Seat No.

PD-24

Total No. of Pages : 13

Shivaji University, Kolhapur
Ph.D. Entrance Examination, February- 2024
STATISTICS
Sub. Code : 58794

Day and Date : Tuesday, 06-02-2024 **Total Marks : 100**

Time : 04:00 pm to 06:00 pm

Instructions:

1) All questions are compulsory.
2) Each question carries 2 marks.
3) Answers should be marked in the given OMR answer sheet by darkening the appropriate option.
4) Use black ball point pen only for marking the circle. Do not make any stray mark on the OMR Answer Sheet.
5) Follow the instructions given on OMR Sheet.
6) Rough work shall be done on the sheet provided at the end of question paper.
7) Only non programmable calculators are allowed.

---

1. Which of the following is /are the objectives of research?

   i) To become familiar with phenomenon

   ii) To test a hypothesis of a causal relationship between variables

   iii) To determine the frequency with which something occurs or whether it is associated with something else.

   A) Only i)

   B) Only ii)

   C) Only i) and iii)

   D) All i), ii) and iii)

[P.T.O...]

2. What is the primary goal of applied research?

A) To test existing theories

B) To advance scientific knowledge

C) To formulate new hypotheses

D) To solve practical problems or address real-world issues

3. A Hypothesis is a.....

A) Tentative statement whose validity is still to be tested

B) Conclusion of the research

C) Statement of facts

D) Statement which researcher wants to prove,

4. Which of the following is NOT a common type of quantitative research design?

A) Correlational B) Experimental

C) Quasi-experimental D) Case study

5 What is the name of the conceptual framework in which the research is carried out?

A) Research hypothesis

B) Synopsis of Research

C) Research paradigm

D) Research design

6 Which of the following is NOT the tool for searching research topics?

A) Books B) Internet

C) Journals D) Microsoft Word

7. Plagiarism in research refers to:

A) Exaggerating research findings

B) Falsifying data.

C) Conducting research without ethical approval

D) Using someone else's work without proper credit

8. What is the significance of the "peer review" process in research ethics?

A) To ensure that only positive results are published

B) To provide a platform for researchers to promote their work

C) To evaluate the quality and validity of research before publication

D) To exclude diverse perspectives from the research community

9. Which of the following is an example of research misconduct?

A) Openly discussing preliminary findings at a conference

B) Collaborating with other researchers to enhance the study's quality

C) Publishing negative results to contribute to scientific knowledge

D) Altering research data to support desired outcomes

10. Informed consent in research involves:

A) Keeping research participants unaware of the study's purpose

B) Obtaining voluntary agreement from participants after providing information about the study

C) Forcing participants to participate without their consent

D) Only seeking consent from certain demographic groups

11. What is one of the key qualities of a good researcher?

A) Speed in completing research projects

B) Ability to conduct experiments without proper planning

C) Open-mindedness and a willingness to consider different perspectives

D) Conducting research by ignoring ethical considerations for the sake of quick results

12 . Find the odd man out: 3, 5, 7, 12, 17, 19

A) 3 B) 7

C) 12 D) 17

13. Which of the following will not hold true in case of a skew symmetric matrix A of even order whose all elements are integers?

A) Det (A) = 9 B) Det (A) = 81

C) Det (A) = 7 D) Det (A) = 4

14. Which of the following statements is true?

A) Bayes estimator is always a function of minimal sufficient statistic

B) Bayes estimator are most efficient

C) Bayes estimator are always asymptotically normal

D) None of the above

15. The equation $x^3 - x^2 + 4x - 4 = 0$ is to be solved using the Newton-Raphson method. If x = 2 is taken as the initial approximation of the solution, then the next approximation using this method will be _____

A) 2/3 B) 4/3

C) 1 D) 3/2

16. In a goodness-of-fit test, the null hypothesis typically assumes:

A) The observed data fits the expected distribution

B) The observed data does not fit the expected distribution

C) The means of two samples are equal

D) The variance of the sample is known

17. Assuming an initial bracket of [0, 3], the second (at the end of 2 iterations) iterative value of the root of $x^2 + x - 2 = 0$ using the bisection method is

A) 1 B) 0.75

C) 2.25 D) 1.5

18. Consider the following statements,

(i) Newton Raphson method for approximating root of nonlinear equation always converges to the real root.

(ii) Gauss Seidel method for approximating the solution of system of linear equations always converges to true solution.

A) Both statements (i) and (ii) are true

B) Only statement (i) is true

C) Only statement (ii) is true

D) Both statements (i) and (ii) are false

19. Which of the following factors will increase the width of a confidence interval?

A) Decreasing the sample size

B) Increasing the level of significance.

C) Reducing the variability of the data

D) Increasing the sample size

20. _____ data analysis approach allows the data to suggest admissible models that fit the data.

