B - 8

Entrance Examination : M.Sc. Statistics, 2013

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. There are a total of 50 questions in Part A and Part B together.
7. **Each question in Part - A has only one correct option and there is negative marking.**
8. **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**
9. 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.

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. Probability that at most one of the events A and B occurs is $7/8$, the probability of both A and B occurring

   (A) is $1/8$. **(B)** is $1/6$.

   (C) is $1/4$. **(D)** can not be determined from the information given.

2. The mean and median of 10 distinct real numbers are M and m respectively,now remove the largest of these numbers , let M' and m' denote the mean and the median respectively of these 9 numbers,then

   (A) $M' < M, m' > m$. **(B)** $M' > M, m' < m$.

   (C) $M' < M, m' < m$. **(D)** $M' > M, m' > m$.

3. 5 balls out of 30 are red how many of the remaining 25 should be blue and how many green so that the number of different possible arrangements in a row is maximized?

   (A) 9 blue & 16 green and vice versa.

   (B) 10 blue & 15 green and vice versa.

   (C) 11 blue & 14 green and vice versa.

   (D) 12 blue & 13 green and vice versa.

4. The sum of 10 positive numbers is 10 , then their product

   (A) is at least 10. **(B)** is at most 1.

   (C) is equal to 2. **(D)** is in the interval $(5, 6]$.

5. From a bag containing 2 balls each of 5 different colours , 2 balls are to be drawn without replacement , the probability that they will be of different colours is

   (A) $7/8$. **(B)** $8/9$. **(C)** $9/10$. **(D)** $11/12$.

6. The random variable $X \sim B(20, 1/4)$ its median is

   (A) in the set $\{12, 13, 14, 15\}$. **(B)** in the set $\{16, 17, 18\}$.

   (C) equal to 10. **(D)** less than 10.

7. $X_1, X_2, \ldots, X_n$ is a random sample from the $exp(\lambda)$ population. Let $\overline{X}, X_{(n)}$ and $X_{(1)}$ denote the sample mean , maximum and minimum respectively of the sample , an unbiased estimator of $\frac{1}{\lambda}$ is

   (A) $\bar{X}$.

   (B) $\frac{1}{\bar{X}}$.

   (C) $\frac{X_{(1)}+X_{(n)}}{2}$.

   (D) None of the above.

8. What is the number of subsets of 3 numbers from the set $\{1, 2, \ldots, 10\}$ that contain at least one multiple of 5 ?

   (A) 48. **(B)** 56. **(C)** 60. **(D)** 64.

9. The random variable X is uniformly distributed over $(-2, 2)$, that is $X \sim U(-2, 2)$, then ,

   (A) $P(-\frac{3}{2} < X \leq 0) > P(0 < X \leq \frac{3}{2})$.

   (B) $P(-\frac{1}{2} < X \leq \frac{1}{2}) > P(0 < X \leq 1)$.

   (C) $P(1 < X \leq \frac{4}{3}) = P(0 < X \leq \frac{1}{3})$.

   (D) $P(1 < X \leq \frac{4}{3}) = P(-\frac{1}{2} < X \leq \frac{3}{2})$.

10. If most of the measurements in a large data set are of approximately the same magnitude except for a few measurements that are quite a bit larger, then

    (A) the mean is smaller than the median and the histogram is skewed with a long left tail.

    (B) the mean is larger than the median and the histogram is skewed with a long right tail.

    (C) the mean is larger than the median and the histogram is skewed with a long left tail.

    (D) the mean is equal to the median and the histogram is symmetric.

11. A sample of 99 distances has a mean of 24 feet and a median of 24.5 feet. Unfortunately, it has just been discovered that an observation which was erroneously recorded as 30 actually had a value of 35. If we make this correction to the data, then:

    (**A**) the mean remains the same, but the median is increased

    (**B**) the mean and median remain the same

    (**C**) the median remains the same, but the mean is increased

    (**D**) we do not know how the mean and median are affected without further calculations; but the variance is increased.

12. According to Chebychev's inequality what can one say about the variance of a random variable for which the probability that it takes values more than 10 away from the mean is 0.1 ?

    (**A**) The variance is more than 10.

    (**B**) The variance is equal to 10.

    (**C**) The variance is less than 10.

    (**D**) Nothing definite can be said.

13. Many professional schools require applicants to take a standardized test. Suppose that 1000 students write the test, and you find that your mark of 63 (out of 100) was the 73rd percentile. This means:

    (**A**) At least 73% of the people got 63 or better.

    (**B**) At least 270 people got 73 or better.

    (**C**) At least 27% of the people got 73 or worse.

    (**D**) At least 270 people got 63 or better.

