Seat No.

Total No. of Pages : 14

Ph.D. Entrance Examination 2025
STATISTICS
Subject Code: 58794

Day and Date : Wednesday, 10-Sep-2025 **Total Marks : 100**
Time : 10.00 am to 12.00 noon

Instructions:

1) All questions are compulsory.
2) Each question carries 2 mark.
3) Answers should be marked in the given OMR answer sheet by darkening the appropriate option.
4) Follow the instructions given on OMR sheet.
5) Rough work shall be done on the sheet provided at the end of question paper.

1 Research is basically .....

i) Collecting the data

ii) Search of truth

iii) A systematic exploration of facts

A) Only iii) B) Only ii)

C) Only i) and iii) D) Only ii) and iii)

2 What is research design in the context of a research study?

A) The overall plan or structure of the study

B) The sampling technique used

C) The statistical methods applied

D) The data collection instruments chosen

3 In a hypothesis testing, what does the null hypothesis typically represent?

A) The absence of an effect or relationship

B) The researcher's expectations

C) A strong positive correlation

D) None of the above

4 Which of the following is not a type of research design?

A) Descriptive design B) Experimental design

C) Correlational design D) Probability design

5 A research tool that measures an individual's abilities, skills, knowledge, or personality is known as a:

A) Test B) Questionnaire

C) Rating scale D) Guttman scale

6 A research study that aims to diagnose the factors contributing to a specific social problem is categorized as:

A) Exploratory research B) Diagnostic research

C) Evaluation research D) Action research

7 Which of the following is NOT considered a form of plagiarism in academic research?

A) Copying and pasting text from a source without proper citation

B) Paraphrasing a source without acknowledging the original author

C) Quoting directly from a source with appropriate citation

D) Referencing widely known facts without citation

8 Self-plagiarism occurs when a researcher:

A) Uses his/her own previously published work without proper citation

B) Accidentally includes someone else's work in their research

C) Fails to conduct a literature review before starting a study

D) Collaborates with other researchers on a project

9 What is the role of 'curiosity' in the qualities of a good researcher?

A) Curiosity helps in exploring new avenues and generating innovative ideas

B) Curiosity avoid asking too many questions and leads to faster publication

C) Curiosity stick researchers to existing knowledge and avoid exploration

D) Curiosity hinders the research process by distracting researchers

10 If all cats are mammals, and Mittens is a cat, what can we logically conclude?

A) Mittens is not a mammal

B) Mittens is a mammal

C) Not all mammals are cats

D) Mittens may or may not be a mammal

11 A person travelled a distance of 610 km in 9 hours. He travelled the first phase at a speed of 40 kmph and the rest at 90 kmph. The distance travelled during the first phase is

A) 140 km B) 150 km

C) 160 km D) 170 km

12 The difference between a number and its three-fifths is 50. What is the number?

A) 75 B) 100

C) 125 D) 150

13 In a colony, there are 50 members. Every member posts a greeting card to all the members. How many greeting cards were posted by them?

A) 990 B) 1225

C) 2500 D) 2450

14 Which one of the following probability distributions is used for determining whether or not a population has a Poisson probability distribution?

A) Normal distribution B) F-distribution

C) Chi-square distribution D) Binomial distribution

15 In the Gauss elimination method for solving a system of linear algebraic equations, triangularization leads to...

A) Diagonal matrix B) Lower triangular matrix

C) Upper triangular matrix D) Singular matrix

16 What is a Bayesian prior?

A) A summary statistic of the data

B) A distribution representing beliefs about a parameter before observing data

C) A test statistic for testing hypothesis

D) The likelihood of observing specific data

17 The purpose of statistical inference is

A) To collect sample data and use them to formulate hypotheses about a population

B) To draw conclusion about populations and then collect sample data to support the conclusions

