Entrance Examination : M.Sc. Statistics - OR, 2017

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.

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 <u>more than</u> 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**

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, $E(X)$-Expected value of the random variable X,
   $V(X)$-Variance of the random variable X, $Cov.(X,Y)$-Covariance of the random variables X and Y, $\rho_{X,Y}$ denotes the correlation coefficient between X and Y, iid-independent and identically distributed, pdf-probability density function, $B(n,p)$ and $N(\mu,\sigma^2)$ denote respectively,the Binomial and the Normal distributions with the said parameters. $Rank(A)$ and $det(B)$ mean rank and determinant of the matrices A and B respectively.

13. This book contains **10 pages** including this page and excluding pages for the rough work. Please check that your paper has all the pages.

N-7

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. Bag1 contains 6 red, 5 blue and 4 green balls while Bag2 contains 6 green, 5 blue and 4 red balls, a ball is drawn from each of the bags, if every ball is equally likely to be drawn, the probability that a ball of the same colour will be drawn from both the bags is

   (A) a little more than 1/2.

   (B) equal to 1/4.

   (C) equal to 1/3.

   (D) a little less than 1/3.

2. If every arrangement of 10 balls numbered $1, 2, \ldots, 10$ in a row is equally likely, the probability that all the even numbers are before any odd number is

   (A) less than 0.004.   **(B)** between 0.005 and 0.007.

   (C) between 0.008 and 0.01.   **(D)** more than 0.011.

3. Arrange the numerals $1, 2, 3, 4, 5$ in a row to get a 5 digit number, the number of such arrangements that are divisible by 3 is

   (A) $\frac{5!}{4}$.   **(B)** $\frac{5!}{3}$.   **(C)** $\frac{5!}{2}$.   **(D)** 5!.

4. From a set of 100 distinct objects what is the number of different non-empty subsets containing an even number of objects?

   (A) 2^{99}.   **(B)** $2^{99} - 1$.   **(C)** $2^{50} - 1$.   **(D)** $2^{49} - 1$.

5. The probabilities of two events A and B, $P(A)$ and $P(B)$ respectively are positive, further $P(A|B) > P(A)$, then

   (A) $P(B|A) > P(B)$.   **(B)** $P(B|A) < P(B)$.

   (C) $P(B|A) = P(B)$.   **(D)** Nothing can be said definitely about $P(B|A)$.

6. A and B are mutually exclusive events and $0 < P(A) < 1$, $0 < P(B) < 1$, so,

   (A) $P(A) \le P(B^c)$.   **(B)** $P(A \cap B) = P(A)P(B)$.

   (C) $P(A^c) < P(B)$.   **(D)** $P(B|A) = P(B)$.

7. A_1, A_2, and A_3 are independent events each of which occur with the same probability p, then, the probability of at most one of A_1, A_2 and A_3 occurring is

   (A) $3p(1-p)^2$.   **(B)** $3p(1-p)^2 + (1-p)^3$.   **(C)** $p^3 + (1-p)^3$.   **(D)** $1 - (1-p)^3$.

N-7

8. X is a non-negative discrete random variable for which $P(X = j) > P(X = j+1)$, $\forall j = 0, 1, 2, \ldots,$ which of the following distributions fits this fact?

   (A) Geometric distribution.

   (B) Binomial distribution.

   (C) Poisson distribution.

   (D) Hypergeometric distribution.

9. The expected value of a random variable is -10 and its variance is 100, the value of the second moment is

   (A) 10.   **(B)** 100.   **(C)** 200.   **(D)** 1000.

10. $V \sim B(10, 1/2)$, the probability that the quadratic equation $x^2 + 6x + V = 0$ will have complex roots is

    (A) $\frac{1}{2^{10}}$.   **(B)** $1 - \frac{1}{2^{10}}$.   **(C)** $\frac{1}{2}$.   **(D)** $1/10$.

11. The probability distribution function of a Poisson random variable with a very large mean is to be approximated as a

    (A) Exponential random variable.

    (B) Negative Binomial random variable.

    (C) Normal random variable.

    (D) Hypergeometric random variable.

12. The 4th head did not occur till the 15th toss of a coin, so,

    (A) A Binomial random variable will be observed to be 15.

    (B) A Negative Binomial random variable will take a value that is at least 15.

    (C) A Binomial random variable will be observed to be more than 15.

    (D) A Negative Binomial random variable will take a value that is more than 15.

13. In garment workshop, shirts are stitched for export, about 10% of the shirts stitched here do not meet specifications and hence are called defective, from a lot of 1000 shirts made in a day, the number of defective shirts in a sample of 100 is a

    (A) Negative Binomial random variable.

