



STATISTICS

TIME ALLOWED: THREE HOURS	(PART-I MCQs) MAXIMUM MARKS: 20
PART-I (MCQs) : MAXIMUM 30 MINUTES	(PART-II) MAXIMUM MARKS: 80
NOTE: (i) First attempt PART-I (MCQs) on separate OMR Answer Sheet which shall be taken back after 30 minutes. (ii) Overwriting/cutting of the options/answers will not be given credit. (iii) There is no negative marking. All MCQs must be attempted.	

PART-I (MCQs)(COMPULSORY)

Q.1. (i) Select the best option/answer and fill in the appropriate Box ☐ on the **OMR Answer Sheet.(20x1=20)**
(ii) Answers given anywhere else, other than OMR Answer Sheet, will not be considered.

- Which of the following best describes what statistics deals with?
(A) Individuals (B) Particular facts (C) Isolated Items (D) Aggregative facts
- Five-point Likert scales (*Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree*) are frequently used to measure motivations and attitudes. A Likert scale is a:
(A) Discrete variable (B) Ordinal variable (C) Categorical variable (D) All of theses
- A numerical value used as a summary measure for a sample, such as sample mean, is known as a:
(A) Parameter (B) Estimator (C) Estimate (D) Statistic
- The general purchasing power of the currency of a country is determined by:
(A) Retail price index (B) Whole sale price index (C) composite index (D) Value index
- If a data set is transformed by adding or subtracting a constant c to every observation, which measure remains unchanged?
(A) Mean (B) Median (C) Range (D) Standard deviation
- The weighted arithmetic mean of the first n natural numbers whose weights are corresponding numbers is:
(A). $\frac{(n+1)(2n+1)}{6}$ (B). $\frac{n(n+1)(2n+1)}{6}$ (C). $\frac{(2n+1)}{3}$ (D). $\frac{(2n+1)}{6}$
- In a random experiment of n trials, with fixed probability of success, the probability to observe at 5th success is modeled by the following distribution:
(A) Binomial (B) Poisson (C) Uniform (D) Normal
- If A and B are two events, the probability of occurrence of either A or B is given as:
(A) $P(A)+P(B)$ (B) $P(A \cup B)$ (C) $P(A \cap B)$ (D) None of these
- The diagonal matrix $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ the determinant of D^{-1} is:
(A). $1/8$ (B). $4/8$ (C). $3/8$ (D). $6/8$
- If X is an exponential random variable with CDF $F(x) = 1 - e^{-2x}$, the mean of the random variable X is:
(A). 2 (B). 0.5 (C). -2 (D). -0.5
- Significance of partial regression coefficients can simultaneously be tested by:
(A) T-test (B) Z-test (C) Chi-square test (D) F-test
- Which of the following is useful for computing the bounds on the probability when the mean and standard deviation of the variable is known.
(A) Markov inequality (B) pdf of X (C) Chebyshev's inequality (D) Central limit theorem
- The sample standard deviation of first five natural numbers is:
(A) 1.414214 (B) 1.581139 (C) 2 (D) 4
- The CDF of the random variable X is defined by:

$$F(x) = \begin{cases} 0 & x < 2 \\ K(x-2) & 2 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$

What is the value of K?

- (A) $\frac{1}{4}$ (B) $\frac{1}{2}$ (C) 2 (D) None of these

15. Given a system of linear equation, this non-homogeneous system has:

$$\begin{aligned} x_1 + x_3 &= 3 \\ x_1 - x_2 - x_3 &= 1 \\ -x_1 + x_2 &= 4 \end{aligned}$$

- (A) No solution (B) Infinite solution (C) Unique solution (D) All of these

16. The probability associated with committing Type-II error is:
 (A) β (B) α (C) $1 - \beta$ (D) $1 - \alpha$
17. The following are the scores a professor gave on the first test in a statistics class: 52, 90, 88, 61, 75, 82, 75, 83, 88, and 86. What was the median score on this test?
 (A) 82 (B) 82.5 (C) 78.5 (D) 52
18. Which of the following is NOT a property of a consistent estimator?
 (A) As $n \rightarrow \infty$, bias $\rightarrow 0$ (B) It converges in probability to the true parameter
 (C) As $n \rightarrow \infty$, variance $\rightarrow 0$ (D) It always attains the Cramér-Rao lower bound
19. In a two-way ANOVA with factors A (3 levels) and B (4 levels), the degrees of freedom for the interaction term are:
 (A) $2 \times 3 = 6$ (B) $(3-1)(4-1) = 6$ (C) $3 + 4 - 2 = 5$ (D) $12 - 1 = 11$
20. A country's infant mortality rate (IMR) is 45 deaths per 1,000 live births. If the number of live births in the year was 250,000, how many infant deaths occurred?
 (A) 13,750 (B) 12,500 (C) 11,250 (D) 10,000

