Entrance Examination : M.Sc. Statistics, 2015

Hall Ticket Number | | | | | | | | |

Time : 2 hours
Max. Marks. 100

Part A : 25 marks
Part B : 75 marks

Instructions

1. Write your Hall Ticket Number on the OMR Answer Sheet given to you. Also write the Hall Ticket Number in the space provided above.
2. Answers are to be marked on the OMR answer sheet.
3. Please read the instructions carefully before marking your answers on the OMR answer sheet.
4. Hand over the OMR answer sheet after the examination.
5. There are plain sheets in the booklet for rough work, no additional sheets will be provided.
6. Calculators are not allowed.
7. There are a total of 50 questions in Part A and Part B together in 11 pages (2-12)
8. **Each question in Part - A has only one correct option and there is negative marking of 0.33**
9. **There is no negative marking in Part - B. Some questions have more than one correct option. All the correct options have to be marked in the OMR answer sheet, otherwise zero marks will be credited.**
10. The appropriate answer(s) should be coloured with either a blue or a black ball point or a sketch pen. DO NOT USE A PENCIL.
11. **THE MAXIMUM MARKS FOR THIS EXAMINATION IS 100 AND THERE WILL BE NO INTERVIEW.**
12. Given below are the meanings of some symbols that may have appeared in the question paper:
    $\mathbb{R}$-The set of all real numbers, $\mathbb{N}$-The set of all natural numbers,
    $\mathbb{Z}$-The set of all integers, $E(X)$-Expected value of the random variable X,
    $V(X)$-Variance o the random variable X, $Cov.(X,Y)$-Covariance of the random variables X and Y.

H-07

Part - A

- **Find the correct answer and mark it on the OMR sheet. Each correct answer gets 1 (one) mark and wrong answer gets -0.33 marks.**

1. Which of the following is a random experiment?

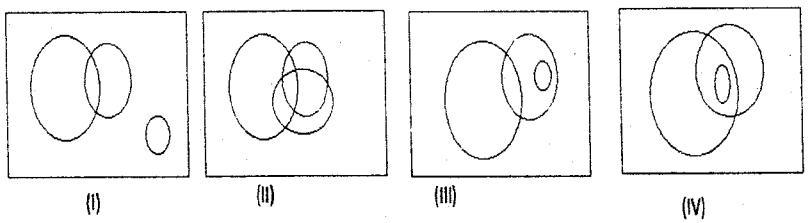
   (A) Place blue litmus paper in acid and observe the colour of the paper.

   (B) From an urn containing some white and some black balls, draw two balls and observe their colours.

   (C) From an urn containing white balls, draw one and observe the colour.

   (D) Find the length of the shadow of a pole fixed at one point at 12.00 noon.

2. Of the 100 students in a class 55 know Telugu, 35 know Hindi, all 16 of those who know Bengali also know Hindi but do not know Telugu, which of the following diagrams represents this data most precisely?



   (A) (I). **(B)** (II). **(C)** (III). **(D)** (IV).

3. 12 red(all alike) and 7 blue(all alike) balls are to be arranged in a row, if every arrangement is equally likely, the probability of getting an arrangement in which the first, middle and last balls are red

   (A) is more than $1/2$ **(B)** is more than $1/3$ but less than or equal to $1/2$

   (C) is not more than $1/4$. **(D)** is more than $1/4$ but less than $1/3$.

4. If X and $-X$ are independent,

   (A) it means that $V(X) = 0$.

   (B) it means that $E(X) = 0$.

   (C) it means that $E(X^2) = 0$.

   (D) none of the above can be said.

5. Which of the following best describes the Geometric random variable?

   (A) the number of attempts it takes Ashok to solve a puzzle.

   (B) the number of throws of a fair die till either 4 or 6 shows up.

   (C) the number of throws of a fair die till 4 and 6 appear consecutively.

   (D) the number of throws of a fair die till the sum of numbers that have shown up is an even number.

6. Let $X \sim B(20, 1/2)$, and $Y = 20 - X$, then

   (A) Y and X are identically distributed.

   (B) $E(Y) = -10$.

   (C) $V(Y) = 20 - V(X)$.

   (D) $Cov(X, Y) = 0$.