    (B) Hypergeometric random variable.

    (C) Binomial random variable.

    (D) Poisson random variable.

N-7

14. The most appropriate diagram to represent data on grades achieved by students in a public exam is

    (**A**) Bar charts.   (**B**) Histogram.

    (**C**) Stem and leaf plot.   (**D**) Ogive.

15. Which of the following is not a measure of dispersion

    (**A**) Range.   (**B**) Mean deviation about median.

    (**C**) Mode.   (**D**) Mean deviation about mean.

16. The correlation coefficient $\rho_{X,Y}$ between the random variables X and Y is 0.8, then the correlation coefficient between X and $U = 20 - 3.2Y$ is

    (**A**) 0.8.   (**B**) −0.8.   (**C**) 1.   (**D**) 0.

17. The correlation coefficient calculated based on n observations on the random variables X and Y was − 0.3, which of the scatter plots given below reveals this correlation?
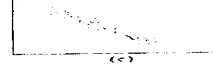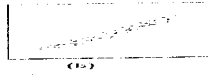
18. The average of n positive numbers is 100 and their product is 100000000, then,

    (**A**) $n = 2$.   (**B**) $n = 3$.

    (**C**) $n \geq 4$.   (**D**) we can't say anything about n based on data given.

19. Consider the function $f(x) = \begin{cases} \frac{3}{2}x(x-1) & 0 < x \leq 2 \\ 0 & elsewhere \end{cases}$

    (**A**) It is a probability density function

    (**B**) It is not a probability density function because $\int_{-\infty}^{\infty} f(x)dx \neq 1$.

    (**C**) It is not a probability density function because $f(x) < 0$ for some values of x.

    (**D**) It is not a probability density function because f is not increasing in x.

20. The heights of adult males in a certain population are normally distributed, the heights of half of them are more than 165cm., while the heights of 5% of them are more than 183cm., the percentage of adult males whose heights are less than 147cm.

    (**A**) is less than 2%.   (**B**) is 5%.

    (**C**) is more than 2%.   (**D**) we can't say based on data given.

N-7

21. X_1 and X_2 are independent standard normal random variables $N(0,1)$, $X_1^2 + X_2^2$ follows:

   (A) Chi square distribution with 1 degree of freedom.

   (B) Exponential distribution with mean 1

   (C) Exponential distribution with mean 2

   (D) None of the above.

22. A sample survey is to be done to estimate the average milk consumption per family in a city, to draw the most representative sample it was decided to stratify the families and then select samples from each stratum, on the basis of which of the following criteria should the families be stratified.

   (A) The last digit of the cell phone number of the head of the family.

   (B) The height of the head of the family.

   (C) The month of birth of the head of the family.

   (D) the total income of the family.

23. In a hypothesis testing problem of H_0 against H_1 based on a sample, type - 2 error occurs when

   (A) the sample is such that it falls in the complement of the critical region when H_1 is true.

   (B) the sample is such that it falls in the critical region when the H_0 is true.

   (C) the sample is such that it falls in the critical region when H_0 is false.

   (D) the sample is such that it falls in the critical region when neither H_0 nor H_1 is true.

24. The function $F(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$ is

   (A) continuous at all $x \in \mathbb{R}$.

   (B) continuous everywhere but not differentiable at some points.

   (C) decreasing in x.

   (D) not continuous at one point.

25. Which of the following is equivalent to the statement 'Ashok did not solve all the problems'

   (A) Ashok did not solve any problem.

   (B) Ashok did not solve at least one problem.

   (C) Ashok solved at least one problem.

   (D) Ashok solved at most one problem.

Part - B

- Questions (26)-(37) have more than one correct option.
- For the answer to be right <u>all the correct options</u> 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. Which of the following are random experiments?

   (A) Put paper in fire and see what happens to it.

   (B) Ask a child to place 10 balls, 1 of which is red and the rest blue in a line and observe the position of the red ball.

   (C) Place 10 distinct objects in three distinct boxes and observe which object is in which box.

   (D) From a large basket of mangoes take out 5 and report their total weight.

27. $X_1 \sim U((-\sqrt{3}, \sqrt{3}])$ and $X_2 \sim N(0, 1)$, so,

   (A) $P(X_1 > 0) = P(X_2 \leq 0)$.

   (B) $P(-1.5 < X_1 \leq -1) = P(-3.5 < X_2 \leq -3)$.

   (C) $P(X_1 > 1) > P(X_2 > 1)$.

   (D) $E(X_1) = E(X_2)$.

