ii. It is believed that the precision (as measured by the variance) of an instrument is no more than0.16. Write down the null and alternative hypothesis for testing this belief. Carry out the test at 1% level given 11 measurements of the same subject on the instrument: **5**
2.5, 2.3, 2.4, 2.3, 2.5, 2.7, 2.5, 2.6, 2.6, 2.7, 2.5
(Given Tabulated value 23.2of chi-square at 1% level of significance)

iii. A manufacturing company has purchased 3 new machines (A, B, C) of different machines and wishes to determine whether one of them is faster than the other in producing a certain machine. From hourly production figures are observed at random from each machine and results are given below: **5**

A	20	21	23	16	20
B	18	20	17	25	15

Use analysis of variance to test whether machines differ significantly. (Two way classification)
(Table value of F at 5% level for v_1=2 and v_2=12 is 3.89)

******

Total No. of Questions: 6 *Total No. of Printed Pages:4*

Enrollment No.....

Faculty of Engineering
End Sem Examination Dec-2023
CS3EL11 Statistical Analysis

Programme: B.Tech. Branch/Specialisation: CSE All

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. If mode of a grouped data is 10 and mean is 4, then median will be: **1**
(a) 1 (b) 4 (c) 6 (d) None of these

ii. Which of these is simply the difference between the maximum and minimum values given in a data set? **1**
(a) Range (b) Mean Deviation
(c) Standard Deviation (d) All of these

iii. What is the probability of getting the sum as a prime number if two dice are thrown? **1**
(a) 5/24 (b) 5/10 (c) 5/12 (d) None of these

iv. What will be the value of P (not E) if P(E) = 0.07? **1**
(a) 0.93 (b) 0.39 (c) 0.45 (d) None of these

v. For a Poisson distribution, parameter λ is having a value of 7. What is the variance for the distribution? **1**
(a) 2.34 (b) 2.45 (c) 7 (d) None of these

vi. Which of these cannot be shown on the continuous distributions? **1**
(a) Length dimension measurement of box
(b) Volume measurement of box
(c) Area measurement of one face of the box
(d) Number of defects on the surface of the box

vii. In the least square method we use _____ to find the value of unknowns. **1**
(a) Regression equations (b) Normal equations
(c) General equations (d) None of these

viii. Which of the following are types of correlation? **1**
(a) Positive and Negative (b) Simple, Partial and Multiple
(c) Linear and Nonlinear (d) All of these

ix. Which of the following is a Non-Parametric test? **1**
(a) T-test (b) F-test (c) Z-test (d) Chi-square test

x. A hypothesis which defines the population distribution is called? **1**
(a) Null Hypothesis (b) Simple Hypothesis
(c) Statistical Hypothesis (d) None of these

Q.2 Attempt any two:

i. Compute the Arithmetic Mean of the marks from the following table: **5**

Marks	0-10	10-20	20-30	30-40	40-50	50-60
No. of Students	5	15	25	35	45	55

ii. Find the Harmonic mean of: **5**

X	0-10	10-20	20-30	30-40	40-50
Y	4	5	11	6	4

iii. Find the Standard Deviation of the following series: **5**

Marks (above)	0	10	20	30	40	50	60	70
Students	100	90	75	50	25	15	5	0

Q.3 Attempt any two:

i. One bag contains 4 white balls and 2 black balls, another bag contain 3 white balls and 5 black balls. If one ball is drawn at random from each bag, find the probability that: **5**
(a) Both are white
(b) One is black, and one is white.

ii. A pair of dice is rolled. If the sum on the two dice is 9, find the probability that one of the dice showed 3. **5**

iii. Suppose the probability mass function of the discrete random variable is: **5**

X=x	0	1	2	3
P(x)	0.2	0.1	0.4	0.3

What is the value of $E(3X + 2X^2)$?

