PUMDET–2017

Subject : Statistics

Time Allowed: 1Hour 30 minutes. *Maximum Marks*: 100

Booklet No. 21600200

INSTRUCTIONS

Candidates should read the following instructions carefully before answering the questions:

1. This question paper contains 50 MCQ type objective questions. Each question has four answer options given, viz. A, B, C and D.
2. Only one answer is correct. Correct answer will fetch full marks 2. Incorrect answer or any combinations of more than one answer will fetch – ½ mark. No answer will fetch 0 mark.
3. Questions must be answered on OMR sheet by darkening the appropriate bubble marked A, B, C, or D.
4. Use only **Black/Blue ball point pen** to mark the answer by complete filling up of the respective bubbles.
5. Mark the answers only in the space provided. Do not make any stray mark on the OMR.
6. Write question booklet number and your roll number carefully in the specified locations of the OMR. Also fill appropriate bubbles.
7. Write your name (in block letter), name of the examination centre and put your full signature in appropriate boxes in the OMR.
8. The OMRs will be processed by electronic means. Hence it is liable to become invalid if there is any mistake in the question booklet number or roll number entered or if there is any mistake in filling corresponding bubbles. Also it may become invalid if there is any discrepancy in the name of the candidate, name of the examination centre or signature of the candidate vis-a-vis what is given in the candidate's admit card. The **OMR** 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.
9. Rough work must be done on the question paper itself. Additional blank pages are given in the question paper for rough work.
10. Hand over the OMR to the invigilator before leaving the Examination Hall.

1. The following gives the probability distribution of a random variable X under two possible parameter values θ_1 and θ_2:

X:	1	2	3	4
θ_1	0.01	0.04	0.05	0.9
θ_2	0.8	0.1	0.04	0.06

For testing $H_0: \theta = \theta_1$ against $H_1: \theta = \theta_2$ consider two critical regions $W_1: \{X = 1 \text{ or } 2\}$ and $W_2: \{X = 3\}$. Then which of the following statements are NOT true?

(A) Both W_1 and W_2 are of size 0.05.

(B) Both W_1 and W_2 are unbiased.

(C) W_1 is unbiased but W_2 is biased.

(D) W_1 is more powerful than W_2.

2. Suppose $X_1, X_2, \dots$ is a sequence of iid random variables with $P(X_i = 1) = p = 1 - P(X_i = 0)$, $i = 1, 2, \dots$

Let $Y = \frac{1}{500}\sum_{i=1}^{500} X_i$ and $\alpha = P(|Y - p| > 0.01)$. Then, for all p,

(A) $\alpha \leq 0.01$

(B) $\alpha \leq 0.05$

(C) $\alpha > 0.01$

(D) $\alpha = 0$

3. Suppose $X_1 \sim Uniform\ (0, \theta)$ independently of $X_2 \sim Uniform\ (0, 1 + \theta)$, then a sufficient statistic for θ is

(A) $\min\{X_1, X_2\}$

(B) $\max\{X_1, X_2\}$

(C) $\max\{X_1, X_2 - 1\}$

(D) $\max\{X_1 + 1, X_2\}$

4. Suppose $X \sim N(0, \sigma^2)$, Y has the exponential distribution with mean $2\sigma^2$ and X and Y are independent. We want to test, at level $\alpha, H_0: \sigma^2 = 1$ against $H_1: \sigma^2 > 1$, based on (X, Y). Then

(A) UMP test does not exit.

(B) UMP test rejects H_0 when $X^2 + Y$ is larger.

(C) UMP test is a chi-square test.

(D) UMP test is a t-test.

5. Suppose X is a random point within 0 and 3 and Y is a random point within 0 and 4. If X and Y are independent, the value of $P(X < Y)$ is

(A) $\frac{5}{8}$

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

(C) $\frac{1}{2}$

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

6. Suppose X_1, X_2, X_3 and X_4 are *iid N (0, 1)*. Which of the following statements is NOT true?

(A) $X_1 + X_2$ is independent of $X_1 - X_2$.