28. For two random variables X and Y, $V(X + Y)$

   (A) is never less than either $V(X)$ or $V(Y)$.

   (B) is never less than either $V(X)$ or $V(Y)$ if X and Y are uncorrelated or positively correlated.

   (C) is always less than one of $V(X)$ and $V(Y)$.

   (D) can be less than both $V(X)$ and $V(Y)$ only if X and Y are negatively correlated.

29. In a certain country, 70% of the households have incomes less than the average income, those households among the highest 10% earners are considered upper class

   (A) the median income is less than the average income.

   (B) the median income is more than the average income.

   (C) About a third of the households with more than average incomes are in the upper class.

   (D) Less than a quarter of the households with more than the median income are in the upper class.

N-7

30. The probability distribution of a random variable X is $P(X = j) = \frac{c}{j(j+2)}$, $j = 1, 2, \ldots$,

   (A) c is equal to 2/3.

   (B) c is equal to 4/3.

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

   (D) The variance of X does not exist.

31. Identify the common properties of the random variables with the following pdfs

$$f_1(x) = \begin{cases} \frac{1}{2} & -1 < x < 1 \\ 0 & e.w \end{cases} \qquad f_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty \qquad f_3(x) = \frac{1}{2} e^{-|x|}, \quad -\infty < x < \infty$$

   (A) All of them have the same mean.

   (B) The third moment of all of them is 0

   (C) X and $-X$ of all of them are identically distributed.

   (D) All of the have the same variance.

32. regarding a simple random sample of size n without replacement from a population of size N, identify the correct statements

   (A) Sample raw moments are unbiased estimators of the corresponding raw moments of the population.

   (B) Every collection of n population units is equally likely to be the selected sample.

   (C) The second central moment of the sample is not an unbiased estimator of the second central moment of the population.

   (D) Every unit of the population is equally likely to be in the selected sample.

33. The probability with which a coin shows heads upon tossing is p, the random variable X_1 takes the values 1 and 0 if the outcome of the first toss is heads or tails respectively, another random variable X_2 is defined in the same way based on the second toss.

   (A) $\frac{X_1 + X_2}{2}$ is an unbiased estimator of p.

   (B) $2X_1 - X_2$ is also an unbiased estimator of p, but not the most efficient.

   (C) $X_1 - X_2$ is a sufficient statistic of p.

   (D) $X_1 + X_2$ is a sufficient estimator, but $X_1 - X_2$ is not a sufficient statistic for p.

34. A, B and C are three events and if $P(A) = P(B) = P(C) = 2/3$, $P(A \cap B) = P(A \cap C) = P(B \cap C) = 1/2$, then $P(A \cap B \cap C)$

   (A) has to be zero.   (B) can be 2/5.

   (C) can not be 1/6.   (D) can be 1/4.

N-7

35. If X and Y are independent real valued random variables, then

   (A) $E(XY) = E(X)E(Y)$.

   (B) $V(X + Y) = V(X) + V(Y)$.

   (C) $V(X|Y = y) = E(X)$ for every $y \in \mathbb{R}$.

   (D) $2X$ and $-3Y$ are also independent random variables.

36. Consider the data given below on the students of a university

   $Number$ of $female$ $students:$ 2000
   $Number$ of $male$ $students:$ 4000
   $Number$ of $female$ $students$ $residing$ in $hostels:$ 1600
   $Number$ of $male$ $students$ $residing$ in $hostels:$ 3000

   from this data one can see that

   (A) only one third of the students are female.

   (B) only one third of the hostel residents are female.

   (C) a larger proportion of female students stay in hostels than male students.

   (D) more than half of the students are hostel residents.

37. A is a 5×5 real matrix whose 5th row is the sum of the first and second rows, let A^T denote its transpose, then certainly

   (A) the rank of A is equal to 4.

   (B) the rank of A^T is at most 4.

   (C) the determinant of $A^T A$ is equal to 0.

   (D) the rank of $-A$ is less than 5.

38. X is a Poisson random variable with mean λ, if $E((2X + 1)(X - 1)) = 0$, then
   (A) $V(X) = 1/2$.   (B) $\lambda = 1$.
   (C) $\lambda = 2$.   (D) λ can not be uniquely determined.

39. The value of the integral $\int_0^\infty x^4 e^{-2x} dx$ is equal to
   (A) $15/4$.   (B) $5/2$.   (C) $3/4$.   (D) $1/2$.

