PUMDET-2024

Subject : Statistics

4032500461

(Booklet Number)

Duration : **90** Minutes | No. of Questions : **50** | Full Marks : **100**

INSTRUCTIONS

1. All questions are of objective type having four answer options for each. Only one option is correct. Correct answer will carry full marks 2. In case of incorrect answer or any combination of more than one answer, ½ mark will be deducted.
2. Questions must be answered on OMR sheet by darkening the appropriate bubble marked A, B, C or D.
3. Use only **Black/Blue ink ball point pen** to mark the answer by complete filling up of the respective bubbles.
4. Mark the answers only in the space provided. Do not make any stray mark on the OMR Sheet.
5. Write question booklet number and your roll number carefully in the specified locations of the **OMR Sheet**. Also fill appropriate bubbles.
6. Write your name (in block letter), name of the examination centre and put your signature (as is appeared in Admit Card) in appropriate boxes in the OMR Sheet.
7. The OMR Sheet is liable to become invalid if there is any mistake in filling the correct bubbles for question booklet number/roll number or if there is any discrepancy in the name/signature of the candidate, name of the examination centre. The OMR Sheet may also become invalid due to folding or putting stray marks on it or any damage to it. The consequence of such invalidation due to incorrect marking or careless handling by the candidate will be sole responsibility of candidate.
8. Candidates are not allowed to carry any written or printed material, calculator, pen, docu-pen, log table, wristwatch, any communication device like mobile phones, bluetooth devices etc. inside the examination hall. Any candidate found with such prohibited items will be **reported against** and his/her candidature will be summarily cancelled.
9. Rough work must be done on the question booklet itself. Additional blank pages are given in the question booklet for rough work.
10. Hand over the OMR Sheet to the invigilator before leaving the Examination Hall.
11. Candidates are allowed to take the Question Booklet after examination is over.

Signature of the Candidate : ____________________
(as in Admit Card)

Signature of the Invigilator : ____________________

SPACE FOR ROUGH WORK

Statistics

1. Let $S = \{(-1)^m + \frac{1}{n} : m \in \mathbb{N}, n \in \mathbb{N}\}$, then the set of limit points of S

(A) is empty (B) has exactly one element

(C) has exactly two elements (D) has infinitely many elements

2. The graphs of the function $f(x) = \cos x$ and the polynomial $g(x) = x^2 - 1$ intersect at

(A) exactly two points.

(B) exactly three points.

(C) exactly four points.

(D) at least four but finitely many points.

3. Let A be a matrix of order 10×20. If nullity $(A^T) = 1$, then the nullity of A is

(A) 10 (B) 20

(C) 1 (D) 11

4. Let A be a 5×5 diagonal matrix such that $|A - \lambda I| + (\lambda - 2)^3 (\lambda + 2)^2 = 0$. The values of α and β such that $A^{-1} = \alpha A + \beta I$ are

(A) $\alpha = 0, \beta = \frac{1}{4}$ (B) $\alpha = 2, \beta = -2$

(C) $\alpha = \frac{1}{4}, \beta = 0$ (D) $\alpha = -2, \beta = 2$

5. If $A = \begin{pmatrix} 1 & 0 & 0 \\ i & \frac{-1+\sqrt{3}\,i}{2} & 0 \\ 0 & 1+2i & \frac{-1-\sqrt{3}\,i}{2} \end{pmatrix}$, then the trace of A^{102} is equal to

(A) zero (B) 1

(C) 2 (D) 3

6. If it is known that a random variable X has mean 25 and variance 16, then the largest possible lower bound for $P(17 < X < 33)$ will be

(A) $\frac{7}{8}$ (B) $\frac{1}{4}$

(C) $\frac{5}{8}$ (D) $\frac{3}{4}$

7. Let X be a random variable having pdf $f(x) = 3(1-x)^2;\ 0 < x < 1$.

Then the distribution of $(1 - X)^3$ will be

(A) Uniform (0, 3) (B) Uniform (0, 1)

(C) Uniform (0, 2) (D) Uniform (2, 3)

8. A four-sided die with faces marked 1, 2, 3, 4 is rolled twice. Let X be the sum of the outcomes. Then which of the following is true ?