(B) $\frac{X_1+X_2}{X_3+X_4}$ has Cauchy (0, 1)

(C) $\frac{X_1^2+X_2^2}{X_3^2+X_4^2} \sim F_{2,2}$

(D) $\frac{X_1}{X_2}, \frac{X_2}{X_3}$ and $\frac{X_3}{X_4}$ follow *iid* Cauchy (0, 1)

7. Suppose $X_1, X_2, \dots, X_n$ are *iid* Bernoulli variables with success probability p and let

$Y = \max\{X_1, X_2, \dots, X_n\}$. Then

(A) Y is Bernoulli variable with success probability equal to p.

(B) Y is Bernoulli variable with success probability equal to $1 - (1 - p)^n$.

(C) Y is Bernoulli variable with success probability equal to $1 - p^n$.

(D) Y is not a Bernoulli variable.

8. $\inf\{2^x + 2^{1/x} : x > 0\}$

(A) is 0

(B) is 2

(C) is 4

(D) does not exist

9. Let S_n be the partial sums of $\sum_{n=1}^{\infty} \sin\frac{n!\pi}{720}$. Then $\lim_{n\to\infty}(S_n - S_4)$ is equal to

(A) 0

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

(C) $\frac{\sqrt{3}}{2}$

(D) None of the above

10. Determine the vector (a, b, c) so that the following function is differentiable on the set of real numbers $\mathbb{R}$.

$$f(x) = \begin{cases} 4x & if\ x \leq 0 \\ ax^2+bx+c & if\ 0 < x < 1 \\ 3-2x & if\ x \geq 1 \end{cases}$$

(A) (–3, 4, 0)

(B) (3, –4, 0)

(C) (3, 4, 0)

(D) (–3, –4, 0)

11. The number of eggs laid on a tree leaf by an insect of a certain type is a Poisson random variable with parameter 1. However, such a random variable can be observed only if it is positive, since if it is 0, then we cannot know that such an insect was on the leaf, If Y denotes the observed number of eggs then E(Y) is

(A) e

(B) $\frac{1}{e}$

(C) $\frac{1}{1-e}$

(D) $\frac{e}{e-1}$

12. Assuming 0.05 probability limits, find the minimum sample size which guarantees a positive lower control limit in a fraction nonconforming control chart when $p = 0.05$.

(A) 70

(B) 71

(C) 72

(D) 73

13. In the case of single sampling plan for attributes, find the value of ATI (average total inspection) assuming $N = 10000, n = 89, c = 2\ and\ p = 0.01$.

(A) 686

(B) 687

(C) 688

(D) Insufficient information

14. For a population with linear trend, which of the following statements is incorrect?

(A) For eliminating trend, systematic sampling is more efficient than simple random sampling.

(B) For eliminating trend, systematic sampling is more efficient than stratified random sampling.

(C) For eliminating trend, stratified sampling is more efficient than simple random sampling.

(D) For eliminating trend, stratified random sampling is more efficient than systematic sampling.

15. Find the stratified sample mean assuming 3 strata where size of the first two strata are same and third one is twice of the first one. Also consider that sample means of 3 strata are 12, 14 and 15 respectively.

(A) 13

(B) 13.5

(C) 14

(D) 14.5

16. In the case of cluster sampling, let us suppose that 5 clusters are selected from a population of 10 clusters, where each cluster contains 21 elements. Assuming the intraclass correlation coefficient as 0.4, find the efficiency of cluster sampling over SRSWR of 105 elements.

(A) $\frac{1}{42}$

(B) $\frac{1}{51}$

(C) $\frac{21}{50}$

(D) $\frac{4}{105}$

17. Let $X_1, X_2, ..., X_n$ be iid *Cauchy* (0,1) and $S_n = X_1 + X_2 + \cdots + X_n$ then $\frac{S_n}{n}$

(A) converges in probability to 0.

(B) converges in distribution to $N(0,1)$.