14. In hypothesis testing, β is the probability of committing an error of Type II. The power of the test, $1 - \beta$ is then:

    (**A**) the probability of rejecting H_0 when H_1 is true

    (**B**)the probability of failing to reject H_0 when H_1 is true

    (**C**) the probability of failing to reject H_0 when H_0 is true

    (**D**)the probability of failing to reject H_0.

B-8

15. The correlation coefficient provides a measure of the

(A) extent to which changes in one variable cause changes in another variable.

(B) strength of the linear association between two quantitative variables.

(C) strength of the linear association between two categorical variables.

(D) strength of the linear association between a quantitative variable and a categorical variable.

16. A fair coin is tossed three times. What is the probability that it lands on heads exactly once?

(A) 0.125 (B) 0.250 (C)0.333 (D) 0.375

17. Identify the correct statements from the following 3 claims. I. The standard error is computed solely from sample attributes. II. The standard deviation is computed solely from sample attributes. III. The standard error is a measure of central tendency.

(A) I only

(B) II only

(C)II and III only

(D) None of the above

18. Suppose X and Y are independent random variables. The variance of X is equal to 16; and the variance of Y is equal to 9. Let $Z = X - Y$. What is the standard deviation of Z?

(A) 2.65. (B) 5. (C) 7. (D) 25.

19. What is a statistical inference?

(A) A decision,estimate,prediction,or generalization about the population based on information contained in a sample.

(B) A statement made about a sample based on the measurements in that sample.

(C) A decision, estimate, prediction or generalization about sample based on information contained in a population.

(D) A set of data that characterizes some phenomenon.

20. Which of the following describes the Binomial random variable with parameters n, p?

(**A**) The number of heads that show up when n coins are tossed one time each.

(**B**) The number of heads that show up when a coin is tossed n times.

(**C**) The number of heads that show up when a coin is tossed $2n$ times.

(**D**) None of the above.

21. T_n is an unbiased estimator for a parameter μ , based on a random sample of size n and $V(T_n)$ is equal to $\frac{1}{n}$, an unbiased estimator for μ^2 based on the same sample is

(**A**) ${T_n}^2$. (**B**) ${T_n}^2 + T_n$.

(**C**) ${T_n}^2 + \frac{1}{n}$. (**D**) ${T_n}^2 - \frac{1}{n}$.

22. If a data set of numbers shows two modes,say m_1 and m_2 which appear n_1 and n_2 times respectively , also suppose $n_1 > n_2$, it means that

(**A**) $|m_1 - m_2|$ is less than the absolute difference between any other pair of numbers in the data.

(**B**) n_1 and n_2 are the highest frequencies.

(**C**) some other number might appear more often than m_2 but it is not close to m_2.

(**D**) m_1 and m_2 have the same largest frequencies.

23. $X \sim U(0,1)$ and $Y = -X$, $Cov.(X,Y)$ is equal to

(**A**) 0. (**B**) -1. (**C**) $\frac{1}{12}$. (**D**) $-\frac{1}{12}$.

24. A is a $n \times n$ singular matrix and every row of the matrix B is a linear combination of rows of A , the determinant of B

(**A**) is less than 0 (**B**) is more than 0

(**C**) is equal to 0. (**D**) can not be determined from the given information.

25. The negation of the statement "Some books in the library are more than 100 years old" is

(**A**) Some books in the library are not more than 100 years old.

(**B**) No book in the library is more than 100 years old.

(**C**) All the books in the library are more than 100 years old.

(**D**) At least one book in the library is at least 100 years old.

Part-B

- Questions (26)-(37) may 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) may 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 two events A and B have positive probabilities and $P(A|B) > P(A)$, then
(**A**) $P(B|A) > P(B)$ (**B**) $P(B|A) < P(B)$
(**C**) $P(A|B^c) < P(A)$ (**D**) $P(A^c|B^c) > P(A^c)$

27. For n unknown positive numbers $x_1, x_2, \ldots, x_n$ the sums of every pair and the differences of every pair are given but it is not known as to of which pairs the sums and differences are,with this information we can determine
(**A**) all the numbers. (**B**) the mean of the n numbers.
(**C**) the median of the n numbers. (**D**) the variance of the n numbers.

28. The probability of heads showing up upon tossing a coin is twice the probability of tails showing up, if this coin is to be tossed 10 times , the probability of

(**A**) all 10 tosses showing up heads is twice the probability of all 10 tosses showing up tails.

(**B**) 6 heads showing up is 16 times the probability of 6 tails showing up.

(**C**) 6 heads showing up is 4 times the probability of 6 tails showing up.

(**D**) 7 heads showing up is 16 times the probability of 7 tails showing up.