(A) X follows Binomial $\left(8, \frac{1}{2}\right)$

(B) $P(X = x) = \frac{9-x}{28};\ x = 2, 3, 4, 5, 6, 7, 8$

(C) $P(X = x) = \frac{|x-5|+1}{19},\ x = 2, 3, 4, 5, 6, 7, 8$

(D) $P(X = x) = \frac{4-|x-5|}{16};\ x = 2, 3, 4, 5, 6, 7, 8$

9. Let X_1 and X_2 be two independent random variables with $X_1 \sim Bin\left(m, \frac{1}{2}\right)$ and $X_2 \sim Bin\left(n, \frac{1}{2}\right)$, $m \neq n$. Which of the following statements is correct ?

(A) $2X_1 + 3X_2 \sim Bin\left(2m + 3n, \frac{1}{2}\right)$ (B) $X_2 - X_1 + m \sim Bin\left(m + n, \frac{1}{2}\right)$

(C) $P(X_1 = X_2) = 0$ (D) $X_2 + X_1 - m \sim Bin\left(n, \frac{1}{2}\right)$

10. Consider two <u>independent</u> random variables X and Y where X ~ Uniform (1, 3) and Y ~ Uniform (2, 4). Then P (X ≥ Y) is

(A) $\frac{1}{8}$ (B) $\frac{3}{8}$

(C) $\frac{5}{8}$ (D) $\frac{7}{12}$

11. Let (X, Y) have a bivariate normal distribution with E(X) = E(Y) = 0, Var(X) = Var (Y) = 1 and corr (X, Y) = ρ, −1 < ρ < 1. Then E(max(X, Y)) is equal to

(A) $\left(\frac{1+\rho}{\pi}\right)^{1/2}$ (B) $\frac{\sqrt{1-\rho}}{\pi}$

(C) $\sqrt{\frac{1+\rho}{\pi(1-\rho)}}$ (D) $\left(\frac{1-\rho}{\pi}\right)^{1/2}$

12. A and B play a game of tossing a fair coin. A starts the game by tossing the coin once and B then tosses the coin twice, followed by A tossing the coin once and B tossing the coin twice and this continues until a head turns up. Whoever gets the first head wins the game. Then

(A) P(B wins) > P (A wins) (B) P(B wins) + P (A wins) < 1

(C) P(A wins) > P(B wins) (D) P(B wins) = P(A wins)

13. Five persons A, B, C, D and E are seated at random on <u>eight</u> numbered chairs which are arranged in a circle. What is the probability that A and B are separated by at least two chairs ?

(A) $\frac{3}{7}$ (B) $\frac{1}{2}$

(C) $\frac{4}{7}$ (D) $\frac{1}{4}$

14. Let X and Y be integer-valued bounded random variables. Then which of the following statements is FALSE ?

(A) $E(X) = \sum_y E(X \mid Y = y)\, P(Y = y)$

(B) $Var(X) = \sum_y Var\,(X \mid Y = y)\, P(Y = y)$

(C) $P(X = x) = \sum_y P\,(X = x \mid Y = y)\, P(Y = y)$

(D) $E(XY) = \sum_y y\, E\,(X \mid Y = y)\, P(Y = y)$

15. Suppose that (X, Y) has a joint distribution with marginal distribution of X being N(0,1) and $E(Y \mid X = x) = x^3 \;\forall\; x \in \mathbb{R}$. Then which of the following statements is true ?

(A) corr (X, Y) = 0

(B) corr (X, Y) > 0

(C) corr (X, Y) < 0

(D) X and Y are independent.