Q.4 i. In litters of 4 mice the number of litters which contained 0,1,2,3,4 females were noted. The figures are given in the table below: **4**

No. of Female mice	0	1	2	3	4	Total
No. of Litters	8	32	34	24	5	103

If the chance of obtaining a female in a single trial is assumed constant, estimate this constant of unknown probability. Find also the expected frequencies.

ii. A car-hire firm has two cars, which it hires out day by day. The number of demand for a car on each day is distributed as a Poisson distribution with mean 1.5. Calculate the proportion of days on which neither car is used and proportion of days on which some demand is refused. (Given $e^{-1.5} = 0.2231$) **6**

OR iii. For some normal distribution the first moment about 10 is 40 and fourth moment about 50 is 48. What is the mean variance and Standard deviation of the Normal distribution. **6**

Q.5 i. For 10 observations on price (x) and supply (y), the following data were obtained $\sum x=130$, $\sum y=220$, $\sum x^2=2228$, $\sum y^2=5506$ and, $\sum xy=3467$. **4**
Obtain the two lines of regression.

ii. Fit a second degree parabola to the following data regarding x as an independent variable: **6**

X	0	1	2	3	4
Y	1	5	10	22	38

OR iii. Calculate the Karl Pearson's coefficient of correlation between X and Y series: **6**

X	17	18	19	19	20	20	21	21	22	23
Y	12	16	14	11	15	19	22	16	15	20

Q.6 Attempt any two:

i. The height of 10 males of a given locality are found to be 70,67,62,68,61,68,70,64,64,66 inches. Is it reasonable to believe that the average height is greater than 64 inches? Test at 5% significance level assuming that for 9 degrees of freedom, P(t>1.83) =0.05. **5**

Faculty of Engg

End Sem Examination Dec-23

Date :

P. No. : 1.

CS3EL11 Statistical Analysis

Programme: B Tech Branch: CSE

Qu. (1) (10)

1. c) 6
2. a) range
3. c) 5/12
4. a) 0.93
5. c) 7
6. d) Number of defects on the surface of the box
7. b) Normal equations
8. d) All of these
9. d) Chi - square test
10. b) simple hypothesis

Qu. (2) i) (2.5)

marks	x	f	fx
0-10	5	5	25
10-20	15	15	225
20-30	25	25	625
30-40	35	35	1225
40-50	45	45	2025
50-60	55	55	3025

$$A.M = \frac{\Sigma fx}{N} = \frac{7150}{180} = 39.72$$ (2.5)

ii) (2.5)

X	Y	x	y/x	f/x
0-10	4	5	0.20	0.8
10-20	5	15	0.07	0.33
20-30	11	25	0.04	0.44
30-40	6	35	0.03	0.18
40-50	4	45	0.02	0.08

$$H.M = \frac{N}{\Sigma f/x} = \frac{30}{1.85} = 16.21$$ (2.5)

iii).

Class	x	f	fx	$f(x-m)^2$
0-10	5	10	50	6760
10-20	15	15	225	3840
20-30	25	25	625	900
30-40	35	25	825	400
40-50	45	10	450	1960
50-60	55	10	550	5760
60-70	65	05	325	5780

(2.5)

$$A.M = \frac{\Sigma fx}{N} = \frac{3100}{100} = 31 = m$$

(2.5)

$$SD = \sigma = \sqrt{\frac{\Sigma f(x-m)^2}{N}} = \sqrt{\frac{25400}{100}} = 15.94$$

Que. (3) i) Both are white.

(2.5)

$$P(A \cap B) = P(AB) = P(A) \times P(B)$$
$$= \frac{4}{6} \times \frac{3}{8} = \frac{1}{4}$$

b) One is black and one is white

(2.5)

$$P(A \cap B) + P(B \cap B) = P(A).P(B) + P(B).P(A)$$
$$= \left(\frac{2}{6} \times \frac{3}{8}\right) + \left(\frac{4}{6} \times \frac{5}{8}\right) = \frac{13}{24}$$

ii). Sample space $S = \{(6,3), (5,4), (4,5), (3,6)\}$

(2.5)

Let A be the event that the sum is 9. Then

$P(A) = 4/36$

Let B be the event that one of the dice show 3, when A has happened.

