

STATISTICS

(Paper I)

Time Allowed : Three Hours
Maximum Marks : 250

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.

Attempts of questions shall be counted in sequential order. Unless struck off, attempt of a question shall be counted even if attempted partly. Any page or portion of the page left blank in the Question-cum-Answer Booklet must be clearly struck off.

SECTION 'A'

1. (a) If for the events A and B , $P(A) > 0$ and $P(B) > 0$ then show that $P(A|B) > P(A)$, if and only if $P(B|A) > P(B)$. Likewise, $P(A|B) < P(A)$, if and only if $P(B|A) < P(B)$. 10

- (b) Three digits 1, 2 and 3 are written down in random order. What is the probability that at least one digit will occupy its proper place? 10

- (c) Construct a 95% confidence interval for μ in $N(\mu, 4)$ from the following observed sample: -5, 0, 2, 15. 10

- (d) Let X_1 and X_2 be i.i.d. random variables and each has probability density function

$$f(x; \theta) = 1/\theta \exp(-x/\theta), x \geq 0, \theta > 0; = 0, \text{ otherwise.}$$

If $U_1 = 0.6 X_1 + 0.4 X_2$ and $U_2 = X_1 + X_2$, decide which one of U_1 and U_2 are sufficient statistics for θ . If $h(u_2) = E(U_1 | u_2)$, then show that it has smaller variance compared to $\text{Var}(U_1)$. 10

- (e) A tourist resort is visited by tourists from many countries. The resort operator has the following data in the month of January, 2014:

Country	USA	UK	Canada	Italy	Germany	France	Japan
No. of tourists	22	12	18	10	20	18	30

The resort operator has a hypothesis that the proportion of tourists visiting in the month of January of any year is 2 : 1 : 2 : 1 : 2 : 2 : 3. Test this hypothesis at 5% level of significance (Given $\chi^2_{6; 0.05} = 12.59$). 10

2. (a) The probability density function of a random variable X is given by

$$f_X(x) = 630x^4(1-x)^4 \text{ for } 0 < x < 1, = 0, \text{ otherwise.}$$

Find the probability that X will take on a value within two standard deviations of the mean ($\mu \pm 2\sigma$) and compare it with the lower bound provided by the Chebychev's Inequality. 20

- (b) For a random variable X with probability density function

$$f_X(x; \lambda) = e^{-x} \cdot x^\lambda / \lambda! \text{ for } x > 0, = 0, \text{ otherwise;}$$

where $\lambda \geq 0$ is an integer, show that

$$P_r \{0 < X < 2(\lambda + 1)\} > \lambda / (\lambda + 1). \quad \text{15}$$

- (c) A die is rolled 15 times with the following results:

Face value	1	2	3	4	5	6
Frequency	0	1	4	0	4	6

15

Use Kolmogorov-Smirnov statistic to test whether the die is unbiased or not. (Given $D_{15;0.05}=0.304$)

3. (a) Suppose n items are put on test simultaneously and the test is continued until r items fail. Assuming an exponential failure distribution with mean life time θ , obtain the maximum likelihood estimator for θ and hence estimate $P(X \geq t)$. Also obtain the Fisher information about the parameter θ and show that the estimator is asymptotically normal. 20
- (b) Let $X \sim N(\mu, 1)$ and the prior distribution of μ is $N(0, 1)$. Assuming squared error loss function, obtain the Bayes estimator for μ . Also obtain Bayes risk. 15
- (c) Define convergence in probability and convergence in distribution of a sequence of random variables. Show that convergence of X_n to X in probability implies convergence of X_n to X in distribution. Is the converse also true? 15
4. (a) Let $X_1, X_2, \dots, X_n$ be a random sample drawn from a uniform $U(0, \theta)$ distribution. Obtain a UMP test of size α for testing $H_0: \theta = \theta_0$ against $H_1: \theta \neq \theta_0$. 20
- (b) Prove the following results: (i) A sequential probability ratio test (SPRT) always terminates with probability one. (ii) For a SPRT with $Z = \log f(x|H_1)/f(x|H_0)$ such that $P_r\{|Z| > 0|H\} > 0$, then prove that
- $$E_H\{e^{tS_N}[M(t)]^{-N}\} = 1 \text{ when } S_N = \sum_{i=1}^N Z_i. \quad \text{15}$$
- (c) Let the probability function of X and Y be
- $$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x=0(1)n; \quad P(Y=y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}, \quad y=r, r+1, \dots$$
- Prove that $P[X \geq r] = P[Y \leq n]$. 15

