

STATISTICS

(Paper I)

Time Allowed : Three Hours
Maximum Marks : 300

QUESTION PAPER SPECIFIC INSTRUCTIONS

(Please read each of the following instructions carefully before attempting questions)

There are TWELVE questions divided in three Sections.

Candidate has to attempt FIVE questions in all, selecting at least ONE and not more than TWO questions from each Section.

The number of marks carried by a question/part is indicated against it.

Wherever any assumptions are made for answering a question, they must be clearly indicated.

Word limit in questions, wherever specified, should be adhered to.

SECTION 'A'

1. (a) A postman delivers 5 letters at random to 5 different addresses in a locality. What is the probability that none of the letters is correctly delivered? 15
- (b) A certain water-supply system consists of three pumping stations and a destination. Each pumping station has probability p ($0 \leq p \leq 1$) of being operable at a specified time t_0 . If the stations are operable or failed independently of one another and if the stations are connected as shown below, find the probability that water is available to the destination at time t_0 .
 [Figure: network diagram. Source (dot with arrow) on the left feeds a vertical junction line from which three parallel paths run left to right towards a vertical junction line on the right, which leads by an arrow to Destination. Top path: arrow to station 1 (circle labelled 1), then on to the right junction. Middle path: arrow from the source junction to station 3 (circle labelled 3); from station 3 two lines branch out, one going up to station 1 and one going down to station 2. Bottom path: arrow to station 2 (circle at the bottom, label partly hidden in the source; read as 2), then on to the right junction. Arrows on the right junction point from the top and bottom paths towards the Destination outlet. So water reaches the destination through station 1 or station 2; station 1 and station 2 are each fed either directly from the source or through station 3.] 25
- (c) State Bayes' theorem and explain its significance. 10
- (d) A pair of fair dice are rolled as long as necessary for a sum of two scores of 7 to turn up. Assuming that 7 does not appear on the first throw, what is the probability that more than two throws will be necessary? 10
2. (a) Let X and Y be jointly distributed with pdf

$$f(x,y) = \begin{cases} 1/4(1+xy), & |x| < 1, |y| < 1; \\ 0, & \text{otherwise} \end{cases}$$
 Find the correlation coefficient between X and Y . 20
- (b) Let X be a binomial random variable with parameters n and p . If Y is a random variable defined by $Y=3X+5$, find the probability function Y , $E(Y)$ and $V(Y)$ 20
- (c) Let X be an integer-valued random variable with probability generating function (pgf) $P(s)$. If $Y=aX+b$, where a and b are non-negative integers, then find the pgf of Y . Hence obtain $E(Y)$ and $V(Y)$. 20
3. (a) Suppose that the number of airplanes arriving at a certain airport in any 20-minute period obeys a Poisson probability law with mean 100. Determine lower bound for the probability that the number of airplanes arriving in a given 20-minute period will be between 80 and 120. 16
- (b) Examine whether the following sequences of independent r.v.'s obey the WLLN: (i) $P\{X_n = \pm 2^n\} = 1/2$ (ii) $P\{X_n = \pm \sqrt{n}\} = 1/2$ 28
- (c) State and prove Kolmogorov's inequality and explain its use. 16
4. (a) Show that $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{L} X$ 15

- (b) Show that under certain conditions (to be precisely stated) the Poisson distribution tends to the normal distribution 20
- (c) State the Lindeberg-Levy central limit theorem. If $X_1, X_2, \dots, X_n$ are i.i.d r.v.'s. with common pdf

$$f(x, \theta) = \begin{cases} \theta e^{-\theta x}, & x > 0, \theta > 0; \\ 0, & \text{otherwise} \end{cases}$$

then find the asymptotic distribution of

$$Y = 1/n \sum_{i=1}^n X_i \quad 25$$

SECTION 'B'