16. Let $X_1, X_2, \ldots\ldots$ be iid N(0, 1) random variables and let $T_n = \sum_{i=1}^{n} X_i^2/n$. Then which of the following statements is correct ?

(A) The limiting distribution of $T_n - 1$ is χ^2 with 1 d.f.

(B) The limiting distribution of $\dfrac{T_n - 1}{\sqrt{n}}$ is N(0, Variance = 2)

(C) The limiting distribution of $\sqrt{n}\,(T_n - 1)$ is χ^2 with 1 d.f.

(D) The limiting distribution of $\sqrt{n}\,(T_n - 1)$ is N(0, Variance = 2)

17. Suppose θ is a random variable with uniform distribution on the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. The value of the distribution function of the random variable $X = \sin\theta$ at $x \in [-1, 1]$ is

(A) $\frac{x+1}{2}$
(B) $\sin^{-1} x + \frac{\pi}{2}$
(C) $\frac{\sin^{-1}x}{\pi} + \frac{1}{2}$
(D) $\frac{\sin^{-1}x}{\pi} + \frac{\pi}{2}$

18. Let $X \sim N(0, 1)$. Then $\frac{X}{|X|}$ has

(A) a Cauchy distribution
(B) a degenerate distribution
(C) a normal distribution
(D) a two-point distribution

19. Suppose a random sample of size 4 is drawn from exponential (1) distribution. Let $Y = \max\{X_1, X_2, X_3, X_4\}$. Then $P(2 < Y)$ is

(A) $\left(1-e^{-\frac{1}{2}}\right)^4$
(B) $1-\left(1-e^{-\frac{1}{2}}\right)^4$
(C) $(1-e^{-2})^4$
(D) $1-(1-e^{-2})^4$

20. Let X_1 and X_2 be iid random variables with pmf

$$P(X_i = \pm 1) = \frac{1}{2}, \ i = 1, 2.$$

Suppose, $X_3 = X_1 X_2$. Identify the correct statement.

(A) X_1 and X_3 are independent and identically distributed.

(B) X_1 and X_3 are independent but not identically distributed.

(C) X_1 and X_3 are identically distributed but not independent.

(D) X_1 and X_3 are neither independent nor identically distributed.

21. Suppose we have 9 distinct observations on a certain variable. Let minimum and maximum of the observations be 7 and 17, respectively. Then which of the following is a possible value of the mean ?

(A) 7.5
(B) 8.1
(C) 8.2
(D) 15.8

22. Consider the data on preference of cold drinks based on a sample of 60 individuals :

Brand number of cold drink	1	2	3	4	5	Total
Frequency	10	11	25	9	5	60

An appropriate measure of central tendency for the given data would be

(A) Arithmetic mean
(B) Median
(C) Mode
(D) Trimmed mean

23. Consider the following rankings of 10 contestants by two judges :

Judge 1	1	2	3	4	5	6	7	8	9	10
Judge 2	2	1	6	5	7	9	10	4	8	3

The difference between the numbers of concordances and discordances is

(A) 10
(B) 11
(C) 12
(D) 13

24. Given a set of observations $(x_1, x_2, \ldots x_n)$ on a variable 'x', consider a function given by $f(A) = \frac{1}{2^n} \exp\left(\frac{-1}{2} \sum_{i=1}^{n} |x_i - A|\right)$. The function will be maximum when

(A) $A = 0$
(B) $A =$ mean of $x_1, x_2, \ldots, x_n$
(C) $A =$ median of $x_1, x_2, \ldots, x_n$
(D) $A =$ mode of $x_1, x_2, \ldots, x_n$

25. Let a variable x assume the values 1, 2, 3 ..., 8, 9 with 'greater-than' type cumulative frequencies 45, 40, 35, ..., 10, 5. Then $\bar{x}$ is

(A) 4 (B) 5
(C) 4.5 (D) 5.5

26. Which of the following is an exit-controlled loop ?

(A) while loop (B) for loop
(C) do-while loop (D) if-else loop

27. For acceptance-rejection method with target density f and proposal density g having common support x, which of the following statements is necessarily true ?

