PART A

- Find the correct answer and mark it on the answer sheet on the **top page**.
- A right answer gets **1 mark** and a wrong answer gets $-\frac{\mathbf{1}}{\mathbf{3}}$ **marks**.

1. If $P(A) = \frac{1}{2}$ and $P(B) = \frac{1}{3}$ then $P(A \bigcap B)$ is

   (a) equal to $\frac{1}{6}$.

   (b) greater than equal to $\frac{1}{6}$.

   (c) equal to $\frac{5}{6}$.

   (d) less than equal to $\frac{5}{6}$.

2. Three numbers are drawn from the set $\{1, 2, \cdots, 100\}$ by SRSWOR. The probability that the largest of the three belongs to the set $\{61, 62, \cdots, 70\}$ and the smallest of the three belongs to the set $\{11, 12, \cdots, 20\}$ is

   (a) at least $\frac{1}{5}$ but less than $\frac{2}{5}$.

   (b) at least $\frac{1}{10}$ but less than $\frac{1}{5}$.

   (c) at least $\frac{2}{10}$.

   (d) less that $\frac{1}{10}$.

3. If boys and girls are born equally likely, the probability that in a family with three children exactly one child is a girl is

   (a) $\frac{1}{3}$.

   (b) $\frac{1}{2}$.

   (c) $\frac{3}{8}$.

   (d) $\frac{5}{8}$.

4. For two dependent random variables X and Y, $E(Y|X = x) = 2x$. Suppose the marginal distribution of X is Uniform on $(0, 1)$, then $E(Y)$

   (a) is 1.

   (b) is 2.

   (c) is $\frac{1}{2}$.

   (d) cannot be determined based on the given information.

5. Let $X_1, ..., X_n$ be i.i.d. random variables from a distribution $F(x; \theta)$, which is continuous in x. Then $Y = -\sum_{i=1}^{n} logF(X_i, \theta)$ follows

   (a) Uniform distribution.

   (b) Beta distribution.

   (c) Gamma distribution.

   (d) Normal distribution.

6. Suppose X is a random variable with p.m.f. $P(X = 3^n) = \frac{2}{3^n}$, $n = 1, 2, ...$ Then

   (a) the moment generating function of X exists.

   (b) the first moment of X exists, but not the second.

   (c) the variance of X exists.

   (d) none of the moments of X exist.

7. $X_1, \cdots, X_8$ are i.i.d. Normal$(0, 1)$ random variables and let $\bar{X}_7$ be the mean of $X_1, \cdots, X_7$. Then the distribution of $k\ (\bar{X}_7 - X_8)^2$ is

   (a) χ^2 with 1 degree of freedom for $k = \frac{7}{8}$.

   (b) χ^2 with 1 degree of freedom for $k = \frac{8}{7}$.

   (c) χ^2 with 7 degrees of freedom for $k = \frac{1}{8}$.

   (d) χ^2 with 8 degrees of freedom for $k = \frac{1}{7}$.

8. If $X_1, ..., X_n$ be i.i.d. random variables from $N(\mu, \sigma^2)$, σ^2 known. The $UMVUE$ of μ^2 is given by

   (a) $\bar{X}^2$.

   (b) $\bar{X}^2 - \frac{\sigma^2}{n}$.

   (c) $\frac{(\bar{X} - \sigma)^2}{n}$.

   (d) none of the above.

9. Let X have a distribution belonging to exponential family with parameter θ. Let $\hat{\theta}$ be the maximum likelihood estimator of θ. Then which of the following is incorrect?

   (a) $\hat{\theta}$ is an unbiased estimator of θ.

   (b) $\hat{\theta}$ is consistent for θ.

   (c) The asymptotic distribution of $\hat{\theta}$ is normal.

   (d) $\hat{\theta}$ is a function of a sufficient statistic for θ.

10. The proportion of households with $0, 1, 2$ and 3 cars are $1 - 6\theta,\ 3\theta,\ 2\theta$ and θ respectively, $\theta < \frac{1}{6}$. In a random sample of 5 households, 3 had no car, 1 had 1 car and 1 had 3 cars. The maximum likelihood estimate for the proportion of households with 2 cars is

    (a) $\frac{1}{5}$.

    (b) $\frac{1}{4}$.

    (c) $\frac{1}{3}$.

    (d) $\frac{1}{2}$.

