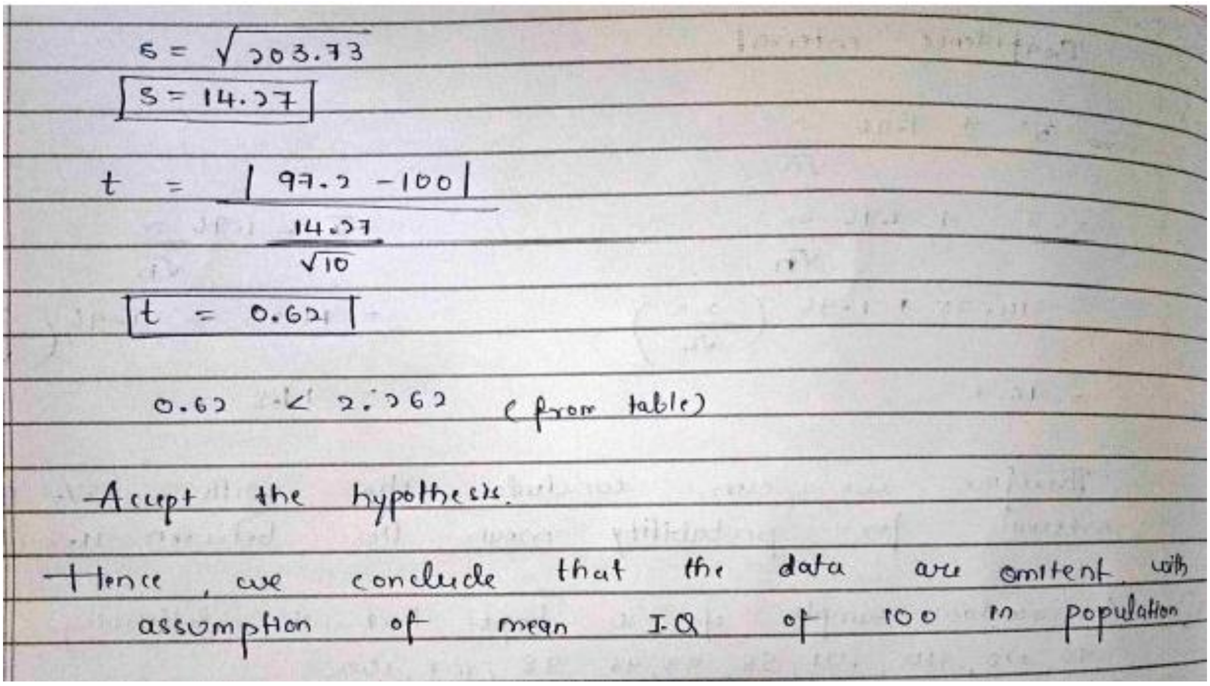
$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ $\quad s = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

$$\bar{x} = \frac{70 + 120 + 110 + 101 + 88 + 83 + 95 + 107 + 100}{10} = 97.2$$

$$s^2 = \frac{1}{10-1} \left[ (70-97.2)^2 + (120-97.2)^2 + (110-97.2)^2 + (101-97.2)^2 + (88-97.2)^2 + (83-97.2)^2 + (95-97.2)^2 + (98-97.2)^2 + (107-97.2)^2 + (100-97.2)^2 \right]$$

$$= \frac{1}{9} \left[ (-27.2)^2 + (22.8)^2 + (12.8)^2 + (3.8)^2 + (-9.2)^2 + (-14.2)^2 + (-2.2)^2 + (0.8)^2 + (9.8)^2 + (2.8)^2 \right]$$

$\boxed{s^2 = 203.75}$



$s = \sqrt{203.73}$

$\boxed{s = 14.27}$

$$t = \frac{|97.2 - 100|}{\frac{14.27}{\sqrt{10}}}$$

$\boxed{t = 0.62}$

0.62 < 2.262 (from table)

Accept the hypothesis.

Hence, we conclude that the data are consitent with assumption of mean I.Q of 100 in population

## 5.

Sol:- The frequencies in the given data are the observed frequencies.

