Probability Questions

G V V Sharma*

1. There are 30 questions in a certain multiple choice examination paper. Each question has 4 options and exactly one is to be marked by the candidate. Three candidates A,B,C mark each of the 30 questions at random independently. The probability that all the 30 answers of the three students match each other perfectly is?
2. Consider a Markov Chain with state space {0,1,2,3,4} and transition matrix

$$P = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

3. Consider the function $f(x)$ defined as $f(x) = ce^{-x^4}, x \in \mathbb{R}$. For what value of c is f a probability density function?
   a) $\frac{2}{\Gamma(\frac{1}{4})}$
   b) $\frac{4}{\Gamma(\frac{1}{4})}$
   c) $\frac{3}{\Gamma(\frac{1}{3})}$
   d) $\frac{1}{4\Gamma(4)}$
4. A random sample of size 7 is drawn from a distribution with p.d.f.

$$f_\theta(x) = \begin{cases} \frac{1+x^2}{3\theta(1+\theta^2)}, & -2\theta \le x \le \theta, \theta > 0 \\ 0, & otherwise \end{cases}$$

   and the observations are 12, -54, 26, -2, 24, 17, -39. What is the maximum likelihood estimation of θ.
   a) 12
   b) 24
   c) 26
   d) 27
5. Let X_1, X_2, X_3, X_4, X_5be i.i.d random variables having a continuous distribution function. Then P$(X_1 > X_2 > X_3 > X_4 > X_5 | X_1 = \max(X_1, X_2, X_3, X_4, X_5))$ equals
   a) $\frac{1}{4}$
   b) $\frac{1}{5}$
   c) $\frac{1}{4!}$
   d) $\frac{1}{5!}$
6. Suppose (X,Y) follows bivariate normal distribution with means $\mu 1, \mu 2$, standard deviations $\sigma 1$, $\sigma 2$ and correlation coefficient ρ, where all parameters are un-known. Then, testing Ho: $\sigma 1 = \sigma 2$ is equivalent to testing the independence of
   a) X and Y
   b) X and X-Y
   c) X+Y and Y
   d) X+Y and X-Y
7. Let $X_1, X_2...$ be the sequence of independent normally distributed random variables with mean1 and variance 1. Let N be the poisson random variable with mean 2, independent of $X_1, X_2.....$ Then the variance of $X_1 + X_2 + ... + X_{N+1}$ is
   a) 3
   b) 4
   c) 5
   d) 9
8. Suppose $r_{1.23}$ and $r_{1.234}$ are sample multiple correlation coefficients of X_1 on X_2, X_3 and X_1 on X_2, X_3, X_4 respectively. Which of the following is possible?
   a) $r_{1.23} = -0.3$ and $r_{1.234} = 0.7$
   b) $r_{1.23} = 0.7$ and $r_{1.234} = 0.3$
   c) $r_{1.23} = 0.3$ and $r_{1.234} = 0.7$

*The author is with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad 502285 India e-mail: gadepall@iith.ac.in.

d) $r_{1.23} = 0.7$ and $r_{1.234} = -0.3$

9. Suppose X~ Cauchy(0,1). Then the distribution of $\frac{1-X}{1+X}$ is?

10. There are two sets of observations on a random vector (X,Y). Consider a simple linear regression model with an intercept for regressing Y on X. Let β_i be the least square estimate of the regression coefficient obtained from the i_{th} (i=1,2) set consisting of $n_i$observations ($n_1, n_2 > 2$). Let β_0 the least square estimate obtained from the pooled sample size $n_1 + n_2$. If it is known that $\beta_1 > \beta_2 > 0$, which of the following statements is true ?

   a) $\beta_2 < \beta_0 < \beta_1$
   b) $\beta_0$may lie outside (β_2,β_1)but cannot exceed $\beta_1 + \beta_2$
   c) β_0 may lie outside (β_2,β_1) but it cannot be negative
   d) β_0 can be negative