11. $X_1, \cdots, X_n$ is a random sample from Normal(μ, σ^2) population where σ^2 is given to be σ_0^2. To test $H_o : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$, the p-value based on the sample observations is p_0. Suppose the actual value of σ^2 is $\sigma_1^2 < \sigma_0^2$. Then the p-value for the same observations

    (a) will be more that p_0.

    (b) will be equal to p_0.

    (c) will be less that p_0.

    (d) will have no specific relation with p_0.

12. Consider the linear regression $E(Y) = \alpha + \beta x$. Let ρ be the correlation coefficient between X and Y. Then

(a) $\beta > \rho$.

(b) $\beta < \rho$.

(c) $-1 \leq \beta\rho \leq 1$.

(d) $\beta\rho \geq 0$.

13. The quadratic form $x_1^2 - 3x_1x_2 + x_2^2 + x_3^2$ is

(a) indefinite.

(b) positive semi definite.

(c) positive definite.

(d) negative definite.

14. In a linear model on $\underline{Y}$ with $E(\underline{Y}) = A\underline{\theta}$ and $D(\underline{Y}) = \gamma^2 I$, suppose $\underline{c}'\underline{\theta}$ is nonestimable. Then

(a) $\underline{c}'\underline{\theta}$ has a non linear unbiased estimator.

(b) $\underline{c}'\underline{\theta}$ has a consistent but biased estimator.

(c) $\underline{c}'\underline{\theta}$ has an unbiased estimator, the variance of which attains Cramer-Rao lower bound.

(d) $\underline{c}'\underline{\theta}$ is not identifiable in the model.

15. Observations on uncorrelated random variables Y_1, Y_2, Y_3 with common variance σ^2 are available.
Suppose $E(Y_1) = \theta_1 - \theta_2 + \theta_3$; $E(Y_2) = \theta_1$; $E(Y_3) = \theta_3 - \theta_2$. Then

(a) $\theta_1, \theta_2, \theta_3$ are all estimable.

(b) $\theta_3 - \theta_2$ is not estimable.

(c) θ_3 is not estimable.

(d) θ_1 is not estimable.

16. Suppose $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$, $|\rho| < 1$.

Then $\text{Var}[X_1 - X_2 | X_1 + X_2 = x]$ is

(a) $(1-\rho)$.

(b) $2(1-\rho)$.

(c) $\rho(1-\rho^2)$.

(d) $(1-\rho^2)$.

17. $\{X_n\}$ is a sequence of independent random variables with probability mass functions $P[X_n = -\sqrt{n}\,] = P[X_n = +\sqrt{n}\,] = \dfrac{1}{n}$ and $P[X_n = 0\,] = 1 - \dfrac{2}{n}$. Let $S_n = X_1 + ... + X_n$. Then

   (a) $\dfrac{S_n}{n}$ does not converge to 0 in probability.

   (b) $\dfrac{S_n}{n}$ converges to 0 in probability, but not with probability 1.

   (c) $\dfrac{S_n}{n}$ does not converge to 0 in mean.

   (d) $\dfrac{S_n}{n}$ converges to 0 with probability 1.

18. ϕ_1 and ϕ_2 are characteristic functions of two random variables. Consider the following functions for $t \in \mathcal{R}$.

$$\psi_1(t) = 1 + \phi_1(t), \qquad \psi_2(t) = \frac{1 + \phi_1(t)}{2},$$
$$\psi_3(t) = \frac{\phi_1(t) + \phi_2(t)}{3} + \frac{2}{3}, \qquad \psi_4(t) = \frac{1 + \phi_1(t)\phi_2(t)}{2}.$$

   Then

   (a) only ψ_1 is not a characteristic function.

   (b) only ψ_2 is a characteristic function.

   (c) only ψ_2 and ψ_3 are characteristic functions.

   (d) all $\psi_1, \psi_2, \psi_3, \psi_4$ are characteristic functions.

19. Consider a Markov chain with four states of 1,2,3 and 4 with the transition probability matrix $\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}$. This Markov chain

   (a) is irreducible.

   (b) has one absorbing state.

   (c) has two absorbing states.

   (d) has a null state.

20. For a Latin Square Design, the error degrees of freedom is 20. Hence the number of treatments is

    (a) 6.
    (b) 5.
    (c) 4.
    (d) 3.

