Entrance Examination : M.Sc. Statistics, 2019

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 of 0.33.**
8. **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.**
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.
10. **The maximum marks for this examination is 100, 25 for Part-A and 75 for Part-B, there will be NO INTERVIEW.**
11. 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. $Rank(A)$ means rank of the matrix A. Members of $\mathbb{R}^n$ are column vectors $\underset{\sim}{x}$ and $\underset{\sim}{0}$ is the column vector of zeros.
12. This book contains **11** pages including this page and excluding pages for rough work. Please check that your paper has all the pages.

T-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. The heights of adult females in a certain country are normally distributed with mean μ and variance σ^2. Let $X_1, X_2, \ldots, X_{10}$ denote the heights of a random sample of 10 ladies from this population. Given below are 3 statements:

   I. $X_1, X_2, \ldots, X_{10}$ are independent random variables.

   II. The height of the tallest lady in the selected sample is a statistic.

   III. The height of the tallest person in any sample is more than μ.

   IV. The average height of the ladies in some samples is more than μ.

   The correct statements are

   (A) All of them. **(B)** only I,III and IV. **(C)** only I and III. **(D)** only I,II and IV.

2. The probability of at least one of the two equally probable events A and B occurring is 0.7 and the probability of their joint occurrence is 0.3. The probability of only A occurring is

   (A) 0.3. **(B)** 0.2. **(C)** 0.1. **(D)** 0.

3. A positive number d, not more than the largest of 10 numbers is added to the smallest of 10 numbers and subtracted from the largest of these 10 numbers, this will result in

   (A) no change in either the mean, the median or the standard deviation.

   (B) increase in the mean, no change in the either the median or the standard deviation.

   (C) no change either the mean or the median, but standard deviation may change.

   (D) no change in the mean, but there may be some change in the median and also in the standard deviation.

4. The words one, two, three, four and five are to be typed once, twice, three times, four times and five times respectively, the number of keyboard clicks is

   (A) 60. **(B)** 64. **(C)** 74. **(D)** 86.

5. The random variable X has Poisson distribution and $P(X = 1) = 2P(X = 0)$, the variance of X is

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

6. For which of the following random variables is the standard deviation always equal to the mean irrespective of the values of the parameters?

   (A) Normal. **(B)** Exponential. **(C)** Uniform over $(a, b]$. **(D)** Poisson.

7. The marks obtained by students who appeared for a public exam are positively skewed with the average being 50, this means that

   (A) More than half of them got more than 50 marks.

   (B) Half of them got more than 50 marks.

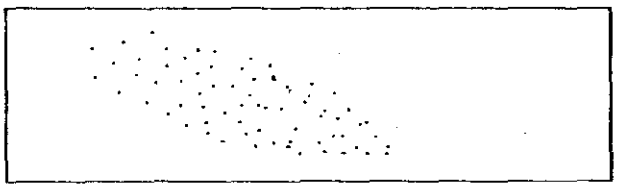
   (C) More than half of them got less than 50 marks.

   (D) None of the above.

8. The numerals $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ are randomly arranged in a row to get a 9 or 10 digit number, if every arrangement is equally likely, the probability that the number so obtained is an odd multiple of 3 is

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

9. What could the correlation coefficient based on n observations on the random variables X and Y be if the scatter plot for them is



   (A) $-1/3$. **(B)** $-99/100$. **(C)** 1/5. **(D)** 99/100

10. Identify the correct statement about the probability distribution of $X \sim P(5)$, that is the Poisson random variable with parameter 5,

    (A) $P(X = 3) < P(X = 4) < P(X = 5) < P(X = 6)$.

    (B) $P(X = 4) < P(X = 5)$ and $P(X = 7) < P(X = 6) < P(X = 5)$.

    (C) $P(X = 5) < P(X = 6) < P(X = 7) < P(X = 8)$.

    (D) $P(X = 3) < P(X = 5) < P(X = 4)$.

