Entrance Examination : M.Sc. Statistics - OR, 2018

Hall Ticket Number | | | | | | | |

Time : 2 hours | Part A : 25 marks

Max. Marks. 100 | 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 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,$E(X)$-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, $B(n,p)$, $N(\mu,\sigma^2)$ and $U((a,b))$ denote respectively,the Binomial, the Normal and the Uniform 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.

9-7

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. Fill in the missing words to the quote:"Statistical methods may be described as methods for drawing conclusions about ... based on ... computed from the ...."

   (A) populations, parameters, samples.

   (B) statistics, parameters, samples.

   (C) parameters, statistics, populations.

   (D) populations, statistics, samples.

2. Regarding two mutually exclusive events A and B, which of the following can happen?

   (A) $P(A) = 0.4$, $P(B) = 0.7$.

   (B) Both $P(A)$ and $P(B)$ are more than $1/2$.

   (C) $P(A) \geq 1/2$ and $P(B) < 1/2$.

   (D) $P(A^c \cup B^c) = 0.8$.

3. the numerals $2, 4, 6, 8$ are to be arranged randomly in a row so as to get a four digit number with even digits, if every arrangement is equally likely, the probability that the number so obtained is a multiple of 6 is

   (A) 0. **(B)** 1/6. **(C)** 1/3. **(D)** 1.

4. The mean of numbers 28, x, 42, 78 and 104 is 62, what is the mean of 48, 62, 98, 124 and x?

   (A) 78. **(B)** 58. **(C)** 31. **(D)** 66.

5. The probability that Ashok will not be able to solve a single problem is $1/10$, the probability that he will be able to solve some problem is

   (A) 0.1. **(B)** 0.5. **(C)** 0.8. **(D)**0.9.

6. For which of the following random variables is the variance always greater than the mean?

   (A) Normal. **(B)** Binomial. **(C)** Negative Binomial. **(D)** Poisson.

7. A data set consisting of say incomes of people in a country is positively skewed, so,

   (A) the median is equal to the mean.

   (B) the median is less than the mean.

   (C) the median is more than the mean.

   (D) nothing can be said about whether the median is more or the mean.

8. There are 20 locks and and their 20 keys, but they are mixed up and the attendant does not know which key opens which lock, so he simply places each key with a lock, in so doing the probability that at least one of locks 1 and 5 will get its key is

   (A) less than $1/20$.

   (B) more than $1/20$ but not more than $1/10$.

   (C) more than $1/10$ but not more than $1/5$.

   (D) more than $1/5$ but less than $1/4$.

9. Three numbers are to be drawn without replacement from $\{1, 2, \ldots, 20\}$ and every outcome is equally likely. If one of the selected numbers is 8, the probability that the sum of the three selected numbers is a multiple of 3 is

   (A) $1/5$. **(B)** $1/4$. **(C)** $1/3$. **(D)** $1/2$.

10. In an aquaculture study, the following variables were measured on each fish:Gender, Initial Weight(grams), Body temperature (°c), Weight Gain (gms.).The scale of these four four variables (in order) are:

    (A) Nominal, Ratio, Ratio, Ratio.

    (B) Nominal, Ratio, Interval, Interval.

    (C) Ordinal, Ratio, Interval, Ratio.

    (D) Nominal, Ratio, Interval, Ratio.

11. 10% of the purchase bills in a mall on any day are for books and stationary.Draw a sample of 100 bills at the end of the day, the number of bills for books and stationary in this sample is a

    (A) Hyper geometric random variable.

    (B) Poisson random variable.

    (C) Negative Binomial random variable.

    (D) Binomial random variable.

12. In a data set of non-negative integers the least observation is 0 and the largest observation is 20, let $n_0, n_1, \ldots, n_{20}$ be the frequencies with which $0, 1, \ldots, 20$ appear in the data, further $n_0 < n_1 < \ldots < n_5$ and $n_5 > n_6 > \ldots > n_{20}$, this data set could be observations of

    (A) a Geometric random variable with $p < 1/2$.

    (B) a Binomial random variable with $p < 1/2$.

    (C) a negative binomial random variable with $p < 1/2$.

    (D) an exponential random variable with mean greater than 5.

13. All except one ball in a bag containing 100 balls are red and the lone other ball is blue. Keep drawing a ball and set aside till the lone blue ball gets drawn and then stop. The number of draws till the blue ball shows up is

(A) a Binomial random variable. **(B)** a Bernoulli random variable.

(C) a Geometric random variable. **(D)** not any of the three mentioned.

14. X is a random variable X follows Poisson distribution with parameter 1, that is $X \sim P(1)$, then $E(X(X-1)(X-2))$ is equal to

(A) 1. **(B)** -1. **(C)** 0. **(D)** 2.