- $H_0$: the die is unbiased
- For unbiased die, the expected number of frequencies are,

$$E_i = \frac{15+6+4+7+11+17}{6} = \frac{60}{6} = 10 \qquad \{n=6\}$$

- Now the data is as follows

| Faces | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Observed Frequency($O_i$) | 15 | 6 | 4 | 7 | 11 | 17 |
| Expected frequency($E_i$) | 10 | 10 | 10 | 10 | 10 | 10 |

$$\therefore \quad \chi^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(15-10)^2}{10} + \frac{(6-10)^2}{10} + \frac{(4-10)^2}{10} + \frac{(7-10)^2}{10} + \frac{(11-10)^2}{10} + \frac{(17-10)^2}{10}$$

$$\chi^2 = \frac{25}{10} + \frac{16}{10} + \frac{36}{10} + \frac{9}{10} + \frac{1}{10} + \frac{49}{10}$$

[illegible], [illegible] $\boxed{\chi^2 = 13.6}$

[illegible]:

degree of freedom $= n-1$ $\quad (n=6$, no of faces$)$

$$= 6-1$$
$$= 5$$

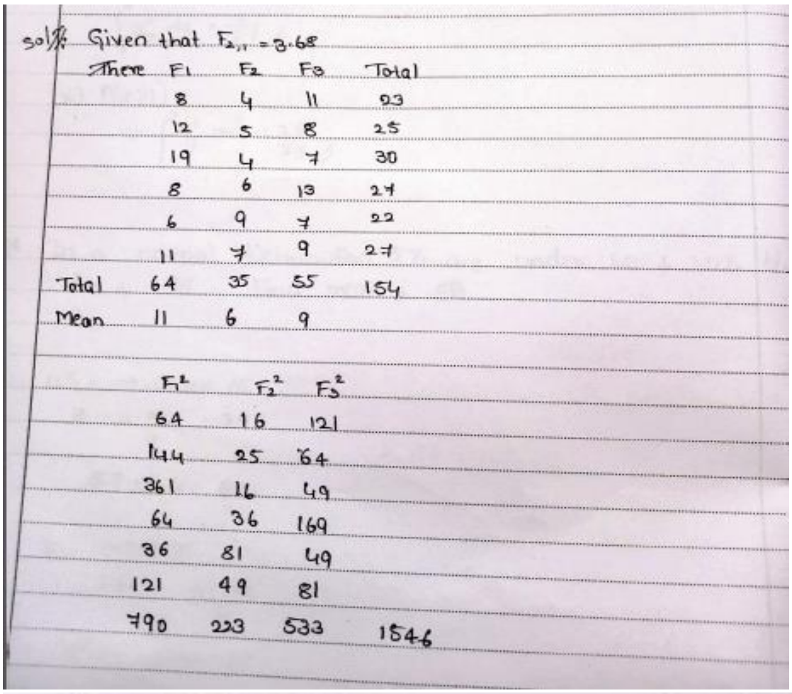
$\chi^2_{0.05} = 11.07$ for 5 d.f

$\therefore$ [illegible]

$$\therefore \quad \chi^2 = 13.6 > \chi^2_{0.05} = 11.07$$

Thus the hypothesis that the die is unbiased is rejected at 5% level of significance.