T-7

11. Write down each of the letters of the word PROBABILITY on 11 pieces of papers and put in a bag, then draw 4 of these slips, the probability that the word BOLT can be formed from the selected letters is

(A) less than $1/200$. **(B)** in the interval $(1/200, 1/100]$

(C) more than $4/11$ **(D)**in the interval $(1/11, 4/11]$

12. Every trial results in success or failure, the probability of success in the i^{th} trial is p_i, the correlation coefficient between number of successes and number of failures in 100 trials is

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

13. The incomes of three people in a start up office are $Rs.20000$ and the incomes of the other two are $Rs.10000$ per month. The probability that the average income of a simple random sample without replacement of three of these people is more than $Rs.15000$ is

(A) 0.58. **(B)** 0.6 **(C)** 0.63. **(D)** 0.7.

14. Let C_1 and C_2 be critical regions for testing a null hypothesis H_0 against an alternate hypothesis H_1 at levels of significance of α_1 and α_2 respectively, if $\alpha_1 < \alpha_2$

(A) $C_1 \subset C_2$. **(B)** $C_2 \subset C_1$. **(C)** $C_1 = C_2$. **(D)** $C_1 \cap C_2 = \emptyset$.

15. T_1 and T_2 are unbiased estimators of μ and μ^2 respectively, an unbiased estimator of $V(T_1)$ is

(A) T_1^2. **(B)** $T_2^2 - T_1^2$. **(C)** T_2. **(D)** $T_1^2 - T_2$.

16. In a public exam 10% of the candidates got 40 or more, but below 50 marks, 65% got 50 or more, but below 65 marks, 20% of them got 65 or more, but below 80 marks and the rest of them got 80 and more, but below 92 marks, therefore, the average marks of all the candidates who appeared in this exam is

(A) at least 53.5%. **(B)** at least 62.5%. **(C)** more than 69%. **(D)** less than 53.5%.

17. A coin for which the probability of heads showing up when tossed is p, $0 < p < 1$ was tossed 12 times and heads showed up 4 times, the maximum likelihood estimate for p

(A) can not be determined . **(B)** is $1/2$. **(C)** is $1/3$. **(D)** is $1/4$.

18. The percentages of votes polled by 7 political parties is available for 3 districts of a state. The most suitable way to display the differences in the percentages polled by the different parties is

(A) Bar charts. **(B)** Histogram.
(C) Stem and leaf plot. **(D)** Pie Chart.

19. X_1, X_2 is a random sample from the Bernoulli random variable $B(p)$, we say $X \sim B(p)$ if $P(X = 1) = p$; $P(X = 0) = 1 - p, 0 < p < 1$, the statistic $X_1 - X_2$

(A) is a sufficient statistic for p because its values tell us as much about p as the sample X_1, X_2.

(B) is not a sufficient statistic for p because its values do not tell us as much about p as the sample X_1, X_2.

(C) is not a sufficient statistic because it can take negative values.

(D) is a sufficient statistic because its expected value is 0.

20. The probability distribution of a random variable X is $P(X = -2) = P(X = 2) = 1/5$; $P(X = -1) = P(X = 1) = 1/4$; $P(X = 0) = 1/10$, $Cov.(X, |X|)$ is

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

21. T_1 and T_2 are two unbiased estimators for $g(\theta)$ - a function of a parameter θ, if $P(|T_1 - g(\theta)| > a) \geq P(|T_2 - g(\theta)| > a)$, $\forall a > 0$, then

(A) one can use any of them to estimate $g(\theta)$.

(B) T_1 is preferable to T_2 as it is more efficient.

(C) T_2 is preferable to T_1 as it is more likely to be close to $g(\theta)$.

(D) One can't say which is more efficient, so one can't say which is preferable.

22. In an hypothesis testing problem, suppose the test criterion is to reject the null hypothesis H_0 vs. H_1at 5% level of significance if $T(X_1. \ldots, X_n) > a$ where $T(X_1, \ldots, X_n)$ is a statistic based on a random sample $X_1, \ldots, X_n$, this means

(A) If H_1 is true, 5% of all samples of size n are such that $T(X_1, \ldots, X_n) \leq a$.