15. $(5, 12), (4, 10), (4, 9), (6, 14), (3.5, 9)$ are observations on the two random variables X and Y on a sample of 5 individuals. The sample correlation $r_{X,Y}$ is equal to or very close to

(A) 1. **(B)** -1. **(C)** 0. **(D)** $1/2$.

16. If the correlation between body weight and annual income were high and positive, we could say that

(A) high incomes cause people to eat more food.

(B) on average high income people tend to spend a greater proportion of their incomes on food than low income people do.

(C) on average, high income people tend to be heavier than low income people.

(D) high incomes are a cause of weight gain in people.

17. For any pair of independent random variables X and Y which of the following is **NOT** correct

(A) $P(X \leq x | Y > 2x) = P(X \leq x | Y \leq 2x) \quad \forall x \in \mathbb{R}$.

(B) $E(Y|X > x) = E(Y) \quad \forall x \in \mathbb{R}$.

(C) $V(X_1 - X_2) < V(X_1) + V(X_2)$.

(D) The correlation coefficient between X and Y, $\rho_{X,Y} = 0$.

18. The Probability distribution of a random variable X is $P(X = 0) = 1/4$; $P(X = 1) = 1/2$; $P(X = 2) = 1/4$, let $Y = 2 - X$, then $Cov(X, Y)$ is

(A) 0. **(B)** 1. **(C)** $1/2$. **(D)** $-1/2$.

19. X is a uniformly distributed random variable in the interval $(a, b]$, that is $X \sim U((a, b))$, consider a positive real number c that is less than $b - a$, suppose $P(a < X \leq a + c) = \alpha$, then, $P(b - c < X \leq b)$ is

(A) α. **(B)** $1 - \alpha$. **(C)** 2α. **(D)** $1 - \frac{\alpha}{2}$.

20. X is a Geometric random variable with parameter $1/5$, which ordering is correct?

(A) $P(X=3) > P(X=7) > P(X=4) > P(X=9)$.

(B) $P(X=2) > P(X=4) > P(X=6) > P(X=9)$.

(C) $P(X=1) < P(X=2) < P(X=3) < P(X=5)$.

(D) $P(X=4) < P(X=5) < P(X=6) > P(X=7)$.

21. The probability of getting the random sample that you have got is 0.021 if the Null Hypothesis H_0 were true, then,

(A) you should accept H_0 at 3% level of significance.

(B) you should reject H_0 at 1% level of significance.

(C) you should reject H_0 at 5% level of significance, but, accept it at 1% level of significance.

(D) you should reject H_0 at 1% level of significance, but, accept it at 5% level of significance.

22. The distribution of heights of students in a large class is roughly bell shaped, the average height is 68 inches and approximately 95% of the heights are between 62 and 74 inches. So, the standard deviation of the distribution of heights is approximately equal to

(A) 5. (B) 3. (C) 6. (D) 1.

23. A new headache remedy was given to a group of 25 subjects who had headaches. Four hours after taking the new remedy, 20 of the subjects reported that their headaches had disappeared. From this information one can conclude

(A) that the remedy is effective for the treatment of headaches.

(B) nothing, because the sample size is too small.

(C) nothing, because there is no control group for comparison.

(D) that the remedy is not effective for the treatment of headaches.

24. How many keystrokes are needed to type all the numbers from 1 to 1000?

(A) 3001. (B) 2893. (C) 2704. (D) 2890.

25. The rank of the matrix $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & 2 \end{pmatrix}$ is

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

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. X_1 and X_2 are independent Poisson random variables with parameters 1 and 2 respectively. Then,

(A) $E(X_2) > E(X_1)$. **(B)** $P(X_1 = X_2) < e^{-2}$.

(C) $V(X_2) > V(X_1)$. **(D)** $P(X_1 = X_2) > e^{-1}$.

27. X is a $X \sim U((-1,1))$, define another random variable $Y = |X|$, then
(A) $E(Y) = 0$. **(B)** $Cov.(X,Y) = 0$. **(C)** $E(Y) = 1$. **(D)** $V(Y) < V(X)$.

28. Which of the following is/are necessary conditions for the Binomial distribution?

(A) All trials are identical.

(B) All trials must be independent.

(C) Each trial must be classified as a success or a failure.

(D) The probability of success in each trial must be 1/2.

29. The *pdf* of a random variable X is $f_X(x) = \begin{cases} \frac{1}{4}|x| & -a < x < a \\ 0 & o.w \end{cases}$, so

(A) $a = 2$. **(B)** $E(X) = 0$. **(C)** $V(X) = 2$. **(D)** $P(X < -1) = P(X > 1)$.