7. $X \sim P(7)$ and let $p_j = Pr(X = j),\ j = 0, 1, \ldots$, then

   (A) $p_4 < p_5 < p_7 < p_8$.

   (B) $p_8 < p_9 < p_{10} < p_7$.

   (C) $p_{15} < p_{11} < p_{10} < p_7$.

   (D) $p_3 < p_5 < p_7 < p_9$.

8. $Y = 2X$, identify the correct statement regarding the distribution of $2Y$.

   (A) $X \sim B(10, 1/2)$, then $Y \sim B(20, 1/2)$.

   (B) $X \sim P(5)$, then $Y \sim P(10)$.

   (C) $X \sim exp(10)$, then $Y \sim exp(20)$.

   (D) $X \sim exp(10)$, then $Y \sim exp(5)$.

9. The probability of getting at least 10 successes in $10,00,000$ iid Bernoulli trials with probability of success being $1/2,00,00,000$ is best obtained

   (A) by the Poisson approximation.

   (B) by the Normal approximation.

   (C) as 0.

   (D) as $1/2$.

10. The haemoglobin(Hb.) levels of Indian women in the $15-45$ years age group are normally distributed with mean $9.5mmg.$ and standard deviation $1.2mmg.$, $a\%$ of these ladies have at least $11.5mmg$(Hb.), what can you say about the percentage of these ladies with less than $7.25mmg$(Hb.)?

(A) it is $100-a$ per cent.

(B) it is less than a per cent.

(C) it is greater than a per cent.

(D) none of the above.

11. X_1 and X_2 are independent and identically distributed (iid) as $U((-1,1))$ (Uniform over $(-1,1)$ random variables, so $Cov.(X_1+X_2\ ,\ X_1-X_2)$ is

(A)$1/12$. **(B)**$1/6$. **(C)**$-1/12$. **(D)**0.

12. $X_1 \sim U((-1,1])$; $X_2 \sim N(0,49)$; $X_3 \sim U((-3,8])$; $X_4 \sim N(-2,16)$, which of these random variables are symmetrically distributed about 0?

(A) Only X_1.

(B) All of them.

(C) Only X_2 and X_3.

(D) Only X_1 and X_2.

13. $Z \sim N(0,1)$, then $E((2Z+1)(Z-3)))$ is

(A) -1. **(B)** 1. **(C)** -2. **(D)** -6.

14. The most appropriate diagram to represent data relating to the monthly expenditure of a family on different items is

(A) Histogram.

(B) Pie-diagram.

(C) Frequency polygon.

(D) Line graph.

15. The median of 21 distinct median was increased by 1, which of the following measures is not effected by this change?

(A) Mean. **(B)**Mean deviation about median. **(C)**Mid-range. **(D)**Standard deviation.

16. The average marks of the top $\frac{1}{3}^{rd}$ of the students of a class is 75% and the average marks of the rest of the class is 45%, the average marks of the whole class

(A) is 50%. (B) is 55%.

(C) is 60%. (D) can not be determined from the data given.

17. if the frequency distribution of a data set containing numbers-say incomes of people is negatively skewed, one can say that

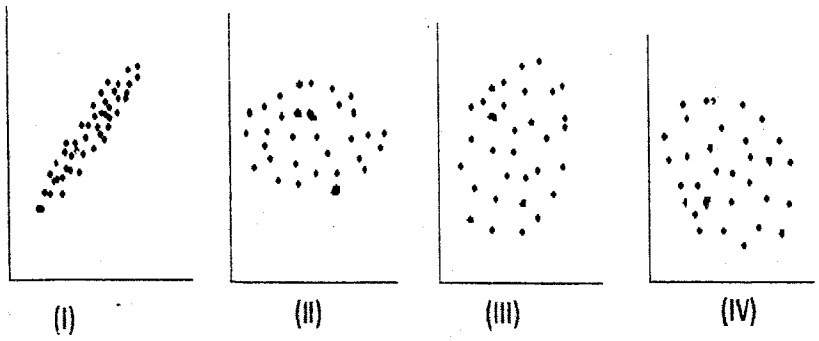
(A) The median is very likely to be closer to the third quartile than to the first.

(B) The median is very likely to be the exactly in the middle of the first and third quartiles.

(C) The median is very likely to be closer to the first quartile than to the third.