(B) If H_1 is true, 5% of all samples of size n are such that $T(X_1, \ldots, X_n) > a$.

(C) If H_0 is true, 95% of all samples of size n are such that $T(X_1, \ldots, X_n) \leq a$.

(D) If H_0 is true, 95% of all samples of size n are such that $T(X_1, \ldots, X_n) > a$.

23. A and B are two subsets of Ω, the complement of the set $A \triangle B$ is the set of all those elements of Ω which

    (A) belong to exactly one of A and B.

    (B) belong to neither A nor B.

    (C) belong either to both A and B or to neither A nor B.

    (D) do not belong to at least one of A and B

24. An estimator T based on a random sample from a certain population for which θ is a parameter is symmetrically distributed about θ, this implies that

    (A) T is a sufficient statistic for θ.

    (B) T is an unbiased estimator for θ.

    (C) T is an efficient estimator for θ.

    (D) T is a consistent estimator for θ.

25. Suppose R_1, R_2, R_3 and R_4 are the first, second, third and fourth rows of a 4×4 real matrix **A** whose rank is 3. Let **B** be the matrix in which the first, second, third and fourth rows are R_1, $R_1 + R_2$, $R_1 + R_2 + R_3$ and $R_1 + R_2 + R_3 + R_4$ respectively, the rank of **B** is

    (A) is 3. **(B)** is 2. **(C)** is 4. **(D)** can not be determined from the information given.

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. A random experiment is one

(A) that has at least two possible outcomes.

(B) whose outcomes can be different when repeated.

(C) whose outcomes are well defined.

(D) that has a unique possible outcome..

27. The heights of Indian females are normally distributed, the average height is $152.2cm.$ and their standard deviation is $10cm.$. If $Z \sim N(0,1)$ and $P(Z \leq 1.25) = 0.8944$; $P(Z \leq 2.21) = 0.9864$; $P(|Z| > 3) = 0.0027$; $P(|Z| > 2) = 0.0456$, identify the correct statements.

(A) The heights of more than 10% of Indian females are at least $164cm.$.

(B) The heights of less than 2% of Indian females are less than $130cm.$.

(C) The heights of more than 2% of Indian females are at least $183cm.$.

(D) The heights of less than 9.3% of Indian females are between 168 and $172cm.$.

28. $X_1 \sim N(0,1)$ and $X_2 \sim U((-a.a])$ where $a > 0$ is such that $V(X_2) = 1$, then

(A) $P(X_1 > 2) > P(X_2 > 2)$.

(B) $P(|X_1| > 1.25) < P(|X_2| > 1.25)$.

(C) $x_1 = x_2$, where $P(X_1 \leq x_1) = P(X_2 \leq x_2) = 1/2$.

(D) $x_1 = x_2$, where $P(X_1 \leq x_1) = P(X_2 \leq x_2) = 0.00135$.

29. Identify the correct statements regarding some measures of central tendency.

(A) Every distribution has a unique mode.

(B) Every distribution has a unique median.

(C) A distribution may have no median.

(D) A distribution may have several modes.

30. Two treatments T_1 and T_2 for hypertension were administered to two sets each of 25 hypertension patients. The mean time to blood pressure becoming normal for patients who received T_1 was less than the same for patients who received T_2.

    (A) This means that T_1 is better than T_2.

    (B) One can not say T_1 is better than T_2 because we don't know how similar or different the two sets of patients are.

    (C) One can not say T_1 is better than T_2 because we don't know how the patients who received T_1 and T_2 were selected.

    (D) One can say that T_1 is better than T_2 if all the 50 patients are of the same age.

31. The *pdf*s of two random variables X and Y are as given below, identify the correct statements

$$f_X(x) = \begin{cases} 20x^3(1-x) & 0 < x < 1 \\ 0 & o.w \end{cases} \qquad f_Y(y) = \begin{cases} 20y(1-y)^3 & 0 < y < 1 \\ 0 & o.w \end{cases}$$

    (A) X is a positively skewed and Y is a negatively skewed random variable.