21. A BIB Design with parameters (v, b, r, k, λ) is

    (a) connected, balanced and orthogonal.
    (b) connected, not balanced and orthogonal.
    (c) connected, balanced and non-orthogonal.
    (d) not connected, balanced and non-orthogonal.

22. For a linear programming problem

$$\begin{aligned} \min \quad & \mathbf{cx}, \\ \mathbf{Ax} & \leq \mathbf{b}, \\ \mathbf{x} & \geq \mathbf{0} \end{aligned}$$

    the second variable in the optimal dual solution is positive. Then in the optimal primal solution

    (a) the second variable must be zero.
    (b) the second variable must be positive.
    (c) the second constraint must be slack.
    (d) slack variable for the second constraint is zero.

23. An example of a function which is continuous but not differentiable at $x = -3$ is

    (a) $f(x) = |x + 3|^2 - 2$.
    (b) $f(x) = |x - 3|^2 + 2$.
    (c) $f(x) = |x - 3| - 2$.
    (d) $f(x) = |x + 3| + 2$.

24. If $f : \mathcal{R} \to \mathcal{R}$ is a continuous function and $f(1) = f(2) = 3$ and $\underset{x\to\infty}{lt} f(x) = -\infty$ then

    (a) f has a maximum between 1 and 2.
    (b) f has a minimum between 1 and 2.
    (c) $f(x_0) = 0$ for some $x_0 > 2$.
    (d) $f(x_0) = 0$ for some $x_0 < 1$.

25. The limit of the sequence $\{a_n\}$ as $n \to \infty$, where

$$a_n = \left(1 - \frac{t^2}{2n} + \frac{e^{-nt^2}}{n}\right)^n, \; t \in \mathcal{R}$$

(a) is 0.

(b) is $\exp\{-\frac{t^2}{2}\}$.

(c) is $\exp\{-\frac{t^2}{2} + e^{-t}\}$.

(d) does not exist.

Part B

- There are **12** questions in this part. Answer any **7** questions.
- Question 1 carries **8 marks** and all the other questions carry **7 marks** each.
- The answers should be written in the separate answer script provided to you.

1. X is a random variable with probability density function (p.d.f)

$$f_X(x) = \frac{1}{2}\, e^{-|x|}, \quad -\infty < x < \infty.$$

Find the value of α for which $Pr(|X-\alpha| > 1) = \frac{1}{2}$.

2. The joint p.d.f. of random variables X and Y is given by

$$f(x,y) = \begin{cases} \frac{1}{2}\,(x+y)e^{-x-y}, & x, y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Find $E[Y|X=x]$.

3. Let X_1 and X_2 be i.i.d random variables from a distribution with p.d.f.

$$f(x;\theta) = \begin{cases} \theta e^{-\theta x}, & x \geq 0,\ \theta > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Let $Y_1 = (X_1X_2)^{1/2}$ and $Y_2 = X_1 + X_2$. Obtain an unbiased estimator of $\frac{1}{\theta}$ based on (a) Y_1 , (b) Y_2. Which estimator will you prefer? Why?

4. $X_1, \cdots, X_n$ are i.i.d. random variables from Normal$(\theta, a\theta^2)$ distribution where a is a known positive constant and $\theta \in \mathcal{R}$. Find a nontrivial sufficient statistic for θ and verify whether it is complete.

5. $X_1, \cdots, X_n$ are i.i.d. random variables from a distribution with p.d.f.

$$f(x;\theta,\lambda) = \begin{cases} \frac{\lambda\theta^\lambda}{x^{\lambda+1}}, & x > \theta;\ \theta,\ \lambda > 0 \\ 0, & \text{otherwise.} \end{cases}$$

   (a) Find the Most Powerful test of size α to test

$$H_o : \theta = 1 \ \text{ against } \ H_1 : \theta = \theta_1 > 1.$$

   (b) For $n = 1$, specify the critical region such that the size of the test is 0.01.

   (c) For $n = 1$, if the observed sample is $X = 2$, specify the corresponding p–value if $\lambda = 2$.

6. Let $X_1, ..., X_n$ be i.i.d from Uniform distribution over the interval $(-\theta, \theta)$. Find the maximum likelihood estimator of θ. What is the maximum likelihood estimate when the observed random sample is $-3,\ 4,\ 7,\ -1,\ 5,\ -6,\ 9$?

