Statistics for Decision Making – I

End Semester Back Paper Examination, 2025

Full Marks: 50 Time : 3 hrs

Answer as many as you like. Maximum you can score is 50

1. Observations are taken on yield of hay (x_1) in 100 cuts/ acre, spring rainfall (x_2) in inches and accumulated spring temperature (x_3) in Fahrenheit for 20 years. The following estimates are obtained.

Sample mean vector: $\bar{x} = (28.02, 4.91, 59.00)'$,

Vector of sample standard deviations: $s = (4.42, 1.10, 85.00)'$

and the correlation matrix is

$$\begin{pmatrix} 1 & 0.80 & -0.40 \\ 0.80 & 1 & -0.56 \\ -0.40 & -0.56 & 1 \end{pmatrix}$$

Find

i) The multiple linear regression equation of x_1 on x_2 and x_3.

ii) The multiple correlation coefficient of x_1 with x_2 and x_3. And hence make a performance analysis of the regression equation

iii) The partial correlation coefficient between x_1 and x_3 eliminating the effect of x_2.

(4+3+3)

2. a) Let X_1 and X_2 be a random sample of size 2 from a distribution having pdf

$$f(x) = \begin{cases} e^{-x}, 0 < x < \infty \\ 0, otherwise \end{cases}$$

Define $Y_1 = X_1 + X_2$ and $Y_2 = \frac{X_1}{X_1 + X_2}$. Show that, Y_1 and Y_2 are stochastically independent. (6)

b) In case of simple random sampling, show that, $P(y_r = Y_i) = \frac{1}{N}$, irrespective of whether the samples are drawn with replacement or without. (4)

3. a) Distinguish between ratio scale and interval scale. (3)

b) Show that, if $\overline{x_1}$ and $\overline{x_2}$ are two subgroup means, then their composite mean will lie between $\overline{x_1}$ and $\overline{x_2}$. (4)

c) What do you mean by a relative measure of dispersion? (3)

4. a) In a group of 500 persons, 220 are literate and employed, 20 are literate and unemployed and 180 are illiterate and unemployed. Is there any association between literacy and employment? (5)

b) Data from a case-control study of 200 esophageal cancer cases and 775 community-based controls are shown below. The data tries to explore relationship between alcohol consumption (dichotomized at 80 grams per day) and esophageal cancer. Use odds ratio and give your comments. (5)

Alcohol g/day	Esophageal cancer	
	+	-
+	96	109
-	104	666

5. Following are the marks obtained by a group of 43 students in a science test:

47	26	45	19	7	30	27	23	12
47	35	28	26	15	36	23	26	29
46	37	39	28	29	37	8	30	36
28	32	29	23	28	21	13	24	
37	38	22	27	32	24	20	13	

Draw a SRSWR and SRSWOR both of size 10. In each case, estimate the average marks and also estimate the corresponding standard error. (10)

6. a) Draw a suitable diagram to show the relative contributions of the different continents to the total world population: (5)

Continent	Population (in millions) in 1968
Africa	336
North America	309
South America	180
Asia	1946
Europe	455
Oceania	19

b) Distinguish between correlation and association. (2)

c) Explain sensitivity and specificity in the context of 2 X 2 contingency table. (3)