29. Regarding the number of different arrangements of 10 red (all alike) and 5 blue (all alike) in a row

(**A**) more than $(1/3)^{rd}$ of the arrangements have red balls at both the ends.

(**B**) more than $(1/3)^{rd}$ of the arrangements have blue balls at both the ends.

(**C**) almost half the arrangements have balls of different colours at either end.

(**D**) less than $(1/10)^{th}$ the arrangements have all the 5 blue balls in the first 10 places.

30. Given below are marks of 6 students in 2 subjects S_1 and S_2

S_1	42	59	62	66	75	52
S_2	46	63	66	70	79	56

(**A**) The variances are the same for the marks in S_1 and S_2 and so are the mean deviations about the respective medians.

(**B**)The variances are the same but the mean deviation about the median in the marks in S_1 is more than the marks in S_2.

(**C**) The correlation coefficient between the marks in the 2 subjects is $+1$.

(**D**)The correlation coefficient between the marks in the 2 subjects is not $+1$ but is greater than 0.

31. X is a discrete random variable for which $P(X \geq 1) = 1, P(X \geq 2) = 0.8, P(X \geq 3) = 0.5, P(X \geq 4) = 0.3, P(X \geq 5) = 0.2, P(X \geq 6) = 0.1$, so

(**A**) $P(X = 7) = 0$ (**B**) $P(X = 2) = 0.3$

(**C**) $E(X)$ can not be determined. (**D**) $E(X)$ can be determined and is less than 3.

32. 10% of the students of a class weigh at least $65kg.$, 40% of them weigh at least $55kg.$ and less than $65kg.$, 35% of them weigh at least $45kg.$ and less than $55kg.$, 15% of them weigh at least $40kg.$ and less than $45kg.$

(**A**) The average weight of the students of this class is less than $45kg.$

(**B**) The average weight is more than $50kg.$

(**C**) The median weight is more than $55kg.$

(**D**) The median weight is about $45kg.$

33. The probability of heads showing up upon tossing a certain coin is p, this coin is tossed 3 times , let $X_i, i = 1, 2, 3$ be 1 or -1 depending on the outcome of the i^{th} toss being heads or tails respectively.

(**A**) $X_1 + X_2 + X_3$ is a sufficient statistic for p.

(**B**) ${X_1}^2 + {X_2}^2 + X_3^2$ is a sufficient statistic for p.

(**C**) $X_1X_2X_3$ is a sufficient statistic for p.

(**D**) ${X_1}^3 + {X_2}^3 + X_3^3$ is a sufficient statistic for p.

34. In general, which of the following statements are true?

(**A**) The sample mean is more sensitive to extreme values than the median.

(**B**) The sample range is more sensitive to extreme values than the standard deviation.

(**C**) The sample standard deviation is a measure of spread around the sample mean.

(**D**)The sample standard deviation is a measure of central tendency around the median.

35. There are 10 girls and 30 boys in M.Sc.(Mathematics) , while in M.Sc.(Statistics) there are 4 girls and 16 boys, 2 students are randomly selected from both the classes,then

(**A**) the probability of selecting both girls is more from M.Sc.(Statistics).

(**B**) the probability of selecting one girl and one boy is more from M.Sc.(Mathematics).

(**C**) the probability of selecting both boys is more from M.Sc.(Statistics).

(**D**) the probability of girls outnumbering boys among the the 4 selected is more than 1/10.

36. Consider the following real valued function defined on $\mathbf{R}$

$$f(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} + \frac{x}{2} & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

The function f is

(**A**) non-decreasing. (**B**) not continuous at $x = 0$.

(**C**) continuous and differentiable everywhere. (**D**) not continuous at $x = 1$.

37. X_1, X_2, X_3 are independent and identically distributed random variables with mean θ and finite variance σ^2 , the two statistics $Y_1 = \frac{X_1+X_2+X_3}{3}$ and $Y_2 = \frac{X_1+2X_2-X_3}{2}$ are both unbiased estimators for θ and also

(**A**) $V(Y_1) \geq V(Y_2)$. (**B**) $V(Y_1) < V(Y_2)$.

(**C**) Y_2 is a better estimator for θ than Y_1 because Y_2 has a smaller denominator.

(**D**) Y_1 is a better estimator for θ than Y_2 because Y_1 is less likely to take values far from θ than Y_2.

The next 13 questions have only one correct option.

38. The number of accidents per day on the Hyderabad - Benguluru highway is a Poisson random variable with parameter λ , on 10 randomly chosen days the number of accidents were observed as $1, 0, 1, 1, 2, 0, 2, 0, 0, 1$, an unbiased estimate of e^{λ} is
(**A**) $e^{0.8}$ (**B**) 0.8 (**C**) 2 (**D**) e^2

39. Of 5 girls and 35 boys of a class 5 were absent on a particular day and a class representative was randomly selected from among those present.

