

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) An urn contains 5 white and 3 black balls, a second urn contains 6 white and 5 black balls. A ball is randomly transferred from the first urn to the second and then from the second to first urn. If a ball is now selected randomly from the first urn, find the probability that it is white. 15
- (b) Describe a priori probabilities, likelihoods and posterior probabilities in defining the Bayes theorem. Explain why, in spite of its easy deductibility from the postulates of probability, it has been the subject of such extensive controversy. 10
- (c) A sportsman's chance of shooting an animal at a distance s ($>r$) is r^2/s^2 . He fires when $s=2r$ and if he misses he reloads and fires when $s=3r, 4r, \dots$. If he misses at distance nr , the animal escapes. Calculate the odds against the sportsman. 15
- (d) (i) For n events $A_1, A_2, \dots, A_n$ show that

$$P(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$$
 (ii) If p be the probability that none of the n events occurs, then

$$p \leq e^{-\sum_{i=1}^n P_i}$$
20
2. (a) (i) What is the relationship between standard deviation and mean deviation of the data 11, 15, 19 and 23 [?]? (ii) if X and Y are symmetric show that

$$E(X/(X+Y)) = 1/2.$$
10
- (b) Let X and Y be jointly distributed with p.d.f.

$$f(x,y) = \begin{cases} [x+y, & 0 < x < 1, 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$
 Find marginals conditional distributions and correlation coefficient. 20
- (c)(i) Discuss discrete and continuous random variable with examples. Indicate the distributions for the following moment generating functions

$$M_1(t) = ((1+e^t)^5)/32, M_2(t) = e^{1/4(e^t-1)}, M_3(t) = e^{2t+1/2t^2}$$
20
- (c)(ii) If the probability of hitting a target is $1/3$ and if 6 shots are fired, what is the conditional probability of the target being hit at least twice assuming that at least one hit is already scored? 10
3. (a) State and prove Kolmogorov's inequalities and explain its uses. 20
- (b) Verify the following: (i) $X_n \xrightarrow{a.s.} C \Rightarrow X_n \xrightarrow{P} C$ (ii) $X_n \xrightarrow{q.m.} C \Rightarrow X_n \xrightarrow{a.s.} C$ 20
- (c) State and prove the Borel-Cantelli lemma and express its significance. 20
4. (a)(i) Do the following functions define distribution functions of random variables?

$$F(x) = \begin{cases} 0, & x < -a \\ 1/2(x/a+2), & -a \leq x \leq a \\ 1, & x > a \end{cases}$$
10

- (a)(ii) $F(x) = \begin{cases} 0, & \text{if } x \leq 1; \\ k/4(x-1)^3, & \text{if } 1 < x \leq 3; \\ 1, & \text{if } x > 3 \end{cases}$
Justify your answer in each case. 10
- (b) Define negative binomial distribution and state its deductions. What will be the limiting distribution? 20
- (c) Explain central limit theorem and weak law of large numbers for a sequence $\{X_n\}$ of i.i.d random variables. State Lindeberg-Levy theorem. 20

SECTION 'B'

5. (a)(i) Define unbiasedness and consistency of the estimator. Obtain unbiased estimate of σ^2 [?] from $N(\mu, \sigma^2)$ if μ is known. 15
- (a)(ii) Does a sufficient statistics for θ exists in

$$f(x, \theta) = 1/2e^{-|x-\theta|}, -\infty < x < \infty.$$
 5
- (b) Let M_0 is an MVUE of $\gamma(\theta)$ and M_1 is any other unbiased estimator of $\gamma(\theta)$ with efficiency $e < 1$. Can any unbiased linear combination of M_0 and M_1 be an MVUE of $\gamma(\theta)$? Justify your answer, 20
- (c)(i) State the conditions under which maximum likelihood estimators are identical with those given by method of moments. 10
- (c)(ii) $f(x, \theta) = e^{-(x-\theta)}, \theta \leq x < \infty, -\infty < \theta < \infty$, then

$$P[X_{(1)} - 1/n \log \alpha \leq \theta \leq X_{(1)}] = 1 - \alpha$$
 10
6. (a) Let (x_1, x_2) be the two independent observations from

$$f(x, \theta) = 1/\theta e^{-x/\theta}, 0 < x < \infty, \theta > 0.$$

For testing $H_0: \theta = 2$, against the alternative $H_1: \theta = 4$, the H_0 is accepted when $x_1 + x_2 < 9.448$. Determine UMP test.