(A) $\frac{g(x)}{f(x)}$ is a bounded function of $x \in x$.

(B) $\frac{g(x)}{f(x)}$ is an unbounded function of $x \in x$.

(C) $\frac{f(x)}{g(x)}$ is a bounded function of $x \in x$.

(D) $\frac{f(x)}{g(x)}$ is an unbounded function of $x \in x$.

28. Suppose $X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N_3\ (\underset{\sim}{0}, \Sigma)$, where $\Sigma = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 1 \\ 3 & 1 & 10 \end{pmatrix}$. What is the distribution of $X_1 + 3X_3$?

(A) N (0, 109) (B) $N_2\left(\underset{\sim}{0}, \begin{pmatrix} 1 & 6 \\ 6 & 9 \end{pmatrix}\right)$
(C) N(0, 91) (D) N(0, 10)

29. Let $(X, Y) \sim N_2\ (0, 0, 1, 1, \rho)$ where $|\rho| \leq 1$. Then $P\ (X > 0, Y > 0)$ is

(A) $\frac{1}{2}$ if $\rho = -1$ and $\frac{1}{2}$ if $\rho = 0$. (B) $\frac{1}{2}$ if $\rho = 1$ and $\frac{1}{4}$ if $\rho = -1$.

(C) $\frac{1}{2}$ if $\rho = 1$ and $\frac{1}{4}$ if $\rho = 0$. (D) $\frac{1}{4}$ if $\rho = -1$ and $\frac{1}{2}$ if $\rho = 0$.

30. Consider the following multiple linear regression equation derived on the basis of a sample of size 100.

$\hat{P} = 119.2 + 0.485 \times BD + 23.4 \times BA + 0.156 \times HS + 0.002 \times PS + 0.090 \times A - 35.6 \times PC$,

where

P	:	Price or value of house (in ₹ 1000)
BD	:	Number of bed rooms
BA	:	Number of bathrooms
HS	:	House size (in sq ft.)
PS	:	Plot size (in sq. ft)
A	:	Age of the flat (in years)

PC : a dummy variable, PC = 1, if the house is in poor condition and = 0, otherwise.

If the homeowner converts a bedroom into a bathroom, what is the expected increase in the value of the house, assuming that no other variable changes ?

(A) ₹ 23,400
(B) ₹ 22,915
(C) ₹ 485
(D) No change

31. Let $(X, Y) \sim N_2 (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ where $\mu_1, \mu_2 \in \mathbb{R}, \sigma_1, \sigma_2 \in \mathbb{R}^+$ and $|\rho| < 1$. Then the expression for Var $(Y | X = x)$ is free of

(A) (x, ρ, μ_2)
(B) (x, μ_1, σ_1)
(C) $(\mu_1, \sigma_1, \sigma_2)$
(D) (μ_1, μ_2, ρ)

32. Let X_1, X_2 and X_3 have multinomial distribution in which n = 25, k = 4 and the unknown probabilities are θ_1, θ_2 and θ_3 respectively. Let $X_4 = 25 - X_1 - X_2 - X_3$ and $\theta_4 = 1 - \theta_1 - \theta_2 - \theta_3$. Suppose the observed values of the random variables are $x_1 = 4$, $x_2 = 11$, $x_3 = 7$. Then the maximum likelihood estimates of θ_1, θ_2 and θ_3 are respectively

(A) $\frac{4}{25}, \frac{11}{25}, \frac{7}{25}$
(B) $\frac{1}{25^4}, \frac{1}{25^{11}}, \frac{1}{25^7}$
(C) $\frac{4}{22}, \frac{11}{22}, \frac{7}{22}$
(D) $\frac{21}{25}, \frac{14}{25}, \frac{18}{25}$