C) To draw conclusions about populations from sample data

D) To draw conclusions about the known value of population parameter

18 The Newton-Raphson method formula for finding the root of the equation $x^2 - 3x + 2 = 0$ is,

A) $x_{i+1} = \frac{1}{2x_i - 3}$

B) $x_{i+1} = \frac{x_i^2}{2x_i - 3}$

C) $x_{i+1} = \frac{x_i^2 + 2}{2x_i - 3}$

D) $x_{i+1} = \frac{x_i^2 - 2}{2x_i - 3}$

19 What is the primary purpose of bootstrap resampling in statistics?

A) To reduce bias in the data

B) To increase the sample size

C) To estimate the sampling distribution of a statistic

D) To eliminate outliers from the dataset

20 Which metric is commonly used to evaluate the performance of a classification model?

A) Mean Squared Error

B) R-square

C) Standard Deviation

D) Accuracy

21 How can a correlation matrix assist in EDA?

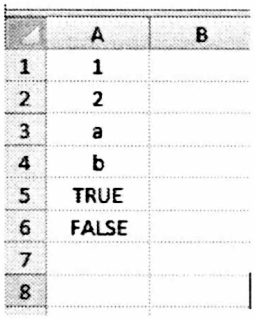
A) It measures the spread of the data

B) It identifies outliers in the dataset

C) It quantifies the strength and direction of relationships between variables

D) It provides a summary of descriptive statistics

22 Consider the following data in EXCEL.



	A	B
1	1	
2	2	
3	a	
4	b	
5	TRUE	
6	FALSE	
7		
8		

What will be the output of the function "-COUNT(A1:A6)"?

A) 2 B) 4

C) 5 D) 6

23 Which of the following statement(s) is/are true about linear Congruential random number generator $X_{n+1} = a * X_n + b \ (mod\ m)\ for\ n = 0, 1, 2, \ldots$?

i) If we choose m=32, a=9 and b = 21 then maximum cycle length will be achieved.

ii) Maximum possible cycle length is always achieved.

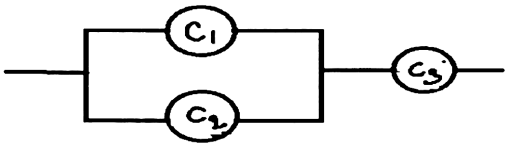
A) Both i) and ii) are true

B) Only i) is true

C) Only ii) is true

D) Both i) and ii) are not true

24 Suppose three components are arranged in a system as follows.



That is, system works if component C_3 and at least one component from C_1 and C_2 works. The life time of component C_i has Exp(i) distribution i = 1, 2, 3. The algorithm to approximate the expected lifetime of the system by simulation is,

- For i = 1, 2 ,.....,1000
  - o Generate random observations $X_{i,1}$ from Exp(1), $X_{i,2}$ from Exp(2) and $X_{i,3}$ from Exp (3)
  - o Compute Y_i = ...
- Compute sample mean of Y_t' s

A) $min\{X_{i,1}, X_{i,2}\} + X_{i,3}$

B) $min\{max\{X_{i,1}, X_{i,2}\}, X_{i,3}\}$

C) $max\left\{min\{X_{i,1}, X_{i,3}\}, min\{X_{i,2}, X_{i,3}\}\right\}$

D) $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$

25 Which of the following is NOT true?

A) Jackknife technique is a resampling technique

B) Resampling techniques provide inference based on a wide range of statistics under very general conditions

C) Resampling methods involve constructing hypothetical populations

D) Jackknife technique cannot be used for robust confidence interval estimation

26 The totals of the response observations of the two replicates of 2^2 factorial experiment with factors A and B are, namely, (1) = 100, a = 50, b = 110 and ab = 120. Then, β_1 in the regression model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ for this experiment is ....

A) -10

B) 10

C) 5

D) -5

27 Let m_1 and m_2 be the first and second sample moments based on a random sample drawn from Gammaa, (α, β) distribution. Then, the consistent estimator of a based on the method of moments is .....

A) $\frac{m_1^2}{m_2 - m_1^2}$ B) $\frac{m_2^2}{m_2 - m_1^2}$

C) $\frac{m_2 - m_1^2}{m_1}$ D) $\frac{m_2^2 - m_1^2}{m_1}$

28 Let T be a sufficient statistic for θ. If both T and θ are vectors, then, ...

A) the dimensions of T and θ are the same.

B) the first component of T is sufficient for the first component of θ.

C) a one-to-one function of T need not be sufficient for θ.

D) in general, a function of T is not necessarily sufficient for θ.

29 If h is an unbiased estimator of θ, then using Rao-Blackwell theorem, we can always obtain an unbiased estimator of θ whose variance is.....

A) the smallest among the variances of all unbiased estimators of

B) ≤ variance of h

C) = variance of h

D) < variance of h

30 Neyman-Pearson lemma provides .....

A) a most powerful test B) an unbiased test

C) an admissible test D) A uniformly most powerful test

31 In GLMs, parameters are estimated using the method of.....

A) ordinary least squares B) maximum likelihood estimation

C) minimum chi-square D) None of A, B, C

32 For an $\overline{X}$ chart with usual assumptions, the probability of type I error.....

A) is less than the probability of type II error

B) is greater than the probability of type II error

C) is less than (1- the probability of type II error)

D) depends on the sample size

33 Which of the following is NOT true?

A) The use of 30 controls limits in a Shewhart chart is based on the assumption of normality of the control statistic.