$$\{P[\chi_4^2 \geq 9.448] = 0.05, P[\chi_4^2 \geq 0.711] = 0.95\}$$
 25
- (b) Define MP and UMP region. Compare them and show that a MP test is necessary unbiased. 15
- (c) Describe real valued function $T(x)$ and the families with MLR. Obtain UMP test for $H_0: \theta \geq \theta_0$ against $H_1: \theta < \theta_0$ 20
7. (a) State the properties of L-R test and suggest the conditions where $-2 \log_e \lambda$ has the asymptotic chi-square distribution. 25
- (b) Let $(x_1, x_2, \dots, x_n)$ be a random sample from $N(2, \sigma^2)$ [?]. Determine the L-R test of $H_0: \sigma^2 = \sigma_0^2$ against $H_1: \sigma^2 \neq \sigma_0^2$ Examine, whether the test is unbiased. 15
- (c) Consider a sequence of i.i.d variable with pdf $f(x, \theta) = \theta e^{-\theta x}, x \geq 0, \theta > 0$. Derive SPRT of strength (α, β) for testing $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$. Also obtain its OC function. 20
8. (a) Explain the main difference between the parametric and non-parametric approaches to the theory of statistical inference. Specify the underlying assumptions and the null hypothesis in median, sign and U tests. 25
- (b) Describe chi-square and Kolmogorov test for goodness of fit and compare. 15
- (c) Discuss complete sufficient statistics and suggest the method to find MVUE for $e^{-\lambda} P(\lambda)$. 20

SECTION 'C'

9. (a) Discuss least square estimate and show that least-square estimates are the same as the best linear estimates in linear estimation. 20
- (b) Let $y_i = m + bx_i + e_i$ ($i=1, 2, \dots, n$), find estimate of m and b and express $\text{Var}(b)$. 20

- (c) In the case of two-way classified data with unequal number of observations per cell derive the expectation of the MS components of the ANOVA table. 20
10. (a) Define contrasts, orthogonal contrasts and show that SS due to n observations is equal to the total SS due to the orthogonal contrasts. 20
- (b) For the model $y = X\beta + e$, we have $\hat{\beta} = (X'X)^{-1}X'y$, $E(\hat{\beta} - \beta)(\hat{\beta} - \beta)' = \sigma^2(X'X)^{-1}$, then test statistic $F = Q_2(\beta_0)/p/Q_1/(n-p)$ for $H_0: \beta = \beta_0$, where $Q_1 = (y - X\hat{\beta})(y - X\hat{\beta})'$, $Q_2 = (X'e)'(X'X)^{-1}(X'e)$, $e = (y - X\hat{\beta})$, find the distribution of F and state its properties. 25
- (c) In the multiple linear regression model with normality assumption obtain $100(1-\alpha)\%$ confidence interval to find out whether the regression on the regressor variables is significant or not 15
11. (a)(i) If $X \sim N(\zeta, \sigma^2)$ then $a'X \sim N(a'\zeta, a'Ba)$ find $B = TT'$ when $T\tilde{y} = (X - \zeta)$. 5
- (a)(ii) If the probability density function of a random vector X is $k \exp\{-1/2(X - \zeta)' \Sigma^{-1}(X - \zeta)\}$ determine $E(X)$ and $V(X)$. 10
- (b) Consider the sets of random variable $X_1, X_2, \dots, X_q$ and $X_{q+1}, X_{q+2}, \dots, X_p$ with vectors $X^{(1)} = ([X_1; X_2; \vdots; X_q])$, $X^{(2)} = ([X_{q+1}; X_{q+2}; \vdots; X_p])$ and $\mu = ([\mu^{(1)}; \mu^{(2)}])$, $\Sigma = ([\Sigma_{11} \ 0; 0 \ \Sigma_{22}])$
- The marginal distribution of the set is multivariate normal with means, variances and covariances obtained by taking the proper components of μ and Σ respectively. 25
- (c) Define partial correlation coefficients in multivariate analysis. Show that
- $$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$$
- 20
12. (a) T^2 statistics is the generalization of t-statistic. Explain its relationship with D^2 -statistic. Explain the uses of T^2 statistic in test of hypothesis. 30
- (b) Let $N_p(\mu_1, \Sigma)$ and $N_p(\mu_2, \Sigma)$ and linear discriminant function $L(U) = (\mu_2 - \mu_1)' \Sigma^{-1} U$ is sufficient for the class of normal densities $N_p(\alpha\mu_1 + \beta\mu_2, \Sigma)$, $\alpha + \beta = 1$, prove it. Discuss the methods used in it. 30