33. Let $X_1, X_2, \ldots, X_n$ be a random sample from Uniform $(\theta, 5\theta)$, $\theta > 0$. Define $X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$ and $X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$. Maximum Likelihood Estimator of θ is

(A) $\frac{X_{(1)}}{5}$ (B) $X_{(1)}$

(C) $X_{(n)}$ (D) $\frac{X_{(n)}}{5}$

34. Let X_1, X_2, X_3 be a random sample from Bernoulli(p) distribution. Which of the following statistics is NOT sufficient for p ?

(A) $X_1 + X_2 + X_3$ (B) $X_1 + 2X_2 + X_3$

(C) $X_1 + 3X_2 + X_3$ (D) $X_1 + 2X_2 + 4X_3$

35. Let $\hat{\alpha}_1, \hat{\alpha}_2$ be two independent unbiased estimators of the parameter α with standard errors σ_1 and σ_2 respectively, with $\sigma_1 \neq \sigma_2$. The linear combination of $\hat{\alpha}_1$ and $\hat{\alpha}_2$ that yields an unbiased estimator of α with the smallest variance is

(A) $\frac{\sigma_1}{\sigma_1 + \sigma_2}\hat{\alpha}_1 + \frac{\sigma_2}{\sigma_1 + \sigma_2}\hat{\alpha}_2$ (B) $\frac{\sigma_2}{\sigma_1 + \sigma_2}\hat{\alpha}_1 + \frac{\sigma_1}{\sigma_1 + \sigma_2}\hat{\alpha}_2$

(C) $\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\hat{\alpha}_1 + \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\hat{\alpha}_2$ (D) $\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\hat{\alpha}_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\hat{\alpha}_2$

36. Let $X_1, X_2, \ldots, X_{100}, X_{101}$ be a random sample from $N(\mu, 1)$ distribution. Let $\bar{X}_{100} = \frac{1}{100}\sum_{i=1}^{100} X_i$. The confidence interval $[\bar{X}_{100} - 0.196, \bar{X}_{100} + 0.196]$ for μ has confidence coefficient 0.95. Then the probability that X_{101} lies in this interval is

(A) equal to 0.95

(B) less than 0.95

(C) greater than 0.95 but less than $\frac{100}{101}$

(D) equal to $\frac{100}{101}$

37. Suppose you have a sample of size one from one of the following densities :

$H_0 : f(x) = 2x \qquad 0 \le x \le 1$

$H_1 : f(x) = 2 - 2x \qquad 0 \le x \le 1$

Let α and β denote the Type I and Type II errors respectively. Consider a test procedure of the form "Reject H_0 if $x < k$" with $\alpha = 0.09$. Then

(A) $k = 0.5, \beta = 0.25$

(B) $k = 0.5, \beta = 0.36$

(C) $k = 0.3, \beta = 0.16$

(D) $k = 0.3, \beta = 0.49$

38. Let $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ be independent random samples from two absolutely continuous cdfs, F and G, respectively. Let,

$T(X_i, Y_j) = \begin{cases} 1 & \text{if } X_i < Y_j \\ 0 & \text{if } X_i \ge Y_j \end{cases}$ and

Q_j = rank of Y_j among the combined m + n observations, for i = 1(1) m and j = 1 (1) n.

Define, $W = \sum_{i=1}^{m}\sum_{j=1}^{n} T(X_i, Y_j)$ and $U = \sum_{j=1}^{n} Q_j$.