(D) The mean is very likely to be the exactly in the middle of the first and third quartiles.

18. Based on data on two variables for n units the correlation coefficient was 1/2, which of the following scatter plots best suits this fact?



(A)(I). (B)(II). (C)(III). (D)(IV).

19. $X_1, \ldots, X_n$ is a random sample from the $N(\mu, 16)$ population, the estimator $T = \frac{2\sum_{i=1}^{n} X_i}{n(n+1)}$ for μ is

(A) is unbiased and sufficient.

(B) neither unbiased nor sufficient.

(C) unbiased and efficient.

(D) unbiased but not sufficient.

20. In a hypothesis testing problem of a null hypothesis H_0 versus the alternate hypothesis H_1, type-II error is

(A) rejecting H_0 when it is true.

(B) rejecting H_1 when it is true.

(C) accepting H_0 when H_1 is true.

(D) rejecting both H_0 and H_1.

21. Consider the 2 hypotheses

- H_1 : The mean of a normal random variable is 0.
- H_2 : The number of red balls in a bag containing red and blue balls is 20

which of the following statements is correct?

(A) neither H_1 nor H_2 is a simple hypothesis.

(B) both are simple hypotheses.

(C) H_1 is a simple hypothesis, but H_2 is not.

(D) H_2 is a simple hypothesis, but H_1 is not.

22. Under a hypothesis H_a, the probability of obtaining a particular data set is 0.7, while the probability of this data set being observed if an other hypothesis were true is 0.01, then we should

(A) accept H_a in favour of H_b.

(B) accept H_b in favour of H_a.

(C) accept both H_a and H_b.

(D) reject both H_a and H_b.

23. A is a 4×4 real non-singular real matrix, the first,second, third and fourth rows of 5×5 matrix B are the first, second, third and fourth rows of A respectively followed by 0 and the fifth row of B is $(0, 0, 0, 0, 1)$, then

(A) the rank of B is equal to the rank of A.

(B) the rank of B is 5.

(C) $det.B \neq det.A$

(D) trace of B is equal to trace of A.

24. The value of $\lim_{n\to\infty}(\frac{1+n+n^2}{n^2})^n$

(A) is 1.

(B) lies in the interval $(1, 2]$.

(C) lies in the interval $(2, 4]$.

(D) is greater than 4.

25. The negation of the statement 'Ashok did not submit at least one of the assignments' is

(A) Ashok submitted all the assignments.

(B) Ashok submitted exactly one assignment.

(C) Ashok submitted at least one of the assignments.

(D) Ashok did not submit any assignment.

Part-B

- **Questions (26)-(37) have more than one correct option. For the answer to be right all the correct options have to be marked on the OMR sheet. No credit will be given for partially correct answers.**
- **Questions (38)-(50) have only one correct option.**
- **Find the correct answers and mark them on the OMR sheet. Correct answers (marked in OMR sheet) to a question get 3 marks and zero otherwise.**

26. The product of 4 positive numbers x_1, x_2, x_3, x_4 is 4096, then

(A) $x_1 + x_2 + x_3 + x_4 \leq 31$.

(B) $x_1 + x_2 + x_3 + x_4$ could be 40.

(C) $\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \frac{1}{x_4} > 1/4$.

(D) $x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 \geq 2048$.

27. The probability of at least one of the three events A, B, C occurring is 0.7 and the probability of at least one of them not occurring is 0.8, then the probability that

(A) at most one of the three events occurs is 0.8.

(B) either exactly one of them or exactly two them occurring is 0.5.

(C) exactly two them occurring is 0.3.

(D) exactly one of them occurring can not be obtained from the data given.

28. The probabilities of getting heads upon tossing the coins A and B respectively are 1/3 and 2/3. Let X and Y be the numbers of heads in 4 tosses of each of the coins A and B respectively, then

(A) X is always less than Y. **(B)** $X = 4 - Y$.

(C) $Pr(X = 2) = Pr(Y = 2)$. **(D)** $Pr(X = 1) = Pr(Y = 3)$.

29. A, B and C are three events with non-zero probabilities further $P(A|B) = P(A) = P(A|C)$,

(A) A, B, and C are mutually independent events.

(B) the events A and C are independent.