5. (a) State and prove Cramer-Rao inequality and explain its importance. 20
- (b) Let $X_1, X_2, \dots, X_n$ be a random sample from a Bernoulli distribution with parameter θ ($0 < \theta < 1$). Obtain uniformly minimum variance unbiased estimator of $\gamma(\theta) = \theta^2(1-\theta)$. 20
- (c) Obtain 95% confidence interval for σ^2 based on a random sample $X_1, X_2, \dots, X_n$, from $N(0, \sigma^2)$. 20
6. (a) State and prove Lehmann-Scheffe theorem and explain its significance in point estimation. 20
- (b) Show that the family of binomial distribution with parameters n and θ is complete. 20
- (c) Let $X_1, X_2, \dots, X_n$ be a random sample from a pdf
- $$f(x, \theta) = \begin{cases} e^{-(x-\theta)}, & x \geq \theta, -\infty < \theta < \infty; \\ 0, & \text{otherwise} \end{cases}$$
- Find method of moment estimator of θ . 20
7. (a) Explain the concept of an unbiased test. Examine that whether a MP test is necessarily unbiased. 16
- (b) When is a family of pdfs said to have MLR in some statistic $T(X)$? If $X_1, X_2, \dots, X_n$ is a random sample from exponential distribution with mean θ , verify whether the family of exponential distribution has MLR in some statistic $T(X)$. Obtain UMP test of size α for testing $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$. 22
- (c) Let $X_1, X_2, \dots, X_n$ be a random sample from a pdf
- $$f(x, \theta) = \begin{cases} \theta x^{\theta-1}, & 0 \leq x \leq 1, \theta > 0; \\ 0, & \text{otherwise} \end{cases}$$
- Obtain a level α UMP test of $H_0: \theta = 2$ against $H_1: \theta > 2$. Determine also its power function and sketch its graph. 22
8. (a) Let $X_1, X_2, \dots, X_n$ be i.i.d. $b(1, \theta)$ random variables. Obtain a level α Likelihood Ratio test of $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$. 20
- (b) Consider a sequence of i.i.d. $N(\mu, \sigma^2)$ σ^2 known, random variables. Derive SPRT of strength (α, β) for testing $H_0: \mu = \mu_0$ against $H_1: \mu = \mu_1 (< \mu_0)$. Also obtain its OC function. 20
- (c) Describe the Sign test for testing the location of a probability distribution. What is its parametric counterpart? 20

SECTION 'C'

9. (a) In connection with the problem of linear estimation, explain the terms 'estimable parametric function', 'error function' and 'best linear unbiased estimator (BLUE)'. 24
- (b) Consider three independent random variables Y_1, Y_2 and Y_3 having common variance σ^2 and
- $$E(Y_1) = 2\mu_1 + 3\mu_2, E(Y_2) = 3\mu_1 + 4\mu_2 \text{ and } E(Y_3) = 4\mu_1 + 5\mu_2$$
- Show that μ_1 and μ_2 are estimable. Find their best linear unbiased estimates. 36
10. (a) Using the linear regression model
- $$E(Y|x_1, x_2, x_3) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \quad 30$$

and n independent sets of observations, derive the tests for the following hypotheses (i) $H_0: \beta_1 = \beta_2 = \beta_3$ (ii) $H_0: \beta_1 + \beta_2 + \beta_3 = 1$ You may make the usual assumptions. In each case state clearly the alternative hypothesis you consider.

- (b) Explain how orthogonal polynomials are used in fitting regression curves. Develop an appropriate test for choosing the appropriate degree of orthogonal polynomial using ANOVA table. 30

11. (a) Let $X \sim N_p(\mu, \Sigma)$ and let

$$X = ([X^{(1)}; X^{(2)}]), \mu = ([\mu^{(1)}; \mu^{(2)}]), \Sigma = ([\Sigma_{11} \quad \Sigma_{12}; \Sigma_{21} \quad \Sigma_{22}])$$

Show that $X^{(1)} \sim N(\mu^{(1)}, \Sigma_{11})$ even if $X^{(1)}$ and $X^{(2)}$ are not independent. 30

- (b) Let X be a random vector with mean vector μ and dispersion matrix Σ . Prove that X is distributed as $N_p(\mu, \Sigma)$ if, and only if, for every coefficient vector l , $l'X$ is distributed as $N_1(l'\mu, l'\Sigma l)$. 30

12. (a) Let X_α ($\alpha=1,2,\dots,N$) be a random sample of size N from $N_p(\mu, \Sigma)$, where Σ is known. Develop a suitable test for testing the null hypothesis $H_0: \mu = \mu_0$, where μ_0 is a given vector. 30

- (b) Define fishers discriminant function and obtain a test of the adequacy of a hypothetical discriminant function. 30