

D 102525

(Pages : 12)

Name.....

Reg. No.....

P.G. ENTRANCE EXAMINATION, APRIL 2024

STATISTICS

Time : Two Hours

Maximum : 300 Marks

Answer all the questions.

Each question carries 4 marks.

1 mark will be deducted for each wrong answer.

1. What is the supremum of the set $A = \{x \in \mathbb{R} : x^2 < 2\}$?
(A) 2. (B) $\sqrt{2}$.
(C) 0. (D) There is no supremum.
2. Which of the following sets is uncountable ?
(A) Integers. (B) Rational numbers.
(C) Algebraic numbers. (D) Real numbers.
3. Which convergence test would be suitable for determining the convergence of the series $\sum_{n=1}^{\infty} \frac{1}{n^3}$?
(A) Ratio test. (B) Comparison test.
(C) Root test. (D) Integral test.
4. Which of the following functions is not continuous at $x = 0$?
(A) $f(x) = 1/x$. (B) $f(x) = |x|$.
(C) $f(x) = \sqrt{x}$. (D) $f(x) = \sin x / x$.
5. The Maclaurin series expansion of e^x is :
(A) $1 + x + x^2/2 + x^3/3 + \dots$ (B) $1 + x + x^2/2! + x^3/3! + \dots$
(C) $1 - x + x^2/2 - x^3/3 + \dots$ (D) $1 - x + x^2/2! - x^3/3! + \dots$
6. Which theorem guarantees the existence of at least one point c in the interval (a, b) such that $f(c) = 0$?
(A) Rolle's theorem. (B) Mean value theorem.
(C) Taylor's theorem. (D) Lagrange's theorem.

Turn over

7. Evaluate the integral $\int_0^1 \frac{\sin x}{x} dx$:

(A) $\frac{\pi}{2}$.

(B) π .

(C) $\frac{\pi}{4}$.

(D) $\frac{\pi}{6}$.

8. Which property is NOT true for symmetric matrices ?

(A) They are always square matrices.

(B) Their eigenvalues are real.

(C) They are always invertible.

(D) They are equal to their transpose.

9. The rank of the matrix $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ is :

(A) 0.

(B) 1.

(C) 2.

(D) 3.

10. Which of the following statements about the solution of a system of linear equation is true ?

(A) A system with fewer equations than variables has a unique solution.

(B) A system with more equations than variables has a unique solution.

(C) A system with more equations than variables can have no solution.

(D) A system with fewer equations than variables always has infinite solutions.

11. What does it indicate about the distribution if the mode, mean and median of a dataset are all equal ?

(A) The distribution is skewed to the left.

(B) The distribution is skewed to the right.

(C) The distribution is symmetric.

(D) The distribution is bimodal.

12. In a dataset, the mean is greater than the median. What does this suggest about the distribution ?

(A) The distributon is symmetric.

(B) The distribution is skewed to the right.

(C) The distribution is skewed to the left.

(D) The distributon is uniform.

13. What type of mean is typically used when averaging rates, such as speed or fuel efficiency ?
- (A) Arithmetic mean. (B) Geometric mean.
(C) Harmonic mean. (D) Median.
14. If the co-efficient of skewness of a dataset is negative, what can you infer about its shape ?
- (A) The left tail is longer than the right tail.
(B) The right tail is longer than the left tail.
(C) The distribution is symmetric.
(D) The distribution is bimodal.
15. The co-efficient of kurtosis measures :
- (A) The peakedness of a distribution.
(B) The symmetric of a distribution.
(C) The Spread of a distribution.
(D) The position of a distribution's mode.
16. What type of skewness does a distribution exhibit if its mean is greater than its median ?
- (A) Positive skewness. (B) Negative skewness.
(C) Symmetrical skewness. (D) No skewness.
17. Which of the following measures of dispersion is resistant to outliers ?
- (A) Range. (B) Quartile deviation.
(C) Standard deviation. (D) Co-efficient of variation.
18. The interquartile range (IQR) covers what percentage of the data in a dataset ?
- (A) 25 %. (B) 50 %.
(C) 75 %. (D) 100 %.
19. If the co-efficient of variation of two datasets is the same, what can you infer about their dispersion ?
- (A) They have the same standard deviation.
(B) They have the same mean.
(C) They have the same range.
(D) They have the same relative variability.