If F = G, then E(W) – E(U) is

(A) $\frac{n(n+1)}{2}$

(B) $\frac{m(m+1)}{2}$

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

(D) $-\frac{m(m+1)}{2}$

39. Let $X_1, X_2 \ldots X_{10}$ be a random sample drawn from the population with continuous cdf F_X and $Y_1, Y_2 \ldots Y_{12}$ be another random sample from the population with continuous cdf F_Y. Under the null case, the probability distribution of the Mann-Whitney 'U' will be symmetric about

(A) 0

(B) 1

(C) 60

(D) 5

40. Let b blocks of a BIBD (v, b, k, r, λ) be B_1, ..., B_b. Consider block B_m ($m \in \{1, 2..., b\}$) and let x_i be the number of treatments common between B_m and B_i ($i \neq m$) and we denote $T_m = \sum_{i(\neq m)=1}^{b} x_i$, m = 1, 2, ..., b. Then

(A) T_m's are not necessarily equal

(B) $T_m = k(r-1)\ \forall\ m = 1, ..., b$

(C) $T_m = (k-1)(r-1)\ \forall\ m = 1, ..., b$

(D) $T_m = (k-1)r\ \forall\ m = 1, ..., b$

41. A 2^4 design is arranged in 4 blocks by confounding the factors ABC and ACD. The other effect that will also be confounded with the blocks is

(A) AC (B) BD

(C) BCD (D) ABCD

42. In time series analysis, the ratio to trend method is used to

(A) estimate the trend values

(B) obtain the seasonal indices

(C) determine the cyclical component

(D) eliminate the trend component

43. The wholesale price indices are compiled and published by

(A) Central Statistical Office

(B) National Sample Survey Office

(C) Directorate General of Commercial Intelligence and Statistics

(D) Office of Economic Adviser

44. If for a c-chart, the central line is equal to 6.15, the upper control limit is

(A) 0 (B) 18.45

(C) 12.30 (D) 13.59

45. Consider a double sampling inspection plan for attributes with large lot size N. If $n_1 = n_2 = 3$, $c_1 = 1$, $c_2 = 2$ and actual proportion of defectives in the lot is p, then the probability that the lot will be accepted after second stage is (N is large enough so that Binomial approximation is valid)

(A) $6p^2(1-p)^4$

(B) $3p^4(1-p)^2$

(C) $3p^2(1-p)^4$

(D) $2p^2(1-p)^4$

46. From a life table it is known that $l_{80} = 2633$, $d_{80} = 170$ and $q_{81} = 0.11004$. Then the value of d_{81} is

(A) 271

(B) 391

(C) 662

(D) 266

47. The annual age-specific fertility rates of two cohorts A and B during a year are i_x^A and i_x^B respectively (expressed as pure rates). Further, assume the followings :

p_x^A = Probability that an individual of the cohort A living at age x will reach age $x + 1$

= p_x^B,

l_x^A = Number of individuals living at age x in cohort A

= $\frac{1}{2} l_x^B$, and

sex ratio remains unchanged for both the cohorts during the period and it is Male : Female = 2 : 1 in cohort A and Male : Female = 3 : 1 in cohort B.

Then age specific fertility rate for the combined cohort is

(A) $\frac{1}{4} i_x^A + \frac{3}{4} i_x^B$

(B) $\frac{2}{5} i_x^A + \frac{3}{5} i_x^B$

(C) $\frac{3}{5} i_x^A + \frac{2}{5} i_x^B$

(D) $\frac{1}{2} i_x^A + \frac{1}{2} i_x^B$

48. If Laspeyer's and Fisher's price indices are 124.03 and 123.49, respectively, then Paasche's price index is

(A) 124.15

(B) 123.85

(C) 123.75

(D) 122.95

49. To estimate the average work experience of MBA students in a management institute, ten students are to be selected at random from commerce, science and engineering backgrounds. Here the appropriate sampling scheme will be

(A) systematic sampling

(B) stratified sampling

(C) two-stage sampling

(D) cluster sampling

50. A simple random sample of size 5 is drawn without replacement from a finite population consisting of 41 units. If the standard error of the sample mean is 2.65, then the population standard deviation was closest to

(A) 8.09

(B) 6.25

(C) 8.20

(D) 3.31