40. Let p be the probability that a coin will show heads upon tossing, further let X denote the number of heads in n tosses of this coin, an unbiased estimator for p^2
   (A) is X^2.   (B) does not exist.
   (C) is $\frac{1}{n^2} X^2$.   (D) is $\frac{1}{n(n-1)} X(X - 1)$.

N-7

41. Let m_0, M_0 and S_0 denote the mean, median and standard deviation respectively of 15 distinct numbers, further suppose the difference between the median and the largest number smaller than it is 5 and the difference between the median and the smallest number greater than it is 4. Now add 2 to each of the 5 smallest numbers in this list and subtract 2 from each of the larger 5 numbers, denote by m_1, M_1 and S_1 the mean, median and the standard deviation respectively of this new set of numbers,

(A) $m_1 > m_0$; $M_1 = M_0$; $S_1 > S_0$

(B) $m_1 = m_0$; $M_1 = M_0$; $S_1 < S_0$

(C) $m_1 < m_0$; $M_1 < M_0$; $S_1 < S_0$

(D) $m_1 = m_0$; $M_1 = M_0$; $S_1 = S_0$

42. A statistic T_n to estimate a parameter θ based on a random sample of size n of a certain random variable is such that for 90% of the samples of size n of that random variable the value of T_n is more than $\theta + 10$ and for the remaining 10% of the samples of size n, the value of T_n is equal to θ. This statistic T_n is

(A) a good estimator for θ as it has low variance.

(B) a good estimator for θ as it seems to be unbiased

(C) is not a good estimator for θ as it is more likely to overestimate θ.

(D) is not a good estimator for θ as it is more likely to underestimate θ.

43. $X_1, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$, an unbiased estimator for $\sigma^2 + \mu^2$ is
(A) $\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.   **(B)** $n\bar{X}^2 - \frac{1}{n}\sum_{i=1}^{n}X_i^2$.   **(C)** $n\bar{X}^2$.   **(D)** $\frac{1}{n}\sum_{i=1}^{n}X_i^2$.

44. Let p be the probability of a coin showing up heads when tossed, the hypothesis $H_0 : p = p_0$ is to be rejected in favour of $H_1 : p > p_0$ if the number of heads-X that show up in 10 tosses of this coin is at least 8.

(A) the power of this test is equal to the size of this test for any $p_1 > p_0$.

(B) for some values of $p_1 > p_0$, the power of this test is more than its size.

(C) the power of this test is more than its size for every $p_1 > p_0$.

(D) the power of this test is less than its size for every $p_1 > p_0$.

45. The sum of squares and the products of every pair of n non-negative real numbers $x_1, \ldots, x_n$ are known, however $x_1, \ldots, x_n$ are not known, based on this information

   (A) both the mean and the standard deviation of these numbers can be determined.

   (B) neither the mean nor the standard deviation of these numbers can be determined.

   (C) the mean can not be determined but the standard deviation of these numbers can be determined.

   (D) the mean can be determined but the standard deviation of these numbers can not be determined.

46. There are 3 True or False questions in an exam, if a candidate knows the answer she/he answers it correctly, otherwise a guess is made and the probability of getting it right is $1/2$, an examiner assumes that every candidate knows no answer, 1 answer, 2 answers, 3 answers with equal probabilities, a candidate answered two of the three questions correctly, what is the probability that this candidate knew the answer to only one of them?

   (A) $1/11$.   (B) $2/11$.   (C) $3/11$.   (D) $4/11$.

47. There are 2 red and 2 blue balls in a bag, balls are to be removed one by one, the probability that the second ball to be drawn will be red is

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

48. X is a non-constant real valued random variable whose expected value is less than 0, and $\rho_{X,X^2} = 0$

   (A) X and $-X$ are independent.

   (B) $E(X^3) > 0$.

   (C) $E(X^3) > E(X)^3$.

   (D) $E(X^3) < 0$.

49. The time to complete a one year project by an organization is a random variable with probability density function $f(x) = \begin{cases} cx^2(1-x) & 0 < x < 1 \\ 0 & e.w \end{cases}$, the probability that a project will get completed within 9 months(3/4 of a year) is

   (A) less than $1/3$.   (B) very close to $1/2$.   (C) almost $3/4$.   (D) almost 1.

50. $5/3, 5/6, 5/12, 5/6, 15/12$ are 5 independent observations of a random variable X whose probability density function is $f_X(x) = \begin{cases} \lambda^2 x e^{-\lambda x} & x \geq 0 \\ 0 & o.w \end{cases}$, $\lambda > 0$, the maximum likelihood estimate of λ based on the given sample

   (A) 2.   (B) 2.5.   (C) 3.75.   (D) 5.

N-7