(C) converges in distribution to *Poisson*(1).

(D) converges in distribution to *Cauchy*(0,1).

18. Let L1 : $3x + 2y = 25$ and L2 : $6x + y = 30$ be the two regression lines in a bivariate dataset. Then

(A) L1 is the regression line of y on x.

(B) L2 is the regression line of y on x.

(C) L1 and L2 are parallel to each other.

(D) L1 and L2 will intersect on y-axis.

19. Let us consider the regression of y on x using 7th degree curve $y = a_0 + a_1x + \cdots + a_7x^7$. If e_i's denote the residuals, $c_5 = Cov(e, x^5)$ and $c_6 = Cov(e, x^6)$ then

(A) $c_5 < c_6$

(B) $c_5 > c_6$

(C) $c_5 = c_6$

(D) None of the above

20. Let $X_1, X_2, \ldots, X_n$ be iid with $P(X_1 = 1) = p = 1 - P(X_1 = -1), 0 < p < 1$. Define $Y_n = 1$ if $X_n = 1$ and $X_{n+1} = 1$ and $Y_n = 0$ otherwise. Then

(A) $\overline{Y}$ converges in probability to p.

(B) $\overline{Y}$ converges in probability to $p(1 - p)$.

(C) $\overline{Y}$ converges in probability to p^2.

(D) $\overline{Y}$ does not converges in probability.

21. The asymptotic variance of the sample median corresponding to a random sample of size n drawn from a population having variance = σ^2, median = μ and pdf = f, is

(A) $\frac{\sigma^2}{n}$

(B) σ^2

(C) $\sigma^2/4nf^2(\mu)$

(D) $1/4nf^2(\mu)$

22. The sample median as an estimator of the population mean from a N $(\mu, 1)$ distribution is

(A) unbiased and efficient

(B) biased and efficient

(C) unbiased and inefficient

(D) asymptotically unbiased

23. Let $(X_1, X_2, \ldots, X_n)$ be a sample from a $N(\lambda, 1)$ population. Define $T = \sum_{i=1}^{n} a_iX_i$ such that $\sum_{i=1}^{n} a_i = 1$ then the Rao-Blackwellisation of T conditioning on sample mean gives

(A) $\sum_{i=1}^{n} X_i$

(B) $\sum_{i=1}^{n} X_i^2$

(C) sample mean

(D) sample median

24. The results of a two-factor analysis of variance produce df = (2, 24) for the F-ratio for factor A. Based on this information, find how many levels of factor A were compared in the study.

(A) 2

(B) 3

(C) 24

(D) 25

25. Let $X_1, X_2, \ldots, X_n$ be a sample from a Shifted Exponential (μ, λ) distribution. A sufficient statistic for the family is

(A) $(X_{(1)}, X_{(n)})$

(B) $\sum_{i=1}^{n} X_i$

(C) $X_{(n)}$

(D) $(X_{(1)}, \sum_{i=1}^{n}(X_{(i)} - X_{(1)}))$

26. Let $\begin{pmatrix} X_1 \\ X^{(2)} \end{pmatrix} \sim N_p\left[\begin{pmatrix} \mu_1 \\ \mu^{(2)} \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma'_{(1)} \\ \sigma_{(1)} & \Sigma_{(2)} \end{pmatrix}\right]$, Σ is p.d., where symbols have their usual significance under partitioning. Then the square of the correlation ratio of X_1 on $X^{(2)}$ is

(A) $\frac{\sigma'_{(1)}\Sigma_{(2)}^{-1}\sigma_{(1)}}{|\Sigma_{(2)}|}$

(B) $1 - \frac{|\Sigma|}{\sigma_1^2|\Sigma_{(2)}|}$

(C) $1 - \frac{|\Sigma|}{|\Sigma_{(2)}|}$

(D) $\frac{\sigma_1^2 - \sigma'_{(1)}\Sigma_{(2)}^{-1}\ \sigma_{(1)}}{|\Sigma|}$

27. Let $X \sim N_p\,[\mu, \Sigma], \Sigma$p.d. Then for constant vectors a and $b, a'X$ and $b'X$ will be independently distributed if