    (B) $Y = 1 - X$.

    (C) X and $1 - Y$ are identically distributed.

    (D) X and Y are independently distributed.

32. Which of the following random variables do not follow Negative Binomial distribution.

    (A) the draw number in which the third red ball appears in draws of one ball each without replacement from a bag containing 5 red and 10 blue balls.

    (B) the draw number in which the third red ball appears in draws of one ball each with replacement from a bag containing 5 red and 10 blue balls.

    (C) the number of problems tried till 4 problems are solved.

    (D) The number of tails till the five heads show up in consecutive tosses of a coin.

33. The null and alternate hypotheses for a random variable X are $H_0 : X \sim U((-1, 1])$ and $H_1 : X \sim U((-2, 2])$, based on a single observation X, the test is to reject H_0 if $|X| > 0.975$.

    (A) H_0 and H_1 are simple hypotheses.

    (B) The size of the test is .05.

    (C) The power of the test is more than 0.95.

    (D) The power of the test is less than 0.6.

34. $X_1, X_2, \ldots, X_n$ is a random sample from the random variable with *pdf* $f(x) = \begin{cases} e^{-(x-\mu)} & x \geq \mu \\ 0 & o.w \end{cases}$

(A) The Maximum Likelihood estimator for μ is also an unbiased estimator for μ.

(B) If a sample of size 5 is $2.1, 3.4, 0.8, 1.5, 3.2$, the maximum likelihood estimate is 0.8

(C) based on the same sample as in **(B)**, an unbiased estimate for μ is 0.6

(D) based on the same sample as in **(B)**, an unbiased estimate for μ is 1.2

35. $\mathbf{A}$ is a $n \times n$ real matrix whose first column is a linear combination of the other columns, this means

(A) there exists a non zero vector $\underset{\sim}{\mathbf{x}} \in \mathbb{R}^n$, such that $\mathbf{A}\underset{\sim}{\mathbf{x}} = \underset{\sim}{\mathbf{0}}$

(B) the determinant of $\mathbf{A}$ is equal to 0.

(C) all the rows of $\mathbf{A}$ may be linearly independent.

(D) the rank of $\mathbf{A}$ is $n - 1$.

36. Which of the following imply that the random variables X and Y are independent?

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

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

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

(D) $P(x_1 < X \leq x_2, Y \leq y) = P(x_1 < X \leq x_2)P(Y \leq y), \quad \forall x_1, x_2, y \in \mathbb{R}$ and $x_1 < x_2$.

37. The function $F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} + \frac{x}{2} & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$

(A) is a *pdf* of some random variable.

(B) is not continuous at 0 and 1.

(C) is not continuous only at 0.

(D) is non-decreasing and bounded.

38. A number is selected from $1, 2, 3, \ldots, 100$ with equal probabilities and the selected number is typed as many times. For example if the number 25 is selected, 25 is typed 25 times. So the expected number of keyboard clicks is

(A) less than 75. (B) 75 or more, but less than 100.
(C) 100 or more, but less than 120. (D) more than 120.

T-7

39. The expected number of different colours in a draw of 3 balls from a bag containing 2 red, 3 blue and 4 green balls is in the interval

(A) $(0.62, 1.3]$. **(B)** $(1.3, 2]$. **(C)** $(2, 2.4]$. **(D)** $(2.4, 2.7]$.

40. A fair 6 faced die is rolled once and the then a fair coin is tossed $j + 1$ times if the die showed j, $j = 1, 2, 3, 4, 5, 6$. The probability of 4 heads occurring is in the interval

(A) $(0, 1/16]$. **(B)** $(1/16, 1/8]$. **(C)** $(1/8, 1/4]$. **(D)** $(1/4, 1/2]$.

41. 2 and 4 are two independent observations of a $B(5, p)$ random variable, an unbiased estimate of $(1+p)^5$

(A) is $(1.6)^5$. **(B)** is 10. **(C)** is 3. **(D)** can not be determined.

