

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) (i) For n events $A_1, A_2, \dots, A_n$ show that

$$\sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j) \leq P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i)$$
 (ii) Let $\{A_n\}$ be a non-decreasing sequence of events. State that

$$\lim_{n \rightarrow \infty} P(A_n) = P\left(\lim_{n \rightarrow \infty} A_n\right)$$
26
- (b) An urn contains R red and W white balls. Balls are drawn one by one without replacement. If A_k is the event that a red ball is drawn for the first time on the k^{th} draw, then determine $P(A_k)$. Let p be the proportion of red balls in the urn before the first draw. Show that

$$P(A_k) \rightarrow p(1-p)^{k-1} \text{ as } R+W \rightarrow \infty$$
20
- (c) Explain clearly why there must be a mistake in each of the following statements: (i) The probability that a student will get an 'A' in a statistics course is 0.32 and the probability that he or she will get either an 'A' or a 'B' is 0.27. (ii) A company is working on the construction of two shopping centers. The probability that the larger one will be completed on time is 0.35 and the probability that both will be completed on time is 0.42.
 14
2. (a) State and prove Bayes' theorem.
 20
- (b) If two random variables have the joint density

$$f(x_1, x_2) = \begin{cases} [x_1 x_2] & \text{for } 0 < x_1 < 1, 0 < x_2 < 2 \\ 0 & \text{elsewhere} \end{cases}$$
 find the probabilities that (i) both random variables will take on values less than 1. (ii) the sum of the values taken on by both will be less than 1.
 20
- (c) Show that under suitable conditions the binomial distribution approaches the Poisson distribution.
 20
3. (a) State and prove Markov's theorem for WLLN.
 20
- (b) Verify the following: (i) $X_n \xrightarrow{P} X, X_n \xrightarrow{P} Y \Rightarrow P[X=Y]=1$ (ii) $X_n \xrightarrow{r} X$, for some $r > 0 \Rightarrow X_n \xrightarrow{P} X$
20
- (c) (i) State Kolmogorov's SLLN for a sequence of independent random variables. (ii) Verify whether the following sequence of independent random variables obey the SLLN:

$$P\{X_k = 1/2^k\} = P\{X_k = -1/2^k\} = 1/2, k=1, 2, \dots$$
20
4. (a) Define the probability generating function of a random variable. Find the pgf of a geometric distribution and hence find its mean and variance.
 20
- (b) State the properties of a characteristic function and show that it is uniformly continuous.
 15
- (c) State and prove the Lindeberg-Levy central limit theorem.
 25

SECTION 'B'

5. (a) State and prove Rao-Blackwell theorem. State its importance. 25
- (b) Let (T_n) be a sequence of estimators such that $E(T_n) \rightarrow \theta$ and $\text{Var}(T_n) \rightarrow 0$ as $n \rightarrow \infty$. Show that T_n is a consistent estimator for θ . 15
- (c) Let T_0 be an MVU estimator of $g(\theta)$ and T_1 an unbiased estimator with efficiency $e_\theta < 1$. Then can any unbiased linear function of T_0 and T_1 be an MVU estimator of $g(\theta)$? Justify your answer. 20
6. (a) Let θ be a real parameter and the random variable x have probability density with MLR in $T(x)$. Determine the size α UMP test of $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$. 30
- (b) Let X_i ($i=1,2,\dots,n$) be a random sample from a rectangular distribution $r(0,\theta)$. Show that for testing $H_0: \theta = \theta_0$ against $H_1: \theta > \theta_0$, any critical region W for which $P_{\theta_0}(W) = \alpha$ and which necessarily includes the sample point if $\max(x_1, \dots, x_n) > \theta_0$ is UMP at level α . 30
7. (a) Let X be a $N(\mu, \sigma^2)$ variable, where both the parameters are unknown. Determine the LR test of $H_0: \mu = \mu_0$ against $H_1: \mu \neq \mu_0$ based on a sample of size n . Determine the power function of the test. 30
- (b) Let the random variable X have the pmf $f_\theta(x) = \theta^x(1-\theta)^{1-x}$, $x=0,1$ and $0 < \theta < 1$. Derive the SPR test of $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ [?]. Determine the OC function of the test. 30
8. (a) Define the empirical distribution function and find its expectation and variance. Hence show that it is a consistent estimator of the population distribution function. 20
- (b) Describe the Kolmogorov test for goodness of fit. Compare this test with the chi-square test for goodness of fit. 20
- (c) Obtain a $100(1-\alpha)\%$ confidence interval for the population median based on the sign test. 20

SECTION 'C'

9. (a) With the usual Gauss-Markov linear model, define the estimation space and the error space. Obtain expressions for the SS due to error and the SS due to all possible BLUEs of parametric functions. Find the expectation of the former. 30
- (b) Let $l_1'\beta$ and $l_2'\beta$ be the LS estimators of two estimable functions $l_1'\beta$ and $l_2'\beta$ respectively. Obtain the variances and covariances of the estimators. 30
10. (a) Given two regression lines
- $$Y_{ki} = \beta_k X_i + e_{ki} \quad (k=1,2; i=1,2,\dots,n)$$
- show that the F-statistic for testing $H: \beta_1 = \beta_2$ can be put in the form
- $$F = ((\beta_1 - \beta_2)^2) / (2S^2(\sum_1^n x_i^2)^{-1})$$
- Write the expression of S^2 . 20
- (b) In the multiple linear regression model with $(p-1)$ independent variables, derive a test to find out whether the regression on the regressor variables is significant or not. Show that the test statistic can be expressed in terms of the sample multiple correlation coefficient. 25
- (c) Write down the model of a complete three-way classified data with equal number of observations per cell. Write down the corresponding ANOVA table. 15
11. (a) Let $Y = (Y_1, Y_2, \dots, Y_n)'$ be a vector of random variables with pdf $K^{-1} \exp[-1/2(y-\theta)' \Sigma^{-1}(y-\theta)]$. (i) Show that $K = (2\pi)^{n/2} |\Sigma|^{1/2}$ and (ii) Determine $E[Y]$ and $D[Y]$. 30
- (b) Show that under usual normal population regression model the joint distribution of the partial regression coefficients is multivariate normal. 30
12. (a) Define Mahalanobis' distance between two populations as estimated from the sample on the basis of p characters. State how the D^2 -statistic is used to test the hypothesis of no difference in mean values of the p characters. 30

- (b) Define Fisher's discriminant function between two p-variate populations. Indicate the connection between the discriminant function coefficients and the D^2 -statistic. Describe a test to judge the adequacy of a preassigned discriminant function.

30