Indian Statistical Institute
M.Tech (QR&OR), First Year (E Stream): 2025-26
Semestral Examination

Statistical Methods I

Full Marks: 100 **Time:** 3.00 hours **Date of examination: 24/11/2025**

Calculator is allowed during the examination.

Answer any five questions

1. a) The weights (in pound) at birth for 15 babies born in a hospital are obtained as follows: 6.2, 6.7, 7.1, 6.9, 7.5, 5.7, 4.8, 6.8, 7.6, 7.8, 8.1, 5.0, 5.8, 7.9 and 8.5. Give two limits between which the mean weight at birth for all such babies is likely to lie. Consider the confidence coefficient is 0.99.

   b) Prove that standard deviation is the minimum root-mean-square deviation for a given set of observations.

   c) Let X be a random variable with p.d.f., $f(x) = \theta e^{-x\theta}$; $x > 0$. Show that the central confidence limits of θ (for large samples) with 95% confidence coefficient are given by: $\left(1 \pm \frac{1.96}{\sqrt{n}}\right) \times \frac{1}{\bar{x}}$.

[6 + 5 + 9 = 20]

2. a) In estimating the mean of a normal population $N(\mu, \sigma^2)$ on the basis of a random sample of size $2n + 1$, what is the asymptotic efficiency of median with respect to the mean?

   b) Suppose $X \sim N(\mu, \sigma^2)$, where μ is known. Determine the $100(1-\alpha)\%$ confidence interval of σ^2 based on a random sample of size n taken from the population of X.

   c) Obtain the rank correlation coefficient for the following data:

X	68	64	75	50	64	80	75	40	55	64
Y	62	58	68	45	81	60	68	48	50	70

[5 + 5 + 10 = 20]

3. a) Derive the expressions for expectation and standard error of sample mean in case of simple random sampling without replacement from a finite population.

   b) Show that if one of the regression coefficient (say, b_{YX}) is greater than unity, the other (say, b_{XY} must be less than unity.

   c) Let $x_1, x_2, \dots, x_n$ are the values of a random sample from the rectangular distribution with p.d.f.: $f(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta - \alpha}, & \alpha < x < \beta \\ 0, & \text{elsewhere} \end{cases}$
   Obtain the maximum likelihood estimators for α and β.

   d) State four characteristics of a good point estimator.

[(2+6) + 2 + 8 + 2 = 20]

4. a) Show that $\bar{X}$ is a sufficient estimator for μ of a $N(\mu, \sigma^2)$ population if σ^2 is known. But $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is not a sufficient estimator of σ^2, if μ is known.

b) Sample of size n_1 is collected from a $N(\mu_1, \sigma_1^2)$ population and sample of size n_2 is collected from a $N(\mu_2, \sigma_2^2)$ population. Derive the sampling distribution of $\frac{s_2^2/\sigma_2^2}{s_1^2/\sigma_1^2}$, where s_1^2 and s_2^2 are the estimators of σ_1^2 and σ_2^2 respectively.

c) Suppose $X_1, X_2, \ldots, X_n$ are random sample of size n from a $N(\mu, \sigma^2)$ population, where μ is known. Show that $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ is an MVB estimator for σ^2.

[6 + 6 + 8 = 20]

5. a) The following are the net weight (in oz) of a dry bleach product. Verify using Q-Q plot if the net weight can be assumed to follow normal distribution.

16.2	16.3	16.1	16.3	16.4	16.4	16.4	16.5	16.5	16.5	16.6
16.1	15.9	15.8	16.2	16.4	16.5	16.1	16.5	16.1	16.3	16.4
16.6	16.3	16.4	16.2	16.1	16.4	16.5	16.2	16.2	16.5	16.4
16.6	16.1	16.1	16.3	16.4	16	15.9	16.4	16.1	16.4	16.2
15.8	16.3	15.9	16.3	16.6	16.3	16.2	16.1	16.3	16.5	16.1

b) Examine whether the following results of a piece of computation for obtaining the second moment are consistent or not: $n = 120, \sum x = -125, \sum x^2 = 128$.

[17+ 3 = 20]

6. a) You are given the following incomplete frequency distribution. It is known that the total frequency is 1000 and that the median is 413.11. Estimate by calculation the missing frequencies.

Values	10-20	20-30	30-40	40-50	50-60	60-70	70-80	Total
Frequency	12	30	?	65	?	25	18	229

b) Suppose R is a random variate generated from the uniform distribution $U(0,1)$. Derive the random variate that can be generated based on the value of R for each of the following probability distribution:

i) Exponential distribution
ii) Weibull distribution
iii) Geometric distribution

c) Suppose (x_1, y_1), (x_2, y_2),..., (x_n, y_n) denote n pairs of observations. Prove that the coefficient of correlation between x and y is independent of change of origin and scale.

[4 + (3+4+4) + 5 = 20]