11. Let X and Y be i.i.d random variables uniformly distributed on (0,4). Then $P(X > Y|X < 2Y)$ is

12. A sample of size $n = 2$ is drawn from a population of size $N = 4$ using probability proportional to size without replacement sampling scheme, where the probabilities proportional to size are
   The probability of inclusion of unit 1 in the sample is

i	1	2	3	4
P$_i$	0.4	0.2	0.2	0.2

TABLE 12

   a) 0.4
   b) 0.6
   c) 0.7
   d) 0.75

13. Suppose that $\begin{pmatrix} X \\ Y \end{pmatrix}$ has a bivariate density $f = \frac{1}{2}f_1 + \frac{1}{2}f_2$, where f_1 and f_2 are respectively, the densities of bivariate normal distribution $N(\mu_1, \sum)$, and $N(\mu_2, \sum)$, with $\mu_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mu_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ and $\sum = \mathbb{I}_2$, the 2×2 identity matrix. Then which of the following is correct ?
   a) X and Y are positively correlated
   b) X and Y are negatively correlated
   c) X and Y are uncorrelated but they are not independent
   d) X and Y are independent

14. Suppose X_n is a Markov chain with 3 states and a transition probability matrix

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{0.0.1}$$

   Then which of the following statement is true:
   a) X_n is irreducible
   b) X_n is recurrent
   c) X_n does not admit a stationary probability distribution
   d) X_n has an absorbing state

15. Let $X_1, X_2, ..., X_n$ be a random sample from $f_\theta(x)$, a probability density function or a probability mass function. Defined
   $s_n^2 = \frac{\sum_{n=1}^{n}(X_i - \overline{X})^2}{n-1}$, $\overline{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$
   Then s_n^2 is unbiased for θ if:
   a) $f_\theta(x) = \frac{e^{-\theta}\theta^x}{x!}; x > 0$
   b) $f_\theta(x) = \frac{e^{\frac{-x^2}{2\theta}}}{\sqrt{2\pi\theta}} - \infty < x < \infty, \theta > 0$
   c) $f_\theta(x) = \frac{e^{\frac{-x}{\theta}}}{\theta} x > 0, \theta > 0$
   d) $f_\theta(x) = \theta e^{-\theta x} x > 0, \theta > 0$

16. Given the observations $0.8, 0.71, 0.9, 1.2, 1.68, 1.4, 0.88, 1.62$ from the uniform distribution on $(\theta - 0.2, \theta + 0.8)$ with $-\infty < \theta < \infty$, which of the following is a maximum likelihood estimate for θ ?
   a) 0.7
   b) 0.9
   c) 1.1
   d) 1.3

17. Consider a linear model $Y_l = \theta_1 + \theta_2 + \epsilon_l$ for $l = 1, 2$ and $Y_l = \theta_1 - \theta_3 + \epsilon$ for l = 3,4, where ϵ_l are independent with $R(\epsilon_i) = 0$, $Var(\epsilon_i) = \sigma^2 > 0$ for i = 1, ... 4, and $\theta_1, ...\theta_3 \in R$. Which of the following parametric function is estimable?
   a) $\theta_1 + \theta_3$
   b) $\theta_2 - \theta_3$

c) $\theta_2 + \theta_3$
d) $\theta_1 + \theta_2 + \theta_3$

18. To test the hypothesis H_0 against H_1 using the test statistic T, the proposed test procedure is not to support H_0 if T is large. Based on a given sample, the p value of the test statistic is computed to be 0.05 assuming that the distribution of T is N(0,1) under H_0. If the distribution of T under H_0 is the t-distribution with 10 degrees of freedom instead, the p-value will be :

19. Let $X \geq 0$ be a random variable on $(\Omega, \mathcal{F}, P)$ with $\mathbb{E}(X) = 1$. Let $A \in \mathcal{F}$ be an event with $0 < P(A) < 1$. Which of the following defines another probability measure on $(\Omega, \mathcal{F})$?
   a) $Q(B) = P(A \cap B)\ \forall B \in \mathcal{F}$
   b) $Q(B) = P(A \cup B)\ \forall B \in \mathcal{F}$
   c) $Q(B) = \mathbb{E}(XI_B)\ \forall B \in \mathcal{F}$
   d) $Q(B) = \begin{cases} P(A/B) & P(B) > 0 \\ 0 & P(B) = 0 \end{cases}$

20. Consider a Markov Chain having state space $S = 1, 2, 3, 4$ with a transation probablity matrix $P = (p_{i,j})$ given by

