

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.

All questions carry equal marks.

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 four parts among parts (a) to (e).

- (a) Choose the correct answer for (i), (ii) and (iv) and answer (iii). (i) A, B and C are three arbitrary events. The expression for the event that "only A occurs" in terms of A, B, C is $[(A \cap B \cap C), (A \cap \bar{B} \cap \bar{C}), (A \cap \bar{B} \cap C), (A \cap B \cap \bar{C})]$ (ii) A coin is tossed three times in succession. If A is an event that there are at least two heads and B is the event that 1st throw gives head, the probability $P(A \cap B)$ is..... [1/8, 5/8, 3/8] (iii) State Neyman-Pearson Lemma for testing a simple hypothesis against a simple alternative. (iv) An estimate T_n based on a sample of size n is called an asymptotically unbiased estimate of θ if.....
 $\lim_{n \rightarrow \infty} P(T_n = \theta) = 1, E(T_n) = \lim_{n \rightarrow \infty} \lim_{n \rightarrow \infty} E(T_n) =$

- (b) Match X with Y:

X	Y
(i) Random experiment	(a) Characteristic function
(ii) Sufficient	(b) Non-parametric method
(iii) Inversion theorem	(c) Probability theory
(iv) Kolmogorov test for goodness of fit	(d) Estimate

- (c) Supply the missing word(s) (i) If A and B are independent events then the conditional probability $P(B/A)$ is equal to..... (ii) If $p=q=1/2$, then the binomial distribution is (iii) If a hypothesis specifies the value of all the parameters of a density function, it is called (iv) Methods which are not concerned with testing or estimating the parameters of a density function of a given type are called.....
- (d) (i) Obtain the probability generating function of the Poisson distribution. (ii) Write the probability mass function of the negative binomial distribution. (iii) Define a Uniformly Most Powerful Test (UMPT). (iv) Write the expression for ASN function in sequential probability ratio test (SPRT).
- (e) (i) Obtain the mean and the variance of the exponential distribution. (ii) Write down the reproductive property of the Poisson distribution. (iii) If x and y have unit variances and correlation 0.4, find the correlation between $x+2y$ and $2x+y$ (iv) When is a statistic called the best unbiased estimate of a parameter?

2. (i) A random experiment consists in taking two balls out of a bag containing 3 white and 2 black balls. Write down the corresponding sample space, enumerating all the elementary events and their probabilities.

- (ii) 8 unbiased coins are tossed once. Find the probability that there are at least 7 heads.

- (iii) Obtain estimates of the parameters μ and σ of a Normal distribution using the method of moments
- (iv) An unbiased coin is tossed 30 times and the following sequence of heads (H) and tails (T) is observed: HHHTTTHTTTTHHHTTHHHTTTTHHTHT Find out the number of runs (r). Also find the mean $E(r)$.
3. (i) An unbiased coin is tossed till 3 heads are obtained. Find the expected value of the number of tosses.
- (ii) State the Weak Law of Large Numbers and write necessary and sufficient conditions for its validity.
- (iii) A bag contains four red and six black balls. One ball is taken out a random, without noting its colour. A second ball is then taken out of the remaining balls in the bag at random and found to be black. What is the conditional probability that the first ball taken out was also black?
- (iv) What are the advantages and disadvantages of non-parametric inference procedures over corresponding parametric procedures?
4. (i) Show that $2s^2 = \sum (x_i - \bar{x})^2 / (n-1)$ is an unbiased estimate of the population variance, where x_i for $i=1,2,\dots,n$ are observed values in SRSWR sampling and $\bar{x}$ is the sample mean.
- (ii) Define Uniformly Minimum Variance Unbiased Estimator of a parameter. Explain the role of Lehmann-Scheffe Theorem in this context.
- (iii) Write the probability density function of Beta distribution of first kind. Find the value of its r-th moment about origin.
- (iv) Obtain the Moment Generating Function of the standard Normal distribution.

SECTION 'B'

5. Answer any four parts among parts (a) to (e)
- (a) Choose the most appropriate answer: (i) The main reason for forming blocks in experimental designs is (error control, replication, increasing inductive scope) (ii) The correlation coefficient of y on x is 0.5. This means the proportion of the variability in p accounted for by x is (75%, 50%, 25%) (iii) In a design, the main effects are confounded with the block effects. (Split Plot, Latin Square Randomized Blocks) (iv) In a sample of size n, drawn from a population of size N, drawn by the method of simple random sampling without replacement, the probability that a particular unit is included is (n/N , N/n , $1/({}^N C_n)$)

- (b) Match X with Y

X	Y
(i) Replication	(a) Stratified sampling
(ii) Neyman allocation	(b) Analysis of variance
(iii) Variance covariance matrix	(c) Design of experiment
(iv) One-way classification	(d) Multivariate analysis

- (c) Fill up the blanks in parts (i), (ii) and (iii) and answer (iv). (i) The arithmetic mean of a positive valued statistical variable is than the reciprocal of the arithmetic mean of the reciprocal of the variable. (ii) If one observation is missing in a Randomized Block Design with v treatments and b blocks, in the Analysis of Variance Table, the number of degrees of freedom for treatments will be and that for error will be (iii) The Gauss-Markov Theorem requires the error components to have mean values, variance and all pair-wise covariance's to be (iv) Write down the structure of the Analysis of Variance Table for a Latin Square Design of order m.

- (d) (i) Write down an approximation for the variance of the usual estimate of the ratio of the means of two correlated variables, in the case of SRSWR sampling. (ii) Write down the variance of the Horvitz-Thomson unbiased estimate of the population total. (iii) Write the linear statistical model for the analysis of Randomized Block Design. (iv) Write down the partial correlation coefficient between the first and second variables, eliminating the effect of the third variable from each of them, in terms of the three pair-wise total correlation coefficients.
- (e) (i) Define two-stage sampling. (ii) Describe the use of critical differences in comparison of pair-wise differences amongst treatment effects. (iii) If regression coefficient of y on x is -0.5 and regression coefficient of x on y is -0.4 , find the value of correlation coefficient. (iv) Explain the concept of confounding in factorial experiments.
6. (i) Calculate the expectation and the variance of the sample mean in SRSWOR sampling.
- (ii) Construct a Latin Square Design of order 5.
- (iii) Describe an appropriate test for equality of means of several populations on the basis of independent samples drawn from them.
- (iv) Interpret the Analysis of Variance Table in two-way mixed effects model with equal number of observations in each cell.
7. (i) Explain how the basic principles of randomization, replication and error control are utilized in a Randomized Blocks Design.
- (ii) What is the difficulty in analyzing a Latin Square Design of size 2?
- (iii) Give your critical comments on the following observation "A great disadvantage of systematic sampling is that the variance of the estimated mean does not necessarily decrease as the sample size is increased."
- (iv) Show that the correlation matrix is at least positive semi-definite.
8. (i) Explain Fixed Effect and Mixed Effect models in two-way classification with examples.
- (ii) For a BIBD (v, b, r, k, λ) , prove that $\lambda(v-1) = r(k-1)$
- (iii) If $x = 4y + 5$ and $y = (1/16)x + 4$ are the regression lines of x on y and y on x respectively, find the value of the 4 correlation coefficient between x and y.
- (iv) In standard notations, given that $r_{12} = 0.5$, $r_{13} = 0.4$ and $r_{23} = 0.2$. calculate $R_{1.23}$