(**A**) The probability that a girl was selected as class representative is 1/8.

(**B**) The probability that a girl was selected as class representative is 1/16.

(**C**) The expected number of girls present is 2.

(**D**) The expected number of absentee boys is less than 4.

40. $\lim_{n\to\infty}\left(1+\frac{e}{n}\right)^{-n}$
(**A**) does not exist. (**B**) is e^{-e}. (**C**) is e^{e}. (**D**) is $e^{e^{-1}}$.

41. $1, 2, 3, 3, 4$ is a random sample from X that has the following probability distribution $P(X = 1) = P(X = 2) = 1/4$, $P(X = 3) = p, P(X = 4) = (1/2) - p, 0 < p < 1/2$, the Maximum Likelihood Estimate of p based on the given sample is
(**A**) 1/2. (**B**) 1/3. (**C**) 1/4. (**D**) 1/6.

42. The random variable X has Poisson distribution and its 2^{nd} raw moment is equal to 3 times the 1^{st} raw moment , so
(**A**) $P(X = 1) = P(X = 0)$. (**B**) $P(X = 1) = 2P(X = 0)$.
(**C**) $P(X = 1) = \frac{1}{2}P(X = 0)$. (**D**) The statements (**A**),(**B**),(**C**) are wrong.

43. A fair coin is tossed till the first heads shoes up , suppose all the first 10 tosses showed tails,then the probability that heads will show up on or before the 20^{th} toss is
(**A**) $\frac{1}{512}$ (**B**) $\frac{1}{1024}$ (**C**) $\frac{511}{512}$ (**D**) $\frac{1023}{1024}$

44. X is a normally distributed random variable with mean 22 and variance 16 , that is $X \sim N(22, 16)$ and suppose $Y \sim N(20, 25)$, then

(**A**) $P(X > 30) < P(Y > 30)$.

(**B**) $P(X > 34) > P(Y > 34)$.

(**C**) $P(X > 30) = P(Y > 30)$.

(**D**) $P(X > 28) = P(Y > 29)$.

45. At least one of the 3 events A, B, C will certainly occur, the probability that exactly one of the 3 will happen is 0.3 and the probability that exactly 2 of the 3 events will happen is 0.6 ,from this information

(**A**) the probability that all the 3 events will happen can be determined and is equal to 0.1.

(**B**) the probability that all the 3 events will happen can not be determined.

(**C**) the probability that none of the 3 events will happen can be determined and is equal to 0.1.

(**D**) the probability that at least one of the 3 events does not occur can not be determined.

46. The probability distribution of a random variable X is $P(X = k) = p_k, k = 0, 1, \ldots$ and $\frac{p_k}{p_{k-1}} = a + \frac{b}{k}$ if $p_0 = 1/9, p_1 = p_2 = 4/27$, what can you say about the sequence of probabilities $\{p_0, p_1, p_2, p_3, \ldots\}$?

(**A**) They are in Arithmetic Progression (**B**) They are in Geometric Progression.

(**C**) They are terms in a certain binomial expansion.

(**D**) Neither of (**A**),(**B**) and (**C**) is correct.

47. The α level critical region for the hypothesis testing problem $H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$ is C where μ is the mean of a normal random variable with variance 2 , if the problem is changed to $H_0 : \mu = \mu_1$ vs. $H_1 : \mu > \mu_1$ where $\mu_1 > \mu_0$,what can we say about the probability of $Type - 1$ error with the same critical region C?

(**A**) It is equal to α. (**B**) It is less than α.

(**C**) It is more than α. (**D**) Nothing definite can be said.

48. $X_1 \sim exp(3)$ and $X_2 \sim exp(6)$, let x_{01} and x_{02} denote the medians of X_1 and X_2 respectively , then

(**A**) $x_{01} = x_{02}$. (**B**) $x_{01} = \frac{1}{2}x_{02}$.

(**C**) $x_{01} = 2x_{02}$. (**D**) $x_{01} = x_{02} + 3$.

49. The non - constant random variables X and Y have the following property: $V(X+Y) = 2V(X-Y)$, it follows that

(**A**) X and Y are independently distributed.

(**B**) X and Y are uncorrelated but not independently distributed.

(**C**) X and Y are negatively correlated.

(**D**) X and Y are positively correlated.

50. The proportion of work that can be completed in a project in the first year is a random variable with density function

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

The probability that more than $(3/4)^{th}$ of the work will be completed in the first year is

(**A**) more than 3/4.

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

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

(**D**) more than 1/4 but less than 1/3.