(C) B and C may not be independent.

(D) A, B and C may not be jointly independent.

30. Identify the correct statements on sums of independent random variables

(A) If X_1 and X_2 are independent Geometric random variables, $X_1 + X_2$ is also a Geometric random variable.

(B) If X_1 and X_2 are independent Poisson random variables, $X_1 + X_2$ is also a Poisson random variable.

(C) If X_1 and X_2 are independent Negative Binomial random variables, $X_1 + X_2$ is also a Negative Binomial random variable.

(D) If X_1 and X_2 are independent Exponential random variables with the same means, $X_1 + X_2$ is also an Exponential random variable.

31. sums and products of every pair of n numbers are given, however it is not known as to which pair's sum and product they are, from this information one can determine

(A) the mean of the n numbers.

(B) the standard deviation of the n numbers.

(C) median of the n numbers.

(D) range of the n numbers.

32. In the Systematic Sampling Scheme of 10 units from 1000 units that are numbered $1, \ldots, 1000$

(A) each of the 1000 has the same probability of being included in the sample.

(B) each of the $\binom{1000}{10}$ collections of 10 units has the same probability of being selected.

(C) there are exactly 100 different samples that can be selected.

(D) units numbers 1 and 2 can not be in the selected sample.

33. The probability density function of a random variable X is

$$f_X(x) = \frac{1}{\pi(1+(x-1)^2)} \qquad -\infty < x < \infty$$

(A) The expected value of X does not exist.

(B) The expected value of X is 1.

(C) The median of X is 1.

(D) X is symmetrically distributed about 1.

34. The variances of T_1 and T_2 are two unbiased estimators of a parameter θ are v_1 and v_2 respectively, further $v_1 < v_2$ and $Cov.(T_1, T_2) = v_1$, then

(A) $\frac{T_1+T_2}{2}$ is more efficient than both T_1 and T_2.

(B) $\frac{T_1+T_2}{2}$ is more efficient than T_2but less efficient than T_1.

(C) $\frac{4T_1+T_2}{5}$ is more efficient than T_1.

(D) $\frac{4T_1+T_2}{5}$ is more efficient than T_2.

35. Consider the function $f : \mathbb{R} \longrightarrow \mathbb{Z}$ as $f(x) = \lfloor x \rfloor$ where $\lfloor x \rfloor$ is the largest integer that is less than or equal to x, then

(A) $\lfloor |x| \rfloor \leq |\lfloor x \rfloor|$, $\forall x \in \mathbb{R}$.

(B) f is non-decreasing and continuous in $\mathbb{R}$.

(C) $-2, 0, 2$ are some of the points at which f is not continuous.

(D) the function g defined as $g(x) = x - \lfloor x \rfloor$ is increasing in $[a, a+1)$, $\forall a \in \mathbb{Z}$.

36. $X_1, \ldots, X_n$ is a random sample from the $N(\mu, \mu^2)$ population, let $\overline{X}_n = \frac{X_1+\ldots+X_n}{n}$ and $T_n = \frac{X_1{}^2+\ldots+X_n{}^2}{n}$ and $s_n^2 = \frac{(X_1-\overline{X}_n)^2+\ldots+(X_n-\overline{X}_n)^2}{n-1}$, so

(A) both s_n^2 and $T_n/2$ are unbiased estimators μ^2.

(B) $\lim_{n\to\infty} E(\overline{X}_n^2) = \mu^2$.

(C) $n\overline{X}_n$ is not a sufficient statistic for (μ, μ^2).

(D) $(n-1)s_n^2$ and nT_n are jointly sufficient for (μ, μ^2)

37. Let $\underset{\sim}{\mathbf{r}}_1^T, \underset{\sim}{\mathbf{r}}_2^T, \ldots, \underset{\sim}{\mathbf{r}}_n^T$ be the n rows of a $n \times n$ real matrix A, suppose $\underset{\sim}{\mathbf{r}}_n^T$ is a linear combination $\underset{\sim}{\mathbf{r}}_1^T, \ldots, \underset{\sim}{\mathbf{r}}_{n-1}^T$, then

(A) Rank of A may be n.

(B) Rank of A is not more than $n-1$.

(C) Rank of A is equal to $n-1$.

(D) The determinant of A is equal to 0.