20. In a dataset with a largen number of outliers, which measure of central tendency is likely to provide the most representative estimate of the center of the data ?
- (A) Trimmed mean. (B) Weighted mean.
(C) Winsorized mean. (D) Median Absolute Deviation (MAD).
21. What does the Karl Pearson correlation co-efficient measure ?
- (A) The strength and direction of linear relationship between two variables.
(B) The strength of association regardless of linearity.
(C) The degree of non-linear relationship between two variables.
(D) The rank correlation between two variables.
22. A company collects data on the number of advertisements and the corresponding sales revenue over several months. The scatterplot shows a linear relationship between the two variables. What would be an appropriate method to model this relationship ?
- (A) Spearman correlation co-efficient.
(B) Least squares regression.
(C) Fisher's ideal index numbers.
(D) Weighted index numbers.
23. In the context of time series analysis, what is the purpose of estimating components ?
- (A) To identify outliers in the data.
(B) To determine the strength of the linear relationship between variables.
(C) To decompose the time series into trend, seasonality and residuals.
(D) To calculate the co-efficient of determination.
24. A researcher wants to study the relationship between temperature and ice cream sales over a year. Which time series component is likely to represent the regular, predictable fluctuations in ice cream sales ?
- (A) Trend. (B) Seasonality.
(C) Residuals. (D) Cyclical variation.
25. What does the co-efficient of determination (R-squared) represent in the context of linear regression ?
- (A) The proportion of variation in the dependent variable explained by the independent variable.
(B) The strength of the linear relationship between the variables.
(C) The degree of association between the variables.
(D) The predictive power of the regression model.

26. In a linear regression model, if the co-efficient of determination (R-squared) is 0.75, what percentage of the variation in the dependent variable is explained by the independent variable ?
- (A) 75 %. (B) 25 %.
(C) 100 %. (D) 50 %.
27. What is the purpose of using index numbers in statistical analysis ?
- (A) The measure the strength of association between variables.
(B) To decompose time series data into components.
(C) To compare changes in variables over time or between different groups.
(D) To estimate the co-efficients of a regression equation.
28. A price index for consumer goods increased from 120 to 125 over a year. If the index is composed of three components-food, housing and transportation-what can be inferred about the change in these components ?
- (A) All components increaed proportionally.
(B) Only one component increased significantly.
(C) There was a mix of increases and decreases across components.
(D) There was no change in the components.
29. When conducting a hypothesis test on index numbers, what is the null hypothesis typically ?
- (A) There is a significant difference between index numbers.
(B) There is no significant difference between index numbers.
(C) The index numbers are normally distributed.
(D) The index numbers have outliers.
30. If the p -value of a correlation test between two variables is 0.05, what can be concluded about the significance of the correlation coefficient ?
- (A) The correlation co-efficient is statistically significant.
(B) The correlation co-efficient is not statistically significant.
(C) The correlation co-efficient is moderate.
(D) The correlation co-efficient is strong.
31. What is the probability of rolling an odd number on a fair six-sided die ?
- (A) $1/6$. (B) $1/3$.
(C) $1/2$. (D) $2/3$.

32. If the probability of event A is 0.4 and the probability of event B is 0.3, what is the probability of either A or B occurring ?
- (A) 0.4. (B) 0.3.
(C) 0.7. (D) 0.2.
33. What is the definition of conditional probability $P(A|B)$?
- (A) $P(A \text{ and } B)/P(A)$. (B) $P(A \text{ and } B)/P(B)$.
(C) $P(A) * P(B)$. (D) $P(A) + P(B)$.
34. In a bag, there are 4 red balls and 6 blue balls. If one ball is drawn at random, what is the probability of getting a red ball ?
- (A) 0.4. (B) 0.6.
(C) 0.5. (D) 0.2.
35. What is Bayes' theorem used for ?
- (A) To calculate the probability of an event occurring given prior knowledge.
(B) To calculate the probability of two independent events occurring simultaneously.
(C) To calculate the probability of mutually exclusive events.
(D) To calculate the probability of complementary events.
36. If X and Y are independent random variables with probability density functions $f(x)$ and $g(y)$ respectively, what is the joint probability distribution of X and Y ?
- (A) $f(x) * g(y)$. (B) $f(x) + g(y)$.
(C) $f(x) / g(y)$. (D) $f(x) - g(y)$.
37. In statistical hypothesis testing, what does the p -value represent ?
- (A) The probability of making a Type I error.
(B) The probability of making a Type II error.
(C) The probability of observing the test statistic or a more extreme value, given that the null hypothesis is true.
(D) The probability of correctly rejecting the null hypothesis.
38. If the joint probability distribution of two discrete random variables X and Y is given by $P(X = 1, Y = 2) = 0.3$, what is the probability of X being 1 and Y being 2 ?
- (A) 0.3. (B) 0.2.
(C) 0.1. (D) 0.6.