PART-II

- NOTE:** (i) Part-II is to be attempted on the separate **Answer Book**.
 (ii) Attempt **FOUR** questions in all by selecting **TWO** Questions each from **SECTION**. **ALL** questions carry **EQUAL** marks.
 (iii) All the parts (if any) of each Question must be attempted at one place instead of at different places.
 (iv) Write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper.
 (v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed.
 (vi) Extra attempt of any question or any part of the question will not be considered.
 (vii) **Use of Calculator is allowed.**

SECTION-A

- Q. No.2. (a)** Define Type I and Type II errors in hypothesis testing. Explain the trade-off between these errors. (10)
- (b)** The first four moments of a frequency distribution about the origin are (10) (20)
 $\mu'_1 = -0.55$, $\mu'_2 = 4.46$, $\mu'_3 = -0.43$ and $\mu'_4 = 68.52$.
 Find skewness and kurtosis. Also comment on the shape of the data distribution.
- Q. No.3. (a)** Explain the concept of exchangeability in probability. How does it differ from independence? (10)
- (b)** Define skewness of a distribution. What are the different measures of skewness? (10) (20)
 (i) In a frequency distribution, the coefficient of skewness based upon quartiles is 0.6. If the sum of the upper and lower quartiles is 100 and median is 38, find the values of Q_1 and Q_3 .
 (ii) Karl Pearson's coefficient of skewness of a distribution is 0.32. If the standard deviation is 6.5 and mean is 29.6, determine the mode and the median of the distribution.
- Q. No.4. (a)** Explain the difference between Parametric and Non-Parametric tests. Provide examples of each in the context of ANOVA. (10)
- (b)** A secretary makes 2 errors per page, on average. What is the probability that on the next page he or she will make (10) (20)
 (i) 4 or more errors? (ii) No error?
- Q. No. 5.(a)** Define conditional probability with examples (10)
- (b)** In a small town only three newspapers, A, B, and C, are available for the readers. (10) (20)
 Suppose that 60% of readers subscribe to newspaper A, 40% subscribe to newspaper B, and 30% to newspaper C. Suppose also that 20% of them subscribe to both A and B, 10% subscribe to both A and C, 20% subscribe to both B and C, and 5% subscribe to all three newspapers.
 (i) Construct Venn Diagram to present the above situation.
 (ii) What percentage of newspaper readers subscribe at least one of the three newspapers?
 (iii) What percentage of newspaper readers subscribe none of the three newspapers?

SECTION-B

- Q. No.6. (a)

Critically evaluate the role of probability theory in statistical inference.

(8)
- (b)

One of the larger components of the cost of living is the amount spent on housing. Housing cost includes rent (for tenants), mortgage payments (for home owners), heating, electricity and water. An Economist undertook a five years study to determine how housing costs have changed. Five years ago, he took a random sample of 200 households and recorded the percentage of total income spent on housing. This year he took a sample of 100 households.

(12) (20)

The statistics from the two samples are:

$$\bar{x}_1 = 32.42; s_1 = 6.08; \bar{x}_2 = 33.72; s_2 = 6.75.$$

Test the hypothesis at 5% level, to determine whether the Economist can infer that housing cost as a percent of total income has increased over the last five years.

- Q. No.7. (a)

Suppose you want a confidence interval for a Normal Mean with small sample size. Explain how would you construct it.

(10)
- (b)

The population and death of two towns, according to age groups are given below:

(10) (20)

Age group	Town A		Town B	
	Population	No. of deaths	Population	No. of deaths
Under 10	100	40	1000	300
11---25	1500	150	500	40
26---50	850	108	400	24
Above 50	50	37	100	6

- (i) Calculate crude death rate for town A
- (ii) Calculate SDR of town A taking the population of town B as a standard population.

- Q. No.8. (a)

Critique the use of p-values in statistical hypothesis testing. Discuss alternatives for interpreting statistical significance.

(10)
- (b)

Suppose in a field experiment there are 4 treatments (T1, T2, T3, and T4) and 4 blocks (B1, B2, B3, and B4). Consider the field is portioned into 4 plots (blocks), each of these blocks is homogeneous. Assign each treatment randomly to each plot (block) by fulfilling the requirements of randomized completely block design conditions (requirements), that you have to complete the layout of the experiment according to the RCBD. What will be the ANOVA table (a general form of RCBD ANOVA table) for this experiment?

(10) (20)

Blacks	T 1	T 2	T 3	T 4
Block I	2	5	4	1
Block II	2	3	3	1
Block III	4	6	6	2
Block IV	1	4	2	3

Perform the analysis of variance to test at the 0.05 level of significance.