B) CUSUM charts cannot be used to monitor dispersion.

C) EWMA charts are more efficient than Shewhart charts in detecting small shifts in process parameters.

D) Use of warning limits in a control chart increases the probability of type 1 error.

34 Which of the following is true?

A) In an finite Markov chain all states are of the same type

B) In an finite reducible Markov chain all states are of the same type

C) In an finite irreducible Markov chain all states are recurrent

D) In an finite Markov chain at least one state has to be transient

35 The output of the following R-command is.....

a = matrix (1:9, nrow-3); sum (a [1,])

A) 18

B) 12

C) 6

D) 1

36 Which of the following statement(s) is/are true about the linear Congruential random number generator $X_{n+1} = a * X_n + b \ (mod\ m)\ for\ n = 0, 1, 2, \ldots$?

i) Maximum possible cycle length is achieved if we select b as any odd number.

ii) If we choose m = 32 a = 21 and b = 17 then maximum cycle length will be achieved.

A) Both i) and ii) are true

B) Only i) is true

C) Only ii) is true

D) Both i) and ii) are not true

37 In PCA, how would you interpret the coefficients of the original features in the principal components?

A) They represent the original feature values

B) They indicate the order of importance of the features

C) They show the percentage of variance each feature contributes

D) They represent the contribution of each feature to the principal component

38 Which of the following statement is always true?

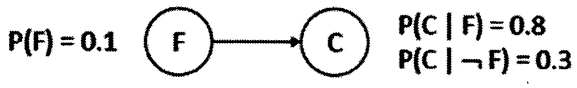
A) systematic sample is superior than stratified random sample

B) simple random sample is inferior than systematic sample

C) stratified random sample is better than systematic sample

D) none of the above

39 Consider the following Bayesian Belief network, where F = having the flue and C = coughing:



where ¬ denotes the negation. Then $P(\neg F, \neg C) = \cdots$

A) 0.08 B) 0.02

C) 0.27 D) 0.63

40 Let 6, 3.5, 6.7, 2, 8 be the observed failures of five units put on study. Based on this complete data, nonparametric estimate of F (7) (distribution function) and S (5) (survival function) are respectively given by.....

A) 0.6, 0.2 B) 0.4, 0.8

C) 0.8, 0.6 D) 0.4, 0.4

41 Diseases that are always present in a community, usually at a low, more or less constant, frequency are classified as having an ..... pattern.

A) epidemic B) endemic

C) pandemic D) Synergistic

42 A set is open if every point of the set is it's .....point.

A) Limit

B) Interior

C) Adjacent

D) Member

43 To find constrained maxima of a function, ... method can be used.

A) Lagrange's B) Leibnitz

C) Bolzano D) Borel

44 Which of the following statements is true regarding a probability measure?

A) A probability measure assigns probabilities to only disjoint events.

B) The probability of the entire sample space is always zero.

C) The probability measure of the empty set is always one.

D) A probability measure is always defined on a non-empty sample space.

45 Regarding the Simplex Method in linear programming, identify the correct statement:

A) The Simplex Method is only applicable to linear programming problems with two decision variables..

B) In each iteration of the Simplex Method, a non-basic variable enters the basis, and a basic variable leaves the basis.

C) The Simplex Method is guaranteed to find the global optimal solution in a finite number of iterations.

D) The Simplex Method is primarily used for solving integer linear programming problems.

46 If $X \sim N(\mu, \sigma^2)$ then distribution of $Y = e^X$ is...

A) Laplace

B) Log normal

C) Exponential

D) Gamma

47 For a matrix N with 5 rows and 3 columns, p(N) is rank of N then .....

A) $\rho(N) \leq 5$

B) $\rho(N) \geq 3$

C) $\rho(N) \geq 5$

D) $\rho(N) \leq 3$

48 The column space of a non-singular matrix A of order 3 has dimension.....

A) is equal to 3.

B) greater than 3.

C) less than 3

D) 3^3

49 What does a unit root imply about a time series?

A) It has no autocorrelation

B) The data is perfectly predictable

C) It is a non-stationary time series

D) It has a stationary trend

50 The ..... test is used to test significance of individual regression coefficient in multiple linear regression

A) x^2 B) t

C) F D) Z

◆◆◆

-- ROUGH WORK --