(A) $a'\Sigma\, b = 0$

(B) $a'b = 0$

(C) $a'\Sigma^{-1}b = 0$

(D) $(a'b)\Sigma = 0$

28. Let $\begin{pmatrix} X_1 \\ X^{(2)} \end{pmatrix} \sim N_p \left[\begin{pmatrix} \mu_1 \\ \mu^{(2)} \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma'_{(1)} \\ \sigma_{(1)} & \Sigma_{(2)} \end{pmatrix} \right]$, Σ is p.d., where symbols have their usual significance under partitioning. Then the distribution of the multiple linear regression $X_{1.23..p}$(i. e. of X_1 on $X_2, X_3, \ldots, X_p$) is

(A) $N(\mu_1, \sigma_1^2 - \sigma'_{(1)}\Sigma_{(2)}^{-1}\sigma_{(1)})$

(B) $N(\mu_1 + \sigma'_{(1)}\mu^{(2)}, \sigma'_{(1)}\Sigma_{(2)}^{-1}\sigma_{(1)})$

(C) $N(\mu_1 + \sigma'_{(1)}\Sigma_{(2)}^{-1}\mu^{(2)}, \sigma_1^2 - \sigma'_{(1)}\Sigma_{(2)}^{-1}\sigma_{(1)})$

(D) $N(\mu_1, \sigma'_{(1)}\Sigma_{(2)}^{-1}\sigma_{(1)})$

29. Let $(X_1, X_2, \ldots, X_K)$ jointly follow *Multinomial* $(m;\, p_1, p_2, \ldots, p_K)$ with $\sum_{i=1}^{K} p_i < 1$. Then the conditional distribution of X_1, X_2 given $X_3 = x_3$, $X_4 = x_4$ and $\sum_{i=1}^{K} X_i = t$ will be

(A) *Multinomial* $\left(m - x_3 - x_4 + t;\, \frac{p_1}{1-\sum_{i=1}^{K} p_i}, \frac{p_2}{1-\sum_{i=1}^{K} p_i}\right)$

(B) *Multinomial* $\left(m - x_3 - x_4 - t;\, \frac{p_1}{1-\sum_{i=1}^{K} p_i}, \frac{p_2}{1-\sum_{i=1}^{K} p_i}\right)$

(C) *Multinomial* $\left(m - x_3 - x_4 - t;\, \frac{p_1}{p_1+p_2+\sum_{i=5}^{K} p_i}, \frac{p_2}{p_1+p_2+\sum_{i=5}^{K} p_i}\right)$

(D) *Multinomial* $\left(m - x_3 - x_4 + t;\, \frac{p_1}{\sum_{i=1}^{K} p_i}, \frac{p_2}{\sum_{i=1}^{K} p_i}\right)$

30. If $X \sim N\,(0,1)$ and $Y \sim \chi_n^2$ which of the following is always correct?

(A) $X^2 + Y \sim \chi_{n+1}^2$

(B) $\frac{X}{\sqrt{\frac{Y}{n}}} \sim t_n$

(C) $E(X^2 + Y) = 1 + n$

(D) $Var\,(X + Y) = 1 + 2n$

31. A clinical trial was conducted among n randomly chosen persons to examine whether two different skin creams A and B have different effects on the human body. The cream A was applied to one of the randomly chosen arms of each person and the cream B to the other. What kind of experimental design is this?

(A) Completely Randomized Design

(B) Randomized Block Design

(C) Latin Square Design

(D) None of the above

32. Let X be a random variable such that $E(X) = E(X^2) = 1$, then the value of $E(X^4)$ is

(A) 0

(B) 2

(C) 1

(D) 4

33. In a $(2^5, 2^2)$ factorial experiment, some treatment combinations of the key block are (1), ac, bc, ade. Then which of the following are the treatment combinations of some block of the experiment?