7. Consider the linear model

$$y_i = E(y_i) + \epsilon_i, \quad i = 1, \cdots, 6$$

where $E(y_1) = \alpha_1 + \alpha_2,\ E(y_2) = E(y_4) = \alpha_1 + \alpha_3,\ E(y_3) = \alpha_2 - \alpha_3,$ $E(y_5) = 2E(y_1),\ E(y_6) = 3E(y_2)$ and $\epsilon_1, ..., \epsilon_6$ are i.i.d $N(0, \sigma^2)$.

   (a) Write down two linear unbiased estimators of $4\alpha_1 + 4\alpha_2$. Which one do you prefer? Why?

   (b) Two different solutions of the normal equations are given below

$$\begin{aligned} \text{Solution 1}: \quad \widehat{\alpha_1} &= 0, \\ \widehat{\alpha_2} &= \frac{1}{71}(12y_1 + y_2 + 11y_3 + y_4 + 24y_5 + 3y_6), \\ \widehat{\alpha_3} &= \frac{1}{71}(y_1 + 6y_2 - 5y_3 + 6y_4 + 2y_5 + 18y_6), \\ \text{Solution 2}: \quad \widehat{\alpha_1} &= \frac{1}{71}(y_1 + 6y_2 - 5y_3 + 6y_4 + 2y_5 + 18y_6), \\ \widehat{\alpha_2} &= \frac{1}{71}(11y_1 - 5y_2 + 16y_3 - 5y_4 + 22y_5 - 15y_6), \\ \widehat{\alpha_3} &= 0. \end{aligned}$$

   Notice that $\widehat{\alpha_2}$ is different in the two solutions but $\widehat{\alpha_1} + \widehat{\alpha_2}$ is the same in both the solutions. Why does this happen?

   (c) What is the minimum variance linear unbiased estimator of $4\alpha_1 + 4\alpha_2$?

8. Let $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim N_3\left(\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0.5 \\ 1 & 4 & 2 \\ 0.5 & 2 & 6 \end{bmatrix} \right)$.

   (a) Find the joint distribution of $Y_1 = 2X_1 - X_2$ and $Y_2 = X_2 - 2X_3$.

   (b) Verify whether Y_1 and Y_2 are independent.

9. Suppose $X_1, X_2, \cdots$ are i.i.d random variables with $\text{Var}(X_1) < \infty$. Show that

$$Z_n \equiv \frac{1}{n(n+1)} \sum_{j=1}^{n} jX_j \xrightarrow{p} EX_1 \quad \text{as} \quad n \to \infty.$$

10. Everyday, a salesman travels from one of the four cities $\{A, B, C, D\}$ to another according to the following scheme.

If he is in city A on a particular day, the next day he travels either to city B or city D with equal probabilities. If he is in city B on a particular day, he travels to city C the next day. If he is in city C on a particular day, the next day he travels to city D. Lastly, if he is in city D on a particular day, the next day he either travels to city A or stays back in city D with equal probabilities.

Let X_n represent the city in which the salesman is on the n^{th} day and assume that $\{X_n, n \geq 1\}$ is a Markov chain.

(a) Identify the states of the Markov Chain and give the transition probability matrix.

(b) Classify the states in to recurrent and transient classes.

(c) If the salesman is in city C on the second day, calculate the probability that he is in city D on the fifth day.

(d) Obtain the probability that the salesman is in city D in the long run.

11. Construct a 2^4 factorial design with factors A, B, C, D in 4 blocks of 4 plots each such that ABC and ACD are confounded. Identify all the confounded effects.

12. Consider a TV company which has 3 warehouses and 2 retail stores in a city. Each warehouse has a given level of supply s_i, $i = 1, 2, 3$ and each retail store has a given level of demand d_j, $j = 1, 2$. The transportation costs between warehouse i and retail store j is given by c_{ij}, $i = 1, 2, 3$; $j = 1, 2$. Given

$$s_1 = 45,\ s_2 = 60,\ s_3 = 35,\ d_1 = 50,\ d_2 = 60,$$

$$c_{11} = 3,\ c_{12} = 2,\ c_{21} = 1,\ c_{22} = 5,\ c_{31} = 2,\ c_{32} = 4,$$

formulate the problem as a Linear Programming problem and obtain an initial basic feasible solution. Is this optimal? Explain.