39. What does it mean for two random variables to be independent ?
- (A) They have the same probability distribution.
 - (B) The value of one variable does not affect the value of the other.
 - (C) They have a joint probability distribution.
 - (D) They have a conditional probability distribution.
40. If the marginal probability distribution of X is $P(X = 1) = 0.4$ and $P(X = 2) = 0.6$, what is the probability of X being 1 given that Y is 2 ?
- (A) 0.4.
 - (B) 0.6.
 - (C) 0.3.
 - (D) 0.2.
41. Which of the following distributions is characterized by a continuous probability density function with two parameters, α and β , and is commonly used in Bayesian statistics and reliability engineering ?
- (A) Beta distribution.
 - (B) Gamma distribution.
 - (C) Exponential distribution.
 - (D) Normal distribution.
42. The moment generating function (MGF) of a random variable X is given by $M(t) = \exp(2t + t^2)$. What is the second moment of X ?
- (A) 2.
 - (B) 4.
 - (C) 6.
 - (D) 8.
43. A random variable X follows a geometric distribution with probability mass function $P(X = k) = (1/2)^{(k+1)}$, $k = 0, 1, 2, \dots$. What is the expected value of X ?
- (A) 1.
 - (B) 2.
 - (C) 3.
 - (D) 4.
44. The sampling distribution that arises from the ratio of two independent Chi-square distributed random variables is known as :
- (A) Chi-square distribution.
 - (B) Student's t -distribution.
 - (C) F-distribution.
 - (D) Normal distribution.
45. In a binomial distribution, if n is the number of trials and p is the probability of success on each trial, then the variance of the distribution is given by :
- (A) np .
 - (B) $np(1 - p)$.
 - (C) n^2p .
 - (D) $n(1 - p)$.
46. Which of the following distributions is not symmetric and has a long right tail ?
- (A) Normal distribution.
 - (B) Lognormal distribution.
 - (C) Exponential distribution.
 - (D) Beta distribution.

47. Which of the following distributions is characterized by a constant probability of success, p , and a constant probability of failure, $q = 1 - p$, over a fixed number of independent trials ?
- (A) Normal distribution. (B) Exponential distribution.
(C) Binomial distribution. (D) Poisson distribution.
48. The distribution of the number of successes in a fixed number of independent Bernoulli trials with constant probability of success is described by the :
- (A) Poisson distribution. (B) Hypergeometric distribution.
(C) Binomial distribution. (D) Exponential distribution.
49. Suppose a random variable X follows a binomial distribution with parameters $n = 10$ and $p = 0.4$. What is the probability that X takes on an even value ?
- (A) 0.37. (B) 0.42.
(C) 0.55. (D) 0.63.
50. Which distribution is used to model the waiting time between independent Poisson events ?
- (A) Exponential distribution. (B) Uniform distribution.
(C) Beta distribution. (D) Gamma distribution.
51. Which property of an estimator is crucial for ensuring that the estimator's expected value converges to the true value of the parameter as the sample size increases ?
- (A) Consistency. (B) Unbiasedness.
(C) Efficiency. (D) Robustness.
52. The Cramer-Rao lower bound establishes a limit on the variance of an unbiased estimator based on :
- (A) The sample size.
(B) The parameter being estimated.
(C) The information provided by the sample.
(D) The population distribution.
53. According to the Neymann factorization theorem, a statistic T is sufficient for a parameter θ if and only if :
- (A) The likelihood function is maximized.
(B) The joint distribution of the sample can be factorized into two functions.
(C) The estimator is unbiased.
(D) The estimator is consistent.

54. The method of moments estimation :
- (A) Involves maximizing the likelihood function.
 - (B) Minimizes the sum of squared errors.
 - (C) Is based on equating sample moments to population moments.
 - (D) Requires knowledge of the sample size.
55. In interval estimation, a confidence interval :
- (A) Guarantees that the true parameter lies within the interval with a certain probability.
 - (B) Provides an exact estimate of the population parameter.
 - (C) Is always symmetric around the sample mean.
 - (D) Is narrower for larger sample size.
56. Type I error refers to :
- (A) Rejecting the null hypothesis when it is true.
 - (B) Failing to reject the null hypothesis when it is false.
 - (C) Rejecting the alternative hypothesis when it is true.
 - (D) Failing to reject the alternative hypothesis when it is false.
57. The power of a hypothesis test is influenced by :
- (A) The level of significance.
 - (B) The sample size.
 - (C) The variability of the data.
 - (D) The null hypothesis.
58. The critical region of a hypothesis test consists of :
- (A) All possible outcomes of the sample.
 - (B) Outcomes that lead to acceptance of the null hypothesis.
 - (C) Outcomes that lead to rejection of the null hypothesis.
 - (D) Outcomes that lead to rejection of the alternative hypothesis.
59. The Chi-square distribution is commonly used in hypothesis testing for :
- (A) Comparing means of two independent samples.
 - (B) Testing the equality of variances in two samples.
 - (C) Comparing proportions in two independent samples.
 - (D) Testing the goodness of fit of a sample distribution to a theoretical distribution.