(A) a, b, abc. ae, b, acde, bcde, abde

(B) d, acd, bcd, de, abd, ce, abce, be

(C) e, ace, bce, ad, abc, cd, abcd, bd

(D) (1), ac, bc, ade, a, cde, abcde, bde

34. A sequence of random variables $\{X_n, n \geq 1\}$ with distribution function F_n is said to converge in law to a random variable X with distribution function F if, as $n \to \infty$,

(A) $F_n(x) \to F(x)$ for all x.

(B) $F_n(x) \to F(x)$ for all continuity points x of F.

(C) $F_n(x) \to F(x)$ for all continuity points x of F_n.

(D) $F_n(x) \to F(x)$ for some continuity points x of F_n.

35. Let X and Y be independent $N(0,1)$ variables. Then which of the following is NOT true?

(A) $X+Y$ and $X-Y$ are independent

(B) $P(X<Y)=P\left(\frac{X}{Y}<1\right)$

(C) $\frac{X}{Y}$ and $\frac{X}{|Y|}$ have the same distribution

(D) $(X,X+Y)\sim N_2\left(0,0,1,2,\frac{1}{\sqrt{2}}\right)$

36. Let A be $n\times n$ real valued matrix such that $A^2=A$. Then

(A) A must be invertible.

(B) A cannot be invertible.

(C) If A is invertible then $A=I$.

(D) Either $A=I$ or $A=0$.

37. Suppose $X=(X_1,X_2,X_3)'$ have $N_3(\mu,\Sigma)$ distribution where $\mu=(-3,1,4)$ and $\Sigma=\begin{pmatrix}1 & -2 & 0\\ -2 & 5 & 0\\ 0 & 0 & 2\end{pmatrix}$. Then which of the following random variables are NOT independent?

(A) X_2 and X_3

(B) $\frac{X_1+X_2}{2}$ and X_3

(C) X_2 and $X_2-\frac{5}{2}X_1-X_3$

(D) X_1 and $\frac{X_2+X_3}{2}$

38. Consider a random variable X with the following p.m.f. $f(x|\theta_1,\theta_2),(\theta_1,\theta_2)\in\theta$ where $\theta=\left\{\left(\frac{1}{5},5\right),\left(\frac{1}{2},2\right),\left(2,\frac{1}{2}\right),\left(5,\frac{1}{5}\right)\right\}$.

(θ_1,θ_2) / x	$\left(\frac{1}{5},5\right)$	$\left(\frac{1}{2},2\right)$	$\left(2,\frac{1}{2}\right)$	$\left(5,\frac{1}{5}\right)$
1	1/11	1/7	1/8	1/9
2	8/11	5/7	3/4	2/3
3	1/11	1/14	1/16	1/9
4	1/11	1/14	1/16	1/9

If a value of X is observed as 2, the MLE of (θ_1,θ_2) would be

(A) $\theta_1=\frac{1}{5},\theta_2=5$

(B) $\theta_1=\frac{1}{2},\theta_2=2$

(C) $\theta_1=2,\theta_2=\frac{1}{2}$

(D) $\theta_1=5,\theta_2=\frac{1}{5}$

39. Let X be a discrete random variable with m.g.f. $M_X(t)=\frac{1}{5}(1+e^t+3e^{2t})e^{-t},t\in\mathbb{R}$. Then the mean and variance of X is

(A) mean $=\frac{2}{5}$, variance $=\frac{16}{25}$

(B) mean $=\frac{1}{2}$, variance $=\frac{2}{3}$

(C) mean $=\frac{2}{7}$, variance $=\frac{13}{25}$

(D) mean $=\frac{3}{2}$, variance $=\frac{2}{3}$

40. From a population of 10 units $\{1,2,\ldots,10\}$, we list down 10 possible distinct samples as $\{1\},\{1,2\},\{1,2,3\},\ldots,\{1,2,3,\ldots,10\}$. From these possible samples, one sample is selected at random. Let π_i be the probability that the i^{th} unit is included in the selected sample. Then which of the following is true?