(2.5)

Then $P(B \cap A) = 2/36$

Now $P\left(\frac{B}{A}\right) = \frac{P(B \cap A)}{P(A)} = \frac{2/36}{4/36} = \frac{1}{2}$

iii). $E(X) = \Sigma x P(x)$ (1.5)

$= 0(0.2) + 1(0.1) + 2(0.4) + 3(0.3)$

$= 1.8$

$E(X^2) = \Sigma x^2 P(x)$ (1.5)

$= 0^2(0.2) + 1^2(0.1) + 2^2(0.4) + 3^2(0.3)$

$= 4.4$

$E(3X + 2X^2) = 3E(X) + 2E(X^2)$ (2)

$= 3(1.8) + 2(4.4) = 14.2$

Qu. (4) i).

x	0	1	2	3	4
f	8	32	34	24	5

$A.M. = \frac{\Sigma fx}{N} = \frac{192}{103} = 1.864$ (1)

By Binomial distribution $AM = np$ (1)

$p = \frac{1.864}{4} = 0.466$ (1)

$q = 1 - p = 0.534$

Expected frequency $= N(q+p)^4$ (1)

$= 103(0.534 + 0.466)^4$

ii). Given $m = 1.5$, $e^{-m} = e^{-1.5} = 0.2231$

Poisson distribution $P(x = r) = \frac{e^{-m} m^r}{r!}$ (2)

a). When neither car is used $r = 0$

$P(x = 0) = \frac{e^{-1.5}(1.5)^0}{0!} = 0.2231$ (2)

b). When some demands refused $r = 3, 4, 5, 6 \cdots \infty$

$P(3 \leq r) = 1 - [P(0) + P(1) + P(2)]$

$= 1 - [0.2231 + 0.3346 + 0.2509]$

$= 1 - 0.8086$

$= 0.1914$ (2)

iii) Let mean $m = ?$, variance $\sigma^2 = ?$, SD $\sigma = ?$

Given that $\mu_1'(10) = 40$ — (1)

$E(x-10) = 40$

$E(x) - 10 = 40$

$m - 10 = 40 \Rightarrow m = 50$ — (2)

$\mu_4'(50) = 48$ — (1)

$3\sigma^4 = 48$

$\sigma^4 = 16 \Rightarrow \sigma^2 = 4$

$\Rightarrow \sigma = 2$ — (2)

Thus mean = 50, var = 4, SD = 2

Ques (5) i) $\bar{x} = \frac{130}{10} = 13$, $\bar{y} = \frac{220}{10} = 22$

$\sigma_x = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2} = \sqrt{222.8 - 169} = 7.33$

$\sigma_y = \sqrt{\frac{\Sigma y^2}{n} - \left(\frac{\Sigma y}{n}\right)^2} = \sqrt{550.6 - 484} = 8.16$ — (1)

$r = \frac{\frac{\Sigma xy}{n} - \frac{\Sigma x \Sigma y}{n^2}}{\sigma_x \sigma_y} = \frac{346.7 - 286}{(7.33)(8.16)} = \frac{60.7}{59.81}$

$= 1.01$ — (1)

Regression line of x on y

$x - \bar{x} = r \frac{\sigma_x}{\sigma_y}(y - \bar{y})$

$x = 0.90y - 6.95$ (app) — (1)

Regression line of y on x

$y - \bar{y} = r \frac{\sigma_y}{\sigma_x}(x - \bar{x})$

$y = 1.12x + 7.44$ (app) — (1)

ii).

x	y	x^2	x^2y	x^3	x^4	xy
0	1	0	0	0	0	0
1	5	1	5	1	1	5
2	10	4	40	8	16	20
3	22	9	198	27	81	66
4	38	16	608	64	256	152
10	76	30	851	100	354	243

(2)

Now $y = a + bx + cx^2$ (2^{nd} degree parabola) — (1)

Normal eq are

$$\Sigma y = am + b\Sigma x + c\Sigma x^2$$
$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$$
$$\Sigma x^2y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$$

(1.5)

where $m = 5$

Solving normal eq we get

$a = 1.43$, $b = 0.24$, $c = 2.21$ (1.5)

Ans. $y = 1.43 + 0.24x + 2.21x^2$

iii).