60. The Kolmogorov-Smirnov test is a nonparametric test used for :
- (A) Comparing means of two independent samples.
 - (B) Testing the equality of variances in two samples.
 - (C) Testing the goodness of fit of a sample distribution to a theoretical distribution.
 - (D) Testing the normality of a sample distribution.
61. Which of the following sampling methods ensures that every subgroup within the population is represented in the sample ?
- (A) Stratified random sampling.
 - (B) Systematic sampling.
 - (C) Probability sampling.
 - (D) Nonprobability sampling.
62. What is the primary advantage of stratified random sampling over simple random sampling ?
- (A) It is easier to implement.
 - (B) It allows for greater precision in estimation.
 - (C) It requires a smaller sample size.
 - (D) It reduces non-sampling errors.
63. A researcher divides a population into several homogeneous groups and then randomly selects samples from each group. What sampling method is being used ?
- (A) Cluster sampling.
 - (B) Systematic sampling.
 - (C) Stratified random sampling.
 - (D) Convenience sampling.
64. In a systematic sampling method, if a researcher wants to sample every 5th item from a population of 1000, what would be the sampling interval ?
- (A) 5.
 - (B) 10.
 - (C) 20.
 - (D) 200.
65. Which of the following is an example of a nonprobability sampling method ?
- (A) Simple random sampling.
 - (B) Stratified random sampling.
 - (C) Convenience sampling.
 - (D) Cluster sampling.
66. What is the purpose of estimating the standard error of the population mean in sampling ?
- (A) To measure the variability of the sample mean.
 - (B) To assess the accuracy of the sample mean as an estimator of the population mean.
 - (C) To determine the margin of error in the sample mean.
 - (D) To calculate confidence intervals for the population mean.

67. Which of the following experimental designs involves only one independent variable with multiple levels ?
- (A) One-way ANOVA.
 - (B) Two-way ANOVA.
 - (C) Randomized Complete Block Design (RCBD).
 - (D) Latin Square Design (LSD).
68. Crude birth rate is calculated as the number of births per ——— people in a given population.
- (A) 100.
 - (B) 1,000.
 - (C) 10,000.
 - (D) 100,000.
69. What is the primary purpose of a national income statistics ?
- (A) To measure the distribution of wealth in a country.
 - (B) To track changes in the Gross Domestic Product (GDP) over time.
 - (C) To estimate the total income earned by individuals and businesses within a country.
 - (D) To assess the level of income inequality within a country.
70. Which organization is responsible for conducting the population census in India ?
- (A) Central Statistical Office (CSO).
 - (B) National Sample Survey Office (NSSO).
 - (C) Ministry of Home Affairs.
 - (D) Registrar General and Census Commissioner of India.
71. In statistical quality control, what does an "X-bar chart" primarily monitor ?
- (A) Individual measurements of a process.
 - (B) Range of measurements in a process.
 - (C) Mean of a sample taken from the process.
 - (D) Variability within a sample from the process.
72. Which type of control chart is used to monitor the dispersion or variability within a sample ?
- (A) X-bar chart.
 - (B) R-chart.
 - (C) S-chart.
 - (D) P-chart.

73. Acceptance sampling is primarily used for :
- (A) Monitoring ongoing processes.
 - (B) Rejecting defective products before they reach the customer.
 - (C) Adjusting process parameters to maintain quality.
 - (D) Estimating process capability.
74. The simplex method is used to solve which type of problem in Operations Research ?
- (A) Linear programming problem.
 - (B) Transportation problem.
 - (C) Assignment problem.
 - (D) Integer programming problem.
75. In transportation problems, the objective is to minimize ?
- (A) The total cost of transportation.
 - (B) The total distance travelled.
 - (C) The number of vehicles used.
 - (D) The time taken for transportation.

(75 × 4 = 300 marks)