A) Exploratory B) Classical

C) Bayesian D) None of the above

21. Let X ~ B (10, 0.4). Which of the following MSEXCEL function returns the value of P(X > 7)?

A) BINOMDIST (7, 10, 0.4, 1)

B) 1-BINOMDIST (6, 10, 0.4, 1)

C) 1-BINOMDIST (7, 10, 0.4, 1)

D) BINOMDIST (10, 7, 0.4, 1)

22. Suppose data of 100 students on the variable "Roll_Number", "GENDER" (Male/Female), "TY_Marks" is stored in EXCEL in column A, B and C respectively. First row contains the variable names. Which of the following function in EXCEL will provide the average TY Marks of Male candidates?

A) = AVERAGEIFS(B2:B101, C2:C101, "=Male")

B) = AVERAGEIFS(B2:B101, "=Male", C2:C101)

C) = AVERAGEIF(B2:B101, "=Male", C2:C101)

D) = AVERAGEIF(C2:C101, "=Male", B2:B101)

23. Let $U_1, U_2, \ldots\ldots\ldots\ldots, U_{10}$ be i.i.d. U (0, 1) random variables

Define $X_i = \begin{cases} 1 & if\ U_i \leq \frac{30-\sum_{j=0}^{i-1} X_j}{51-i} \\ 0 & otherwise \end{cases}$, $i = 1,2,\ldots,10$, $X_0 = 0$ and $X = \sum_{i=1}^{10} X_i$ then distribution of X is .....

A) H (N = 50, M = 30, n = 10)

B) H (N = 51, M = 30, n = 10)

C) H (N = 51, M = 20, n = 10)

D) H (N = 50, M = 20, n = 10)

24. Let $X_1, X_2,..., X_{20}$ be a random sample of size 20 from N (θ, 1). Let M be the sample median. Proposed interval estimate for θ is (M-3, M+3). In order to estimate the coverage probability (by simulation) of the proposed interval estimate when true value of θ is 5,.....

A) we generate 1000 sets of 5 observations from N (20, 1), for each sample compute the sample median M, and compute the proportion of times the interval (M-3, M+3) covers 5

B) we generate 1000 sets of 20 observations from N (3, 1), for each sample compute the sample median M, and compute the proportion of times the interval (M-3, M+3) covers 5

C) we generate 1000 sets of 20 observations from N (5, 1), for each sample compute the sample median M, and compute the proportion of times the interval (M-3, M+3) covers 3

D) we generate 1000 sets of 20 observations from N (5, 1), for each sample compute the sample median M, and compute the proportion of times the interval (M-3, M+3) covers 5

25 Which of the following statement(s) is/are true about Bootstrap estimate of standard error?

i) It does not require knowledge of the theoretical form of an estimator's standard error.

ii) it does not require knowledge of the parent distribution.

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

26. The totals of the response observations of the three replicates of a 2^2 factorial experiment with factors A and B are, namely, (1)=80, a=100, b =60, and ab = 90. Then, the main effect B is....

A) 8.33

B) -5

C) 1.67

D) -15

27. Let q_1 be the sample first quartile of a random sample from exponential distribution with mean θ. The consistent estimator of θ based on q_1 using the method of percentile is .....

A) $\frac{q_1}{\log(\frac{1}{4})}$

B) $\frac{q_1}{\log(\frac{3}{4})}$

C) $\frac{q_1}{\log(\frac{4}{3})}$

D) $\frac{q_1}{\log(4)}$

28. Suppose, $X_1, X_2, ..., X_n$ is random sample from P(λ) distribution. Then, ...

A) $\sum_{i=1}^{n} X_i$ is not sufficient for λ.

B) $\bar{X}$ is sufficient for λ.

C) $\prod_{i=1}^{n} X_i$ sufficient for λ.

D) sufficient statistics for λ does not exist.

29. Suppose, T is a sufficient statistic for θ, and h is a function of T such that $E_\theta h = \theta$ and $E_\theta h^2 < \infty$ for all $\theta \in \Theta$. Then, ...

A) $Var_\theta(h|T) \leq Var_\theta(h)$

B) $Var_\theta(h|T) = Var_\theta(h)$

C) $Var_\theta(h|T) \geq Var_\theta(h)$

D) None of above

30. In hypothesis testing, what does the p-value represent?

A) The probability of rejecting the null hypothesis when it is true

B) The probability of accepting the null hypothesis when it is true

C) The probability of obtaining a value of the test statistic at least as extreme as the observed one, given that the null hypothesis is true

D) The probability of committing a Type II error

31. In a GLM, the distribution of a response variable .....

A) is a member of Pitman family

B) is a member of exponential dispersion family

C) is continuous distribution

D) is such that its range depends on its parameter

32. Consider following statements.

I. In product control, the desired quality is achieved through detection method.

II. In process control, no attempt is made to find the causes of defective production

A) Both ate true

B) Both are false

C) (I) is true and (II) is false

D) (I) is false and (II) is true

33. The process capability index C_{pk} coincides with the index C_p when....

A) the process is normally distributed.