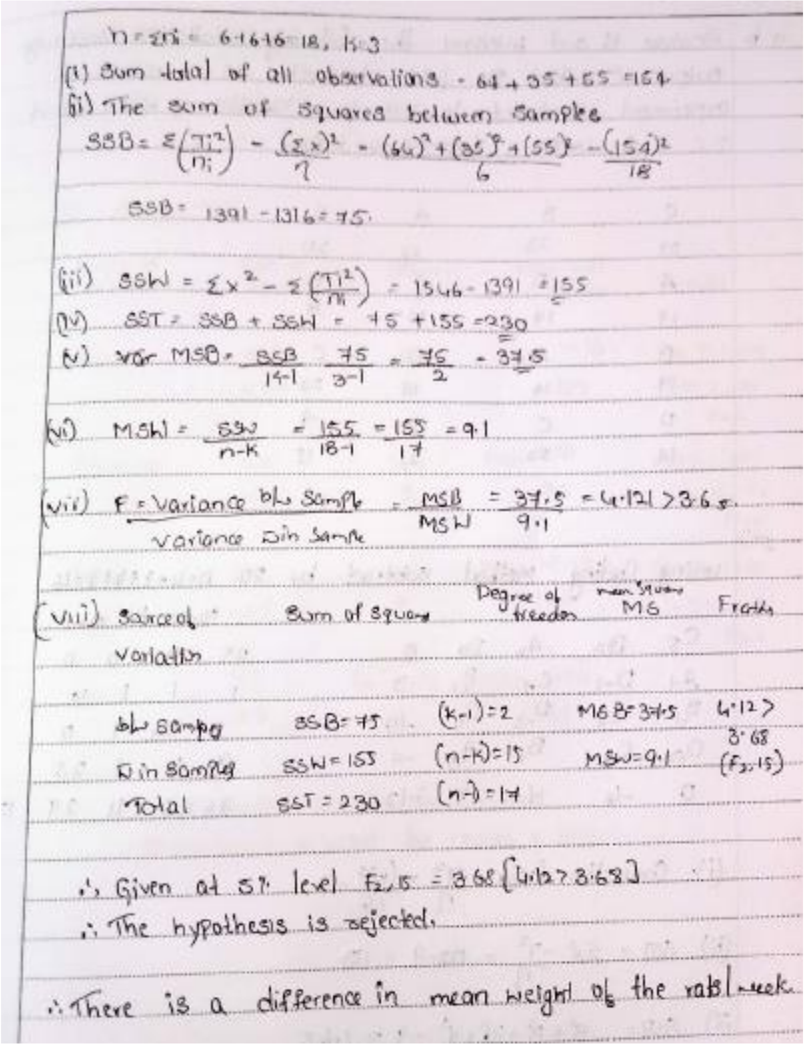
## 6.



sol: Given that $F_{2,1} = 3.68$

| There | $F_1$ | $F_2$ | $F_3$ | Total |
|---|---|---|---|---|
| | 8 | 4 | 11 | 23 |
| | 12 | 5 | 8 | 25 |
| | 19 | 4 | 7 | 30 |
| | 8 | 6 | 13 | 27 |
| | 6 | 9 | 7 | 22 |
| | 11 | 7 | 9 | 27 |
| Total | 64 | 35 | 55 | 154 |
| Mean | 11 | 6 | 9 | |

| $F_1^2$ | $F_2^2$ | $F_3^2$ | |
|---|---|---|---|
| 64 | 16 | 121 | |
| 144 | 25 | 64 | |
| 361 | 16 | 49 | |
| 64 | 36 | 169 | |
| 36 | 81 | 49 | |
| 121 | 49 | 81 | |
| 790 | 223 | 533 | 1546 |



$n = \Sigma n_i = 6+6+6 = 18, \; k = 3$

(i) Sum total of all observations $= 64 + 35 + 55 = 154$

(ii) The sum of squares between samples

$$SSB = \Sigma\left(\frac{T_i^2}{n_i}\right) - \frac{(\Sigma x)^2}{7} = \frac{(64)^2 + (35)^2 + (55)^2}{6} - \frac{(154)^2}{18}$$

$SSB = 1391 - 1316 = 75$

(iii) $SSW = \Sigma x^2 - \Sigma\left(\frac{T_i^2}{n_i}\right) = 1546 - 1391 = \underline{155}$

(iv) $SST = SSB + SSW = 75 + 155 = \underline{230}$

(v) var $MSB = \frac{SSB}{k-1} \; \frac{75}{3-1} = \frac{75}{2} = \underline{37.5}$

(vi) $MSW = \frac{SSW}{n-k} = \frac{155}{18-1} = \frac{155}{17} = 9.1$

(vii) $F = \frac{\text{Variance b/w Sample}}{\text{Variance w/in Sample}} = \frac{MSB}{MSW} = \frac{37.5}{9.1} = 4.121 > 3.68$.