38. Two fair dice-D_1 and D_2 are tossed, let X_1 and X_2 denote the numbers that D_1 and D_2 showed up respectively. Let $U = |X_1 - X_2|$ and $V = X_1 + X_2$, the conditional probability $Pr(V = 5|U = 3)$ is

(A) 3/4. (B) 2/3. (C) 1/2. (D) 1/3.

39. 10 adults and 3 children are to be seated in a row if every arrangement is equally likely, the probability that exactly two of the children are next to each other is

(A)1/13. (B)5/26. (C)3/26. (D)5/13.

40. The probability generating function of a discrete random variable is $P(s) = k(2+3s+5s^2)$, its variance is

(A) is 0.21. **(B)** 0.41. **(C)** 0.61. **(D)** 0.81.

41. The mean and variance of the first 9 numbers in a data set are 38 and $21\frac{1}{3}$ respectively, further the mean of the first 10 numbers in the data set is 39, the variance of the first 10 numbers

(A) can not be determined from the given information.

(B) is $\frac{282}{10}$.

(C) is $\frac{231}{10}$.

(D) is 30.

42. X and Y are independently and identically distributed as $N(0, 1/2)$, then the distribution of $Z = (X-Y)^2 + (X+Y)^2$ is that of

(A) $\frac{1}{2}\chi^2_2$. **(B)** χ^2_2. **(C)** $\frac{1}{4}\chi^2_2$. **(D)** $2\chi^2_2$.

43. The value of $E(X|Y=3)$ if the joint probability distribution function of the random variables X and Y is

$$Pr(X=i,\ Y=j) = \tfrac{1}{6},\ j>i,\ \{i,j\} \subset \{1,2,3,4\}$$

(A) is 0.5. **(B)** is 1. **(C)** is 1.5. **(D)** is 2.

44. The median of the random variable X whose *pdf* for the suitable positive real number c is

$$f(x) = \begin{cases} cx^2(1-x) & 0<x<1 \\ 0 & o.w \end{cases}$$

(A) is less than 1/4. **(B)**is at least 1/4 but less than 1/2.

(C)at least 1/2 but less than 2/3. **(D)** at least 2/3.

45. The variance of a random variable is less than 4, so the probability that the random variable will take values within 7 units from its expected value

(A)is less than 0.5. **(B)**is more than 0.9.

(C)is in the interval $[0.5, 0.9]$. **(D)** nothing can be said from the information given.

46. $-1,\ -1.5,\ 2,\ 0.75,\ 1.7,\ 1.2$ is a random sample from the Uniform distribution over $(-\theta,\ \theta)$, $\theta > 0$, the maximum likelihood estimate of θ

(A) is 2. (B) is 1.5.

(C) is 1. (D) can not be obtained from the given data.

47. Of the 20 balls in a bag M are red and the rest are blue. 5 draws of 3 balls were made-that is 3 balls were drawn without replacement which are the put back, this is done 5 times, the number of red balls that showed up in each of these draws were $2, 1, 1, 0$ and 2, an unbiased estimate of M is

(A)6. (B)8. (C)10. (D)12.

48. $e^{-1},\ e^{-0.5},\ e^{-0.35},\ e^{-0.15}$ are 4 independent observations of a random variable X whose probability density function is $f_X(x) = \begin{cases} (m+1)x^m & 0 < x < 1 \\ 0 & o.w \end{cases}$, $m > 0$, the maximum likelihood estimate of $m > 0$

(A) is 0.5. (B) is 1.

(C) is0.2. (D) can not be determined from the data given.

49. The value of $\int_0^\infty \frac{e^{-x^2}}{\sqrt{\pi}} dx$ is

(A)$\frac{1}{2}$. (B)$\frac{1}{\sqrt{2}}$.

(C)$\sqrt{2}$. (D) $\frac{1}{2\sqrt{2}}$.

50. A bag contains N balls numbered $1, 2, \ldots, N$, to test the hypothesis $H_0 : N = 10$ vs. $H_1 : N = 20$ draw two balls from the bag without replacement, let X denote the larger of the two numbers on the drawn balls and reject H_0 if $X > 9$. The power of this test is close to

(A)0.22. (B)0.56. (C)0.623. (D)0.76.