SECTION 'B'

5. (a) Let Y_1, Y_2, Y_3 be three independent observations having expectations
- $$E(Y_1) = \beta_0 - \beta_1 + \beta_2, \quad E(Y_2) = \beta_0 - 2\beta_2, \quad E(Y_3) = \beta_0 + \beta_1 + \beta_2 \text{ and } V(Y_i) = \sigma^2 \text{ for } i=1,2,3.$$
- Obtain least square estimates of β_0, β_1 and β_2 . Can you obtain unbiased estimate of σ^2 ? Justify your claim. 10
- (b) If π_1 and π_2 denote two bivariate normal populations where $\pi_1 \sim N_2(\mu_1, \Sigma)$ and $\pi_2 \sim N_2(\mu_2, \Sigma)$, $\mu'_1 = [10, 15]$, $\mu'_2 = [10, 25]$ and $\Sigma = \begin{bmatrix} 18 & 12 \\ 12 & 32 \end{bmatrix}$. Compute the Mahalanobis distance between π_1 and π_2 . 10
- (c) For a population size 10, show that in SRS without replacement the probability of drawing a specified unit at 5th draw is equal to the probability of drawing it at the first draw. 10
- (d) Consider a RBD with 4 treatments arranged in 5 blocks. Show that RBD is orthogonal. 10
- (e) Let X_1, X_2, X_3, X_4 and X_5 be independent and identically distributed random vectors with mean vector μ and covariance matrix Σ . Find the distribution of $Y = X_1 - X_2 + X_3 - X_4 + X_5$. 10
6. (a) Let $X \sim N_p(\mu, \Sigma)$, where
- $$X = \begin{bmatrix} X_{q \times 1}^{(1)} \\ X_{p-q \times 1}^{(2)} \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_{q \times 1}^{(1)} \\ \mu_{p-q \times 1}^{(2)} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$
- Obtain the conditional distribution of $X^{(1)}$ given $X^{(2)} = x^{(2)}$. 20
- (b) Consider the area under wheat for a sample of 44 clusters selected from 11 different villages. Four clusters were selected from each of the 11 villages and each cluster consists of 8 consecutive survey numbers (fields). Here sum of squares (SS) due to between villages is 2000, SS due to between clusters within villages is 8250 and Total SS is 30000. Write the ANOVA table. 15
- (c) Construct an LSD of size 5. Delete one column from this LSD. Prove that the resulting design is a symmetrical BIBD with parameters $v=b=5, r=k=4$ and $\lambda=3$. 15
7. (a) Explain Horvitz-Thomson estimator of the population mean. Show that H-T estimator is unbiased estimator of population mean. Also obtain its variance. 20

- (b) Define Hotelling T^2 . State any one of its applications. A random sample of size 3 from a bivariate normal $N_2(\mu, \Sigma)$ distribution gave following unbiased estimates of μ and Σ

$$\bar{x} = ([8; 6]), S = ([4 \ -3; -3 \ 9])$$

Compute T^2 -statistic to test the hypothesis $H_0: \mu = \mu_0 = ([7; 7])$.

15

- (c) Construct a key block of 2^7 confounded factorial experiment into a block of size 8 by choosing suitable interactions only. Write all the independent and generalized confounded interactions. Further obtain two or more blocks using key block.

15

8. (a) Consider a BIBD with parameters $v=11=b$, $r=k=5$ and $\lambda=2$. Obtain c matrix of this design and its non-zero eigen value. Hence obtain $v(\hat{f}_i - \hat{f}_m)$, $i \neq m=1, \dots, 11$.

20

- (b) 200 boys and 100 girls of a college appeared in an examination. Means and variances of their scores are as given below:

Category	No. of Students N_i	Mean Marks $\bar{Y}_{N_i}$	Variance $\sigma_{N_i}^2$
Boys	200	40	10
Girls	100	50	20

How will you draw a random sample of size 30 using proportional allocation? Hence obtain the variance of the estimator of the population mean.

15

- (c) In the Gauss-Markov set up $(Y, X\beta, \sigma^2 I)$ a least square estimate is given by a solution of the system:

$$X'X\beta = X'Y$$

- (i) Justify the above statement. (ii) Establish that the above system of equations is always consistent.

15