(A) Expected sample size = 5

(B) Expected sample size = 5.5

(C) $\sum_{i=1}^{10}\pi_i=1$

(D) $\sum_{i=1}^{10}\pi_i=5$

41. Let A and B be $n \times n$ real matrices and $C = AB - BA$. Then we must have

(A) $C = 0$

(B) $C = I$

(C) $C = -I$

(D) $C \neq I$

42. Which of the following is NOT a subspace of $\mathbb{R}^2$?

(A) $\{(x,y): x^2 + y^2 = 0, x, y \in \mathbb{R}\}$

(B) $\{(x,y): x^2 + y^2 \leq 1, x, y \in \mathbb{R}\}$

(C) $\{(x,y): x - y = 0, x, y \in \mathbb{R}\}$

(D) $\{(x,y): x + y = 0, x, y \in \mathbb{R}\}$

43. In a $(2^3, 2)$ factorial experiment, the treatment combinations in a specific block are b, c, ab, ac. Then

(A) AC is confounded

(B) Both AC and BC are confounded

(C) BC is confounded

(D) Both C and AB are confounded

44. Let $X \sim Binomial\left(2m, \frac{1}{2}\right)$. Define for each $k \in \mathbb{N}, a(m,k) = \frac{4^m}{\binom{2m}{m}} \Pr\{X = m + k\}$. Then which of the following statements is true?

(A) $\lim_{m \to \infty} (a(m,k))^m = e^{-k(k+1)}$

(B) $\lim_{m \to \infty} (a(m,k))^m = e^{-k(k-1)}$

(C) $\lim_{m \to \infty} (a(m,k))^m = e^{-k^2}$

(D) $\lim_{m \to \infty} (a(m,k))^m = e^{-k^2/2}$

45. Let (X,Y) be uniformly distributed over a circle of radius 1. Then the expected distance of (X,Y) from the origin is

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

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

(C) $\frac{2}{3}$

(D) $\frac{1}{3}$

46. Let (X,Y,Z) have joint density given by $f(x,y,z) = \begin{cases} \frac{1}{243} xy e^{\frac{-(x+y+z)}{3}} & if\ x,y,z > 0 \\ 0, & otherwise \end{cases}$

Then the value of $\Pr\{X < Y < Z\}$ is

(A) $\frac{7}{108}$

(B) $\frac{10}{243}$

(C) $\frac{1}{18}$

(D) $\frac{1}{27}$

47. If $(X,Y) \sim N_2(0,0,1,1,1)$ then the best predictor of X based on $Y = y$ is

(A) $y + 1$

(B) y

(C) $-y + 1$

(D) $y - 1$

48. Suppose X_1, X_2, X_3 are $i.i.d.$ $N(0,1)$ random variables and U_1, U_2, U_3 are $i.i.d.$ $U(0,1)$ random variables independent of X_1, X_2, X_3. Define the variables $W = \frac{U_1X_1 + U_2X_2 + U_3X_3}{\sqrt{U_1^2 + U_2^2 + U_3^2}}$, $V = U_1 + U_2 + U_3$.

Then the correlation between W and V is

(A) 1

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

(C) 0

(D) None of the above

49. A fair die is rolled repeatedly. Then the probability that a "2" will show up before a "5" or "6" appears is

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

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

(C) $\frac{1}{2}$

(D) $\frac{2}{3}$

50. Suppose $X_1, X_2, \ldots, X_n$ are *iid* with common pdf $f(x|\mu,\sigma) = \frac{1}{\sigma} e^{-\frac{1}{\sigma}(x-\mu)}, x > \mu, \sigma > 0$. Then the MLE of μ has the pdf

(A) $f(y|n\mu,\sigma)$

(B) $f\left(y|\frac{\mu}{n},\sigma\right)$

(C) $f(x|\mu,n\sigma)$

(D) $f\left(x|\mu,\frac{\sigma}{n}\right)$

Space for rough work

Space for rough work

Space for rough work