(viii)

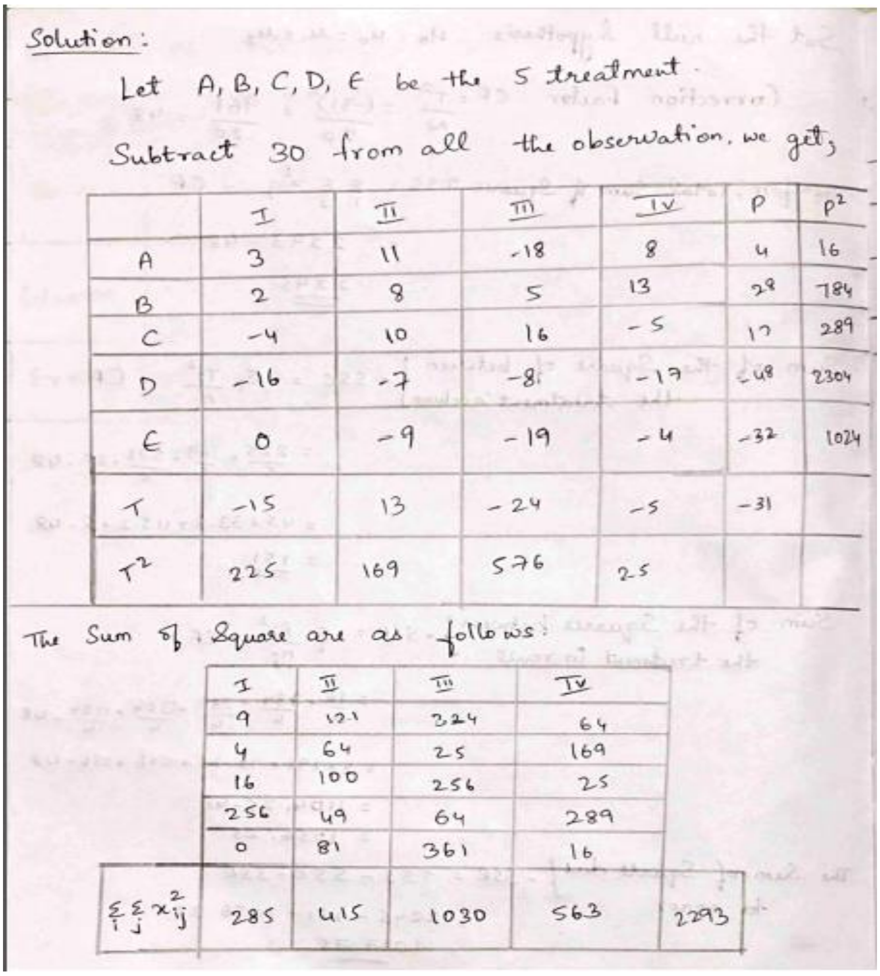
| Source of variation | Sum of square | Degree of freedom | mean square MS | F ratio |
|---|---|---|---|---|
| b/w samples | SSB = 75 | (k-1) = 2 | MSB = 37.5 | 4.12 > 3.68 $(F_{2,15})$ |
| w/in samples | SSW = 155 | (n-k) = 15 | MSW = 9.1 | |
| Total | SST = 230 | (n-1) = 17 | | |

∴ Given at 5% level $F_{2,15} = 3.68$ [4.12 > 3.68]

∴ The hypothesis is rejected.

∴ There is a difference in mean weight of the rats/week.

# 7.



Solution:

Let A, B, C, D, E be the 5 treatment.

Subtract 30 from all the observation, we get;

| | I | II | III | IV | P | $P^2$ |
|---|---|---|---|---|---|---|
| A | 3 | 11 | -18 | 8 | 4 | 16 |
| B | 2 | 8 | 5 | 13 | 28 | 784 |
| C | -4 | 10 | 16 | -5 | 17 | 289 |
| D | -16 | -7 | -8 | -17 | -48 | 2304 |
| E | 0 | -9 | -19 | -4 | -32 | 1024 |
| T | -15 | 13 | -24 | -5 | -31 | |
| $T^2$ | 225 | 169 | 576 | 25 | | |

The Sum of Square are as follows:

| | I | II | III | IV | |
|---|---|---|---|---|---|
| | 9 | 121 | 324 | 64 | |
| | 4 | 64 | 25 | 169 | |
| | 16 | 100 | 256 | 25 | |
| | 256 | 49 | 64 | 289 | |
| | 0 | 81 | 361 | 16 | |
| $\sum_i \sum_j x_{ij}^2$ | 285 | 415 | 1030 | 563 | 2293 |