42. X_1 and X_2 are *iid* Poisson random variables parameter 2. $P(X_1 = 2|X_1 + X_2 = 4)$ is equal to

(A) is 1/8. **(B)** is 1/2. **(C)** is 3/8. **(D)** 3/4.

43. X_1 and X_2 and are independent geometric random variables with parameters 1/2 and 1/3 respectively, that is $P(X_2 = j) = \frac{1}{3}(\frac{2}{3})^{j-1}$, $j = 1, \ldots$, the probability distribution of X_1 can be similarly written. $P(X_2 > X_1)$ is equal to

(A) is 1/8. **(B)** is 1/2. **(C)** is 3/8. **(D)** 3/4.

44. 98% of all random samples of size 100 are such that their mean is within 1 unit from the population mean μ, about the population variance one can say

(A) nothing, based on the information given.

(B) that it could be 1.5.

(C) that it is strictly less than 1.

(D) that it is at least 2.

45. How should the numbers $1, 2, \ldots, n, n+1, \ldots, 4n$ be divided into two sets with $2n$ numbers each such that the sum of the variances of numbers in the two sets is the least?

(A) the numbers $1, 2, \ldots, 2n$ in one set and the rest in the other set.

(B) the numbers $1, 2, \ldots, n, 3n+1, 3n+2, \ldots, 4n$ in one set and the rest in the other set.

(C) All the even numbers in one set and all the odd numbers in the other set.

(D) None of the above will achieve what is needed.

T-7

46. A fair 6 faced die is rolled once and the then a fair coin is tossed $j+1$ times if the die showed j, $j=1,2,3,4,5,6$, the expected number of heads that will show up is .

(A) 27/6. **(B)** 21/12. **(C)** 21/12. **(D)** 27/12.

47. $0.7, 2.3, 1.8, 3.0, 1.5, 0.4, 2.6, 3.2$ is a random sample from the $U((0,\theta])$ distribution, an unbiased estimate for θ is

(A) 3.2. **(B)** 3.6. **(C)** 1.9375. **(D)** 3.1.

48. The total number of cellphones in a simple random sample without replacement of 50 households out of 1000 households in a locality is 148, an unbiased estimate of the total number of cellphones in all the households of the locality is

(A) 2960. **(B)** 1480. **(C)** 740. **(D)** 3700.

49. The product of 4 distinct positive numbers is 81, the sum of their squares

(A) is less than 30.

(B) could be 36.

(C) could be 30.

(D) is more than 36.

50. The value of $\lim_{n\to\infty} \sum_{j=1}^{n} \frac{j}{2^j}$

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

T-7

University of Hyderabad

Revised Key

Entrance Examinations - 2019

School/Department/Centre : Mathematics and Statistics

Course/Subject :M.Sc.(Statistics)

Q.No.	Answer	Q.No.	Answer	Q.No.	Answer	Q.No.	Answer
1	D	26	ABC	51		76	
2	B	27	ABD	52		77	
3	D	28	ABC	53		78	
4	A	29	D	54		79	
5	C	30	BC	55		80	
6	B	31	C	56		81	
7	C	32	ACD	57		82	
8	A	33	AD	58		83	
9	A	34	BC	59		84	
10	XXXX cancel	35	AB	60		85	
11	B	36	D	61		86	
12	A	37	C	62		87	
13	D	38	C	63		88	
14	A	39	C	64		89	
15	D	40	B	65		90	
16	A	41	B	66		91	
17	C	42	C	67		92	
18	D	43	B	68		93	
19	B	44	D	69		94	
20	D	45	A	70		95	
21	C	46	D	71		96	
22	C	47	B	72		97	
23	C	48	A	73		98	
24	B	49	D	74		99	
25	A	50	C	75		100	

Note/Remarks : Q.No-10, Benefit will be given to all candidates.

DEAN
School of Mathematics & Statistics
University of Hyderabad
HYDERABAD - 500 046, T.S.

Faculty Incharge

Signature
School/Department/Centre