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix} \end{array}$$

   Then
   a) $\lim_{n\to\infty} p_{2,2}^{(n)} = 0, \qquad \sum_{n=0}^{\infty} p_{2,2}^{(n)} = \infty$
   b) $\lim_{n\to\infty} p_{2,2}^{(n)} = 0, \qquad \sum_{n=0}^{\infty} p_{2,2}^{(n)} < \infty$
   c) $\lim_{n\to\infty} p_{2,2}^{(n)} = 1, \qquad \sum_{n=0}^{\infty} p_{2,2}^{(n)} = \infty$
   d) $\lim_{n\to\infty} p_{2,2}^{(n)} = 1, \qquad \sum_{n=0}^{\infty} p_{2,2}^{(n)} < \infty$

21. Let X_1, X_2... be a sequence of independent normally distributed random variables with mean 1 and variance 1. Let N be a Poisson random variable with mean 2 , independent of X_1, X_2.... Then the variance of $X_1 + X_2 + ... + X_{N+1}$ is
   a) 3
   b) 4
   c) 5
   d) 9

22. Let $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3) \ldots (X_n, Y_n)$ be n independent observations from a distribution. Let r_p be the product moment correlation coefficient and r_s be the rank correlation coefficient computed based on n observations. Which of the following statements is correct.
   a) $r_p \geqslant 0$ implies $r_s \geqslant 0$
   b) $r_s \geqslant 0$ implies $r_p \geqslant 0$
   c) $r_p = 1$ implies $r_s = 1$
   d) $r_s = 1$ implies $r_p = 1$

23. Let X and Y be i.i.d uniform (0,1) random variables. Let Z = max(X,Y) and W = min(X,Y). Then P((Z-W)> 0.5) =

24. In a 2^4 experiment with two blocks and factors A, B, C and D, one block contains the following treatment combinations $a, b, c, ad, bd, cd, abc, abcd$. Which of the following effects is confounded?
   a) ABC
   b) ABD
   c) BCD
   d) $ABCD$

25. Suppose that the lifetime of an electric bulb follows an exponential distribution with mean θ hours. In order to estimate θ, n bulbs are switched on at the same time. After t hours, $n-m(>0)$ bulbs were found to be in functioning state. If the lifetimes of the other $m(>0)$ bulbs are noted as $x_1, x_2, x_3, ..., x_m$, respectively, then the maximum likelihood estimate of θ is given by -

26. If $X \sim N_p(0, I)$ and A_{pxp} is an idempotent matrix with $rank(A) = k < p$, then which of the following statements is correct?
   a) $\frac{X'AX}{X'X} \sim \frac{k}{p} F_{k,p}$
   b) $\frac{X'AX}{X'X} \sim \frac{k}{p-k} F_{k,p-k}$
   c) $\frac{X'AX}{X'X} \sim Beta\left(\frac{k}{2}, \frac{p}{2}\right)$
   d) $\frac{X'AX}{X'X} \sim Beta\left(\frac{k}{2}, \frac{p-k}{2}\right)$

27. A standard fair die is rolled until some face other than 5 or 6 turns up. Let X denote the face value of the last roll, and A = [X is even] and B = [X is at most 2]. Then
   a) $P(A \bigcap B) = 0$
   b) $P(A \bigcap B) = 1/6$
   c) $P(A \bigcap B) = 1/4$
   d) $P(A \bigcap B) = 1/3$

28. A simple random sample of size n will be drawn from a class of 125 students, and the

mean mathematics score of the sample will be computed. If the standard error of the sample mean for ”with sampling” is twice as much as the standard error of the sample mean for ”without replacement sampling”, the value of n is-

