

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 EIGHT questions divided in two Sections.

Candidate has to attempt FIVE questions in all.

Question Nos. 1 and 5 are compulsory and out of the remaining, any THREE are to be attempted choosing at least ONE question 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. Answer any five parts:
 - (a) An unbiased coin is tossed three times. Write down the sample space of this experiment. Define a random variable on this sample space and also its probability distribution. 12
 - (b) The random variables X_i ($i=1,2$) are distributed independently with p.d.f. $f_i(x)=\lambda_i e^{-\lambda_i x}$ $i=1,2$. Determine the probability $P[\min(X_1, X_2) \geq t]$ 12
 - (c) Find the characteristic function of $Y=aX+b$ where X has the p.i.f. $f(x)=\lambda^n/\Gamma(n)e^{-\lambda x}x^{n-1}$ $0 < x < \infty, \lambda > 0$ 12
 - (d) Given a random sample of n observation from a Poisson distribution $P(x; \lambda)$, derive an estimate of λ based on the frequency of $x=0$ and work out its asymptotic variance. 12
 - (e) Explain the terms prior and posterior distributions. How are prior distributions selected? What is the role of conjugate prior distribution? 12
 - (f) Indicate the use of Wald-Wolfowitz run test and state the test-statistic along with its mean and variance under the null hypothesis. 12
2.
 - (a) X follows an exponential distribution with parameter λ . Prove the following relations: (i) $P(X > x_1 + x_2) = P(X > x_1)P(X > x_2)$ (ii) $E(X) = E(X - x | X \geq x)$ 20
 - (b) State and prove Chebyshev's inequality. The random variable X takes values $-a, 0, a$ with probabilities $1/8, 3/4, 1/8$ respectively. Compute $P[|X| \geq 2\sigma_x]$ and compare it with Chebyshev's bound. 20
 - (c) $X_1, X_2, \dots, X_n$ are random samples from a uniform distribution over $(0, 1)$. Find the p.d.f. of $U = (\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n)^{1/n}$. 20
3.
 - (a) Let $X_1, X_2, \dots, X_n$ be i.i.d. rectangular $(\theta - 1/2, \theta + 1/2)$ random variables. Is $T = (X_{(1)}, X_{(n)})$ sufficient for θ ? Is it complete? Give reasons for your answer. 20
 - (b) Given a random sample from $N(\mu, \sigma^2)$, where both μ and σ^2 are unknown, find the uniformly most accurate unbiased confidence interval σ^2 . 20
 - (c) Show that if T_1 and T_2 are UMVUE's of their expectations, then $T_1 T_2$ is UMVUE of $E(T_1 T_2)$. 20
4.
 - (a) Let $X_1, X_2, \dots, X_n$ be i.i.d. random variables with p.d.f. $f(x)$ having median θ . Show that for testing $H_0: \theta = 0$ against alternatives $\theta > 0$, the sign test is uniformly most powerful 20
 - (b) Let (x_i, y_i) $i=1, 2, \dots, n$ be a random sample from a bivariate normal population with parameters $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$. Find a UMPU test for $H_0: \rho = \rho_0$ vs. $\rho \neq \rho_0$. 20

- (c) Briefly explain Wald's SPRT. Also state its merits and disadvantages. 20

SECTION 'B'

5. Answer any five parts:
- (a) What is multicollinearity? Give the methods of overcoming this problem. 12
 - (b) Describe Tukey's test for non-additivity for a two-way layout with one observation per cell. 12
 - (c) Derive the Shapiro-Wilk statistic for testing normality of errors in a regression model. 12
 - (d) Explain Lahiri's method for selection of a PPS sample. 12
 - (e) Describe a method for detecting heteroscedasticity in a model. 12
 - (f) Give the ANOVA table for a 2^4 experiment in blocks of 8 plots each, such that no main effect or first order interaction is confounded. 12
6. (a) Consider the model
- $$E(y_1) = 2\beta_1 + \beta_2, E(y_2) = \beta_1 - \beta_2, E(y_3) = \beta_1 + \alpha\beta_2$$
- with the usual assumptions. Determine α so that the BLUE's of β_1 and β_2 are uncorrelated. 20
- (b) Explain the method of generalized least squares for estimating the parameters of a linear model, stating the assumptions underlying it. 20
 - (c) In order to test the equality of mean vectors of $N_p(\mu_i, \Sigma)$ $i=1,2$ define Hotelling's T^2 -statistic and sketch the derivation of its null distribution. 20
7. (a) Discuss how one can unbiasedly estimate the population variance using stratified simple random sampling. 20
- (b) Let Π_i and Π_{ij} denote respectively the inclusion probability of population unit i and the joint inclusion probability of two distinct units i and j in a PPSWOR sample of size 2. Express Π_i and Π_{ij} in terms of selection probabilities and verify that $\sum_1^N \Pi_i = 2$ and that $\Pi_i \Pi_j \geq \Pi_{ij}$ where N is the population size. 20
 - (c) Briefly explain the idea behind Randomized Response Technique and indicate its possible uses. 20
8. (a) Explain the concept of balance in the context of a confounded factorial experiment. 20
- (b) Define an m -ple lattice design with its parameters and give its analysis. 20
 - (c) What are 'mixed effects' models? Indicate the analysis of two-way classified data with equal number of observations per cell, using a mixed-effects model. 20