X	Y	$x = X - m_x$	$y = Y - m_y$	x^2	y^2	xy
17	12	−3	−4	9	16	12
18	16	−2	0	4	0	0
19	14	−1	−2	1	4	2
19	11	−1	−5	1	25	5
20	15	0	−1	0	1	0
20	19	0	3	0	9	0
21	22	1	6	1	36	6
21	16	1	0	1	0	0
22	15	2	−1	4	1	−2
23	20	3	4	9	16	12

(2)

Here $m_x = \frac{\Sigma x}{n} = \frac{200}{10} = 20$, $m_y = \frac{\Sigma y}{n} = \frac{160}{10} = 16$ (1)

$\sigma_x = \sqrt{\frac{\Sigma x^2}{n}} = \sqrt{\frac{30}{10}} = 1.73$

$\sigma_y = \sqrt{\frac{\Sigma y^2}{n}} = \sqrt{\frac{108}{10}} = 3.28$ (2)

$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \, \Sigma y^2}} = \frac{\Sigma xy}{n \sigma_x \sigma_y}$

$= \frac{35}{(10)(1.73)(3.28)} = 0.616$ (1)

iii)

Given $\mu = 64$ inches & $\bar{x} = 66$

Qu (6) i). Step I. H_0: There is no differences i.e. $\mu = 64$
H_1: There is differences i.e. $\mu > 64$ (1)

Step II

$$S = \sqrt{\frac{\Sigma (x-\bar{x})^2}{n-1}} = \sqrt{\frac{90}{9}} = \sqrt{10}$$

$$t = \frac{(\bar{x} - \mu)\sqrt{n}}{S} = \frac{(66-64)\sqrt{10}}{\sqrt{10}}$$

$$= \frac{(66-64)(3.16)}{3.16} = \frac{66 - 202.24}{3.16}$$

$$= 2.$$ (2)

Step III Level of significance = 5%.
Step IV Degree of freedom $v = n-1 = 9$
Step V Tabular value = 1.83
Step VI Since calculated value > tabular value
then H_0 is rejected.
Conclusion Average height is greater than 64. (1)

ii) Required table –

x	$x - \bar{x}$	$(x-\bar{x})^2$
70	4	16
67	1	1
62	-4	16
68	2	4
61	-5	25
68	2	4
70	4	16
64	-2	4
64	-2	4
66	0	0

(1)

ii) Step I : H_0 : no differences $\sigma^2 = 0.16$
H_1 : differences $\sigma^2 > 0.16$ (1)

Step II

$$\chi^2 = \frac{\Sigma(x - \bar{x})^2}{\sigma^2}$$

where $\bar{x} = \frac{27.6}{11} = 2.51$, $\sigma^2 = 0.16$ (2)

then $\chi^2 = \frac{0.1891}{0.16} = 1.182$

Step III level of significance 1%.
Step IV Degree of freedom $n = n-1 = 10$
Step V Tabular value = 23.2 (1)
Step VI Since calculated value < tabular value then H_0 may be accepted

Conclusion The data are consistent with the hypothesis that the precision instrument is 0.16.

Required table — (1)

x	$x - \bar{x}$	$(x - \bar{x})^2$
2.5	−0.01	0.0001
2.3	−0.21	0.0441
2.4	−0.11	0.0121
2.3	−0.21	0.0441
2.5	−0.01	0.0001
2.7	0.19	0.0361
2.5	−0.01	0.0001
2.6	0.09	0.0081
2.6	0.09	0.0081
2.7	0.19	0.0361
2.5	−0.01	0.0001
		0.1891

iii) Step I : H_0 : Machine do not differ significantly
H_1 : Machine differ significantly

Step II

A	B
20	18
21	20
23	17
16	25
20	15

$$C = \frac{T^2}{N} = \frac{195^2}{10} = \frac{38025}{10} = 3802.5$$

$$SSC = \left(\frac{100^2 + 95^2}{5}\right) - 3802.5 = 2.5$$

$$SST = (400 + 441 + 529 + 256 + 400 + 324 + 400 + 289 + 625 + 225) - 3802.5 = 86.5$$

$$SSE = SST - SSC = 86.5 - 2.5 = 84$$

(2)

Sources of variation	Sum of Squares	DOF	Mean square	Calculation of F
Between samples	2.5	1	$MSC = \frac{2.5}{1}$	$F = \frac{MSC}{MSE}$
within errors	84	8	$MSE = \frac{84}{8} = 10.5$	$= \frac{2.5}{10.5}$
				$= 0.23$
		9		

(2)

Step III Table value $(1, 8) = 5.32$
Step IV level of significance 5%.
Step V Degree of freedom $v_1 = 1$, $v_2 = 8$
Step VI calculated value < tabular value
then accept the H_0

(1)

Conclusion: Machines do not differ significantly