Set the null hypothesis $H_0 : \mu_0 = \mu_1 = \mu_2$

Correction factor $CF = \frac{T^2}{N} = \frac{(-31)^2}{20} = \frac{961}{20} = 48$

Therefore, Total Sum of Squares $TSS = \sum_i \sum_j x_{ij}^2 - CF$

$= 2293 - 48$

$= \underline{\underline{2245}}$

Sum of the Square of between the treatment in column $\Big\}$ - $SSC = \sum_i \frac{T_i^2}{n_i} - CF$

$= \frac{225}{5} + \frac{169}{5} + \frac{576}{5} + \frac{25}{5} - 48$

$= 45 + 33.8 + 115.2 + 5 - 48$

$= \underline{\underline{151}}$

Sum of the Squares between the treatment in rows $\Big\}$ - $SSR = \sum_i \frac{P_i^2}{n_i} - CF$

$= \frac{16}{4} + \frac{784}{4} + \frac{289}{4} + \frac{1304}{4} + \frac{1024}{4} - 48$

$= 4 + 196 + 72.25 + 576 + 256 - 48$

$= 1104.25 - 48$

$= 1056.25$

The Sum of Square due to error $\Big\}$ - $SSR = TSS - SSC - SSR$

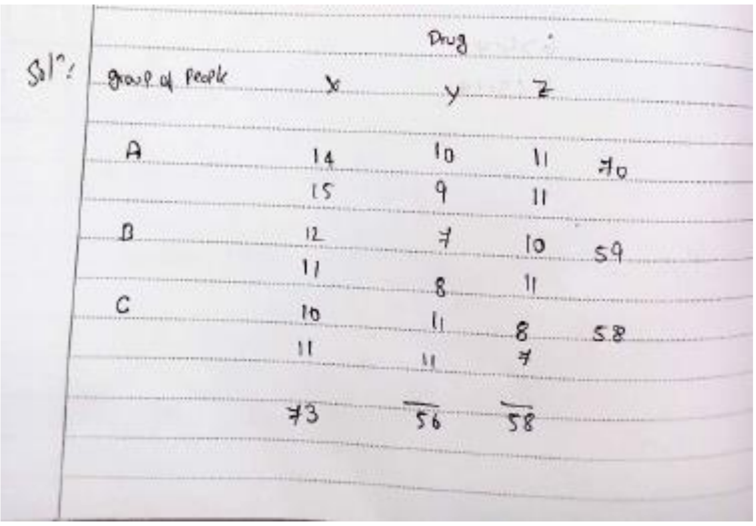
$= 2245 - 151 - 1056.25$

$= \underline{1037.75}$

We know that $F(4,12) = 3.26$ & $F(3,12) = 3.49$

| Source variation | d.o.f | SS | MSS | F Ratio | Conclusion |
|---|---|---|---|---|---|
| Rows | $5-1=4$ | $SSR = 1056.25$ | $MSR = \frac{1056.25}{4}$<br>$= 264.06$ | $F_r = \frac{264.06}{86.48}$<br>$= 3.053$ | $F_r < F(4,12)$<br>$H_0$ is accepted |
| Columns | $4-1=3$ | $SSC = 151$ | $MSC = \frac{151}{3} = 50.33$ | $F_c = \frac{86.48}{50.33}$<br>$= 1.718$ | $F_c < F(3,12)$<br>$H_0$ is Accepted |
| Error | $4 \times 3 = 12$ | $SSE = 1037.75$ | $MSE = \frac{1037.75}{12}$<br>$= 86.48$ | - | - |