B) the process mean is located at the lower or the upper specification limit.

C) the process mean is located at the midpoint of the specification range.

D) the process standard deviation is 1.

34. Let $\{X_n, n \geq 0\}$ be a markov chain with state space S = {0, 1, 2}, t.p.m

$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.1 & 0.8 & 0.1 \end{bmatrix}$ and (0,1,0) as initial distribution.

Then $P(X_0 = 0, X_1 = 1, X_2 = 2, X_3 = 0)$ is.....

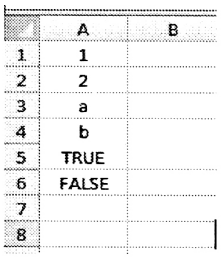
A) 1

B) 0

C) 0.1

D) 0.8

35. 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 "= SUM(A1:A6)"

A) 3 B) 4

C) 6 D) #N/A

36. Suppose you have provided with 50 green and 50 red chits of same size and empty box. From which of the following distributions, you can generate the true random numbers using those 100 chits?

i) G(1/2)

ii) NB (2, 0.15)

iii) NB (5, 1/3)

A) Only i) B) Only i) and ii)

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

37. If X is a multivariate normal random vector with mean vector μ, and Y= AX + b, where A is matrix and b is column vector, what can be said about the distribution of Y?

A) Y is not normally distributed

B) Y is normally distributed with the same mean vector as X

C) Y is normally distributed with a mean vector equal to Aμ + b

D) Y is normally distributed only if A is an identity matrix

38. Which of the following is NOT true about stratified random sampling?

A) It involves a random selection process from identified subgroups

B) Sample proportion of each stratum must always match to corresponding population proportion

C) Sample units are selected from all strata

D) Proportional stratified random sampling yields a representative sample

39. Consider the following data. Suppose we wish to predict the value of the variable Y given the values of the X1 and X2.

X1	X2	Y
150	50	F
110	10	M
140	40	F
130	30	M
120	20	M
160	60	F

Then, predicted label of the new observation (X1=139, X2=39) using k-NN algorithm with k=3 (with Euclidian distance) is with probability

A) F, 1 B) F, 2/3

C) M, 1 D) M, 2/3

40. Life time of component has edf *F()*. Let n identical components were put on test and experiment conducted with type-I censoring scheme with t_0 as a censoring time. If R denotes the number of uncensored observation in type-1 censoring scheme then distribution of R is .....

A) $B(n, F(t_0))$ B) $B(n, 1\text{-}F(t_0))$

C) $G(F(t_0))$ D) $NB(n, 1\text{-}F(t_0))$

41. Which design is appropriate to compare cholesterol levels in a single population of daily walkers over an extended period of time?

A) Cross-sectional study B) Longitudinal Design

C) Crossover design D) None of the above

42. A set is compact if and only if it is closed and.....

A) Finite B) Uncountable

C) Bounded D) Unbounded

43. A convergent sequence have.....

A) Only one limit B) At most two limits

C) At most n limits D) Infinite limits.

44. Which of the following statements is true regarding measurable functions?

A) Every continuous function is measurable

B) A measurable function is always differentiable

C) The indicator function of a measurable set is measurable

D) Measurable functions are always Riemann integrable

45. Regarding methods of solving Linear Programming Problems (LPP), identify the correct statement:

A) The Simplex method can only be applied to lincar programming problems with two variables

B) The graphical method is suitable for solving LPPs with two decision variables

C) The interior-point method is primarily used for discrete optimization problems

D) Sensitivity analysis is used to determine the feasible region of an LPP

46. Which of the following is not a property of distribution function?

A) $\lim_{x\to-\infty} F_X(x) = 0$

B) $\lim_{x\to\infty} F_X(x) = 1$

C) $\lim_{h\to0^+} F_X(x+h) = F_X(x)\ \forall\ x$

D) $F_X(x)$ is an increasing function

47. Null space.....

A) has dimension zero

B) contains no vector

C) is a vector space

D) contains only one vector (0,0,....., 0)

48. If transpose of the given matrix is equal to the matrix itself, then it is called.....

A) Orthogonal matrix B) Symmetric matrix

C) Scalar matrix D) Identity matrix

49. In an AR(p) model, what does 'p' represent?

A) Lag order

B) Moving average order

C) Coefficient of the autoregressive terms

D) Number of observations

50. $MSS_{residual} = \frac{SS_{residual}}{n-p}$ is _____ , where p is the number of parameters

A) an unbiased estimator of error variance σ^2

B) a biased estimator of error variance σ^2

C) an unbiased estimator of error variance σ^2/n

D) an unbiased estimator of error mean

◆◆◆