30. $X_1 \sim N((0,1)$ and X_2 is another random variable whose *pdf* is $f(x) = \frac{1}{2}e^{-|x|}$ $-\infty < x < \infty$, identify the correct statement about the probability distributions of X_1 and X_2.

(A) $E(X_1) = E(X_2)$. **(B)** Medians of X_1 and X_2 are the same.

(C) $V(X_1) = V(X_2)$. **(D)** $P(-3 < X_1 \leq 3) > P(-3 < X_2 \leq 3)$.

31. In a Hypothesis testing problem H_0 - the null hypothesis against the alternate hypothesis H_1 regarding a parameter μ,

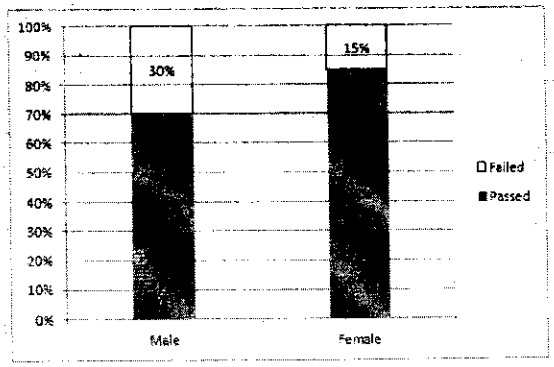
    (A) the test statistic and the test criterion are determined by both H_0 and H_1.

    (B) if μ is the mean of a normal random variable, the problem $H_0 : \mu = 0$ is a simple hypothesis and $H_1 : \mu > 0$ is a composite hypothesis.

    (C) based on the same data H_0 may be rejected against the alternate H_1 at α level of significance, but H_0 may be accepted against some other alternate hypothesis H_2 at the same level of significance.

    (D) if H_0 is rejected at α level of significance based on data, then based on the same data one will have to accept H_1 against H_0 at the same level of significance in the hypothesis test problem with the null and alternate hypotheses interchanged.

32. There are 3500 male and 2500 female students in a university, the bar diagram below displays information on pass and failed percentages.



From this chart one can see that

(A) More female students passed.

(B) A larger proportion among the failures is of males.

(C) A larger proportion among those who passed is of males.

(D) More than 70% of the students have passed.

33. The average of 10 positive numbers is 20, so,

(A) the sum of their squares could be 3500.

(B) their product could be 2048×10^{10}.

(C) the sum of their squares could be 5000.

(D) the sum of their square roots could be 40.

34. For a continuous random variable X, the hypothesis testing problem is to test the null hypothesis H_0 : $f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & o.w \end{cases}$ versus the alternate hypothesis $H_1 : f_X(x) = \begin{cases} 6x(1-x) & 0 < x < 1 \\ 0 & o.w \end{cases}$, the level α test based on one observation or a sample X of size 1 is to reject H_0 if $X(1-X) > a$. Identify the correct statements

(A) H_0 is a simple hypothesis, but H_1 is a composite hypothesis.

(B) The test criterion is to reject H_0 if $\frac{1}{2}(1-\alpha) < X < \frac{1}{2}(1+\alpha)$.

(C) The test criterion is to reject H_0 if $X > \frac{1}{2}(1+\alpha)$.

(D) The power of this test is more than α.

35. To estimate the number of fish say N in a lake, it was decided to throw a net and count the number of fish caught, let X denote the number of fish that will be caught this way, then

(A) X is an unbiased estimator of N.

(B) $2X$ is an unbiased estimator of N.

(C) $V(X)$ is a linear function of N.

(D) The unbiased estimator of N is consistent.

36. The 1000 members of a population are identified by numbers $1, 2, \ldots, 1000$, a sample of 20 members is to be drawn from this population, identify the correct statements

(A) There are 50 possible systematic samples and each one is equally likely to be the selected sample.

(B) Units bearing numbers 873 and 875 could be in the selected systematic sample.

(C) The mean of the systematic sample will give an unbiased estimate of the mean of the variable of interest of the population.

(D) The systematic sample mean is always a more efficient estimator for the population mean than the mean of the simple random sample without replacement of the same size.

37. $\mathbf{A}$ is a $n \times n$ real matrix whose rank is m and $m < n$. then

(A) all the columns of $\mathbf{A}$ are linearly independent.

(B) there exist $n-m$ linearly independent vectors $\underset{\sim}{\mathbf{x}}_1, \ldots, \underset{\sim}{\mathbf{x}}_{n-m}$, of $\mathbb{R}^n$ that are not equal to $\underset{\sim}{\mathbf{0}}$ and $\mathbf{A}\underset{\sim}{\mathbf{x}}_i = \underset{\sim}{\mathbf{0}}, \quad i = 1, \ldots, n-m$.