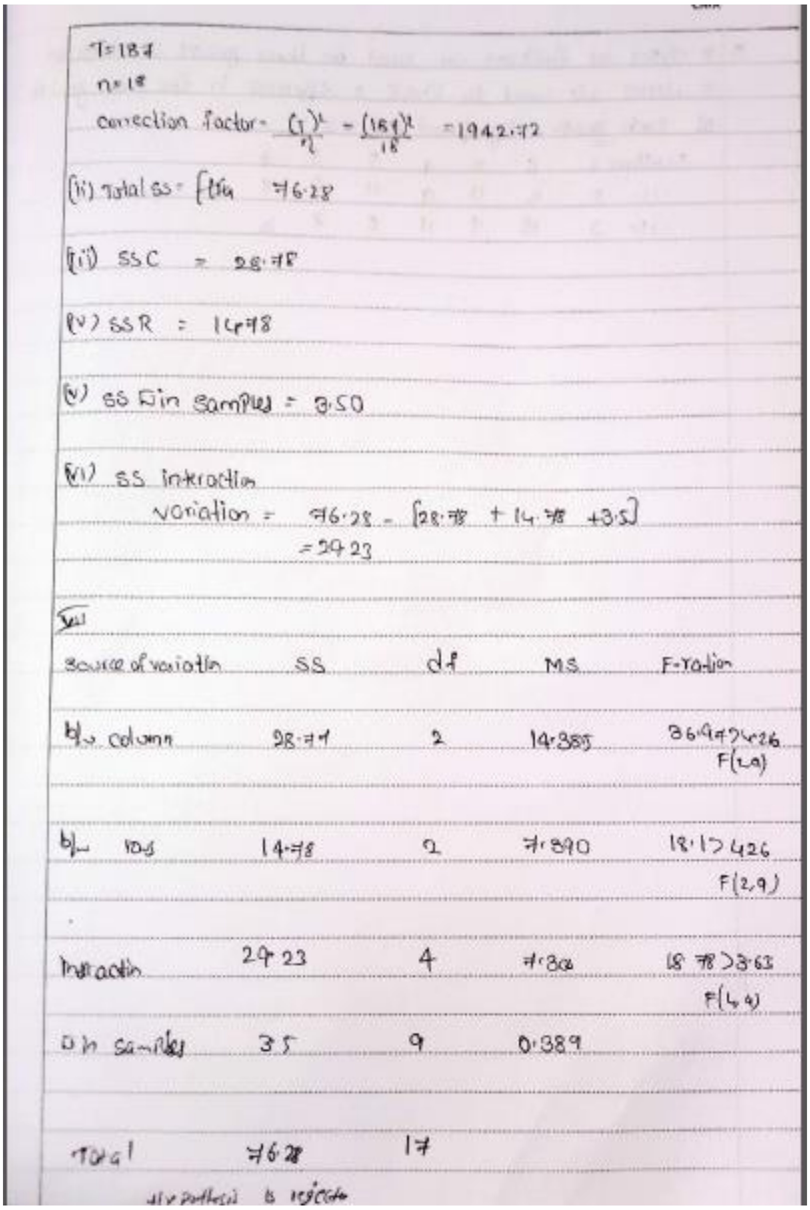
# 8.



Sol^n:

Drug

| group of people | x | y | z | |
|---|---|---|---|---|
| A | 14 | 10 | 11 | 70 |
| | 15 | 9 | 11 | |
| B | 12 | 7 | 10 | 59 |
| | 11 | 8 | 11 | |
| C | 10 | 11 | 8 | 58 |
| | 11 | 11 | 7 | |
| | 73 | 56 | 58 | |



T=187

n=18

correction factor= $\frac{(T)^2}{n} = \frac{(187)^2}{18}$ =1942.72

(ii) Total SS= [illegible] 76.28

(iii) SSC = 28.78

(iv) SSR = 14.78

(v) SS Win samples = 3.50

(vi) SS interaction

variation = 76.28 - [28.78 + 14.78 +3.5]

=29.23

(vii)

| Source of variation | SS | df | MS | F-ration |
|---|---|---|---|---|
| b/w column | 28.77 | 2 | 14.385 | 36.97>4.26<br>F(2,9) |
| b/w rows | 14.78 | 2 | 7.390 | 18.17>4.26<br>F(2,9) |
| Interaction | 29.23 | 4 | 7.30 | 18.78>3.63<br>F(4,9) |
| W/n samples | 3.5 | 9 | 0.389 | |
| Total | 76.28 | 17 | | |

H₀ hypothesis is rejected