(C) $\mathbf{A}$ is a non-singular matrix.

(D) For a given $\underset{\sim}{\mathbf{b}} \in \mathbb{R}^n$, The linear equations $\mathbf{A}\underset{\sim}{\mathbf{x}} = \underset{\sim}{\mathbf{b}}$ may have infinitely many solutions.

38. If every arrangement in a row of 10 alike red balls and 8 alike blue balls is equally likely, the probability that the first and last balls in a randomly made arrangement are of the same colours is

(A) less than 1/3 (B) more than 1/3 but less than 1/2.

(C) more than 1/2, but less than 2/3. (D) more than 2/3.

39. 4 had to be added to the smallest of 25 numbers and 4 had to be subtracted from the largest of the original 25 numbers, for this new data of 25 numbers

(A) there is no change in the median.

(B) there is a change in the mean.

(C) The range and standard deviation will certainly reduce.

(D) the mean is the same as before and there may be a change in the range and standard deviation.

40. The random variable X takes values $1, 2, \ldots$ in accordance with the distribution $P(X = j) = \frac{1}{j(j+1)}, \quad j = 1, 2, \ldots$, given that X is more than 4, the probability that it is more than 9 is
(A) 0. (B) 1/2. (C) 2/3. (D) 3/4.

41. T_1 and T_2 are unbiased estimators for a parameter, $V(T_2) = 2V(T_1)$ and $Cov.(T_1, T_2) = \frac{V(T_1)}{2}$, then the most efficient estimator for the parameter is
(A) T_1. (B) $\frac{T_1+T_2}{2}$. (C) $\frac{T_1+3T_2}{4}$. (D) $\frac{3T_1+T_2}{4}$.

42. The probability distribution of a random variable X is $P(X = -j) = P(X = +j) = \frac{1}{15}$, $j = 1, 2, 3, 4, 5$, $P(X = 0) = \frac{1}{3}$, what is the probability that the quadratic equation $y^2 + Xy + 3$ has real roots?
(A) 1/15. (B) 2/15. (C) 4/15. (D) 8/15.

43. $\lim_{n\to\infty} \frac{1}{n}(n + \frac{n-1}{2} + \frac{n-2}{2^2} + \ldots + \frac{1}{2^{n-1}})$ is

(A) equal to 0. (B) is equal to 1. (C) equal to 2. (D) equal to ∞.

44. Vertices of a quadrilateral $ABCD$ are $A(0,0)$, $B(4,5)$, $C(9,9)$ and $D(5,4)$, what is the shape of the quadrilateral?

(A) Rectangle but not a square. **(B)** Rhombus.

(C) Parallelogram but not a rhombus. **(D)** Square.

45. A good way to recognize whether or not a variable is growing exponentially over time is by:

(A) plotting the variable against time and looking for a straight line pattern.

(B) calculating least squares regression line of the variable against time and examining the residuals.

(C) plotting the logarithm of the variable against time and looking for a straight line pattern.

(D) plotting the exponential of the variable against time and looking for a straight line pattern.

46. $1,3,3,2,4$ is a random sample from X that has the following probability distribution $P(X=1) = P(X=4) = 1/6$, $P(X=2) = p, P(X=3) = (2/3) - p, 0 < p < 2/3$, the Maximum Likelihood Estimate of p based on the given sample is

(A) 1/9. **(B)** 2/9. **(C)** 1/3. **(D)** 4/9.

47. The probability that the sample mean lies within 2 units from the population mean is 0.96, suppose the population variance is 16, what can you say about the sample size?

(A) It is equal to 104. **(B)** It is not more than 100. **(C)** It is at least 144

(D) Nothing can be said based on the information given.

48. Each of the events A, B, C occur with probabilities $1/2$, the joint occurrence of each pair that is $A \cap B$, $A \cap C$, $B \cap C$ have probabilities $1/4$ each, and all three, that is $A \cap B \cap C$ occur with probability $1/8$, the probability of at least two of them not occurring

(A) 1/16. **(B)** 1/8. **(C)** 3/8. **(D)** 1/2.

49. If an estimator T_n based on a sample of size n for a parameter θ is consistent, it means that

(A) for any sample of size n, the values of T_n will be close to θ.

(B) T_n are identically distributed for every sample size n.

(C) as the sample size increases the probability that T_n takes values really close to θ increases to 1.

(D) T_n is the same for every sample of size n.

50. A coin for which the probability of heads showing up on tossing is p was tossed 10 times, heads showed up 4 times and tails showed up 6 times, an unbiased estimate of p^2,

(A) 2/15. **(B)** 4/25. **(C)** 1/15. **(D)** can not determine.