a) 32
b) 63
c) 79
d) 94

29. Twenty items are put in a life testing experiment starting at time 0. The failure times of the items are recorded in a sequential manner. The experiment stops if all the twenty items fails or a pre-fixed time T$\geqslant$ 0 is reached, whichever is earlier. If the lifetimes of the items are independent and identically distributed exponential random variables with mean θ, where $0 < \theta < 10$, then which of the following statements are correct ?
   a) The MLE of θ always exists.
   b) The MLE of θ does not exist.
   c) The MLE of θ is an unbiased estimator of θ if it exists.
   d) The MLE of θ is bounded with probability 1, if it exists.
30. Suppose that the lifetime of an electric bulb follows an exponential distribution with mean θ hours. In order to estimate θ, n bulbs are switched on at the same time. After t hours, n-m(>0) bulbs are found in functioning state. If the lifetimes of other m(>0) bulbs are noted as $x_1, x_2, ..., x_m$ respectively, then the maximum likelihood estimate of θ is given by -
31. In an airport, domestic passengers and international passengers arrive independently according to Poisson processes with rates 100 and 70 per hour. If it is given that the total no. of passengers (domestic and international) arriving in that airport between 9:00 AM and 11:00 AM on a particular day was 520, then what is the conditional distribution of the no. of domestic passengers arriving in this period ?
   a) Poisson (200)
   b) Poisson (100)
   c) Binomial (520,$\frac{10}{17}$)
   d) Binomial (520,$\frac{7}{17}$)
32. The covariance matrix of four dimensional random vector X is of the form
   $$\begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}, \text{ Where } \rho < 0$$
   If v is the variance of the first principle component then,
   a) v cannot exceed $\frac{5}{4}$
   b) v can exceed $\frac{5}{4}$ but cannot exceed $\frac{4}{3}$
   c) v can exceed $\frac{4}{3}$ but cannot exceed $\frac{3}{2}$
   d) v can exceed $\frac{3}{2}$
33. Consider the problem of estimation of a parameter θ on the basis of X, where X~N(θ,1) and $-\infty < \theta < \infty$. Under squared error loss, X has uniformly smaller risk than that of kX, for
   a) $k < 0$
   b) $0 < k < 1$
   c) $k > 1$
   d) no value of k
34. Suppose A, B, C are events in a common probability space with $P(A) = 0.2$, $P(B) = 0.2$, $P(C) = 0.3$, $P(AB) = 0.1$, $P(AC) = 0.1$, $P(BC) = 0.1$ Which of the following are possible values of $P(A + B + C)$?
   a) 0.5
   b) 0.3
   c) 0.4
   d) 0.9
35. Let $X_1, X_2, X_3..X_n$ be i.i.d. uniform (θ_1, θ_2) random variables, where $\theta_1 < \theta_2$ are unknown parameters. Which of the following is an ancillary statistic?
   a) $\frac{X_{(k)}}{X_{(n)}}$ for any $k < n$
   b) $\frac{X_{(n)}-X_{(k)}}{X_{(n)}}$ for any $k < n$
   c) $\frac{X_{(k)}}{X_{(n)}-X_{(k)}}$ for any $k < n$
   d) $\frac{X_{(k)}-X_{(1)}}{X_{(n)}-X_{(k)}}$ for any $k < n$ where $1 < k < n$
36. To test the equality of effects of 10 schools against all alternatives, we take a random sample of 5 students from each school and note their marks in a common examination. ”Between sum of squares” and ”total sum of squares” are found to be 180 and 500 respectively. What is the p-value for the standard F-test?
37. Suppose $X_1, \ldots, X_n$ are i.i.d. random vectors from $N_p(0, \Sigma)$. Let $l \in R^p$, $E(\Sigma_{i=1}^n l^t X_i X_i^t l) = c$ and $E(\Sigma_{i=1}^n X_i X_i^t) = A$.

Which of the following statements are necessarily true?

a) $c = l^t l$

b) $l^t(\Sigma_{i=1}^{n} X_i X_i^t) l$ follows a chi-squared distribution

c) $l^t(\Sigma_{i=1}^{n_1} X_i X_i^t) l$ and $l^t(\Sigma_{i=n_1+1}^{n} X_i X_i^t) l$ are independently distributed for $1 \leq n_1 \leq n-1$

d) $A = \Sigma$

38. Arrival of customers in a shop is a Poisson process with intensity $\lambda = 2$. Let X be the number of customers entering during the time interval (1, 2) and let Y be the number of customers entering during the time interval (5,10). Which of the following are true?

a) $P(X = 0|(X + Y = 12)) = (\frac{5}{6})^{12}$

b) X and Y are in-dependant

c) $X + Y$ is a Poisson with parameter 6

d) $X - Y$ is a Poisson with parameter 8

39. Let the joint distribution of (X,Y) be bivariate normal with the mean vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and variance-covariance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, where -1 $< \rho <$ 1. Let $\phi_\rho(0,0) = P(X \leq 0, Y \leq 0)$. Then the Kendall's τ coefficient between X and Y equals

a) $4\phi_\rho(0,0) - 1$

b) $4\phi_\rho(0,0)$

c) $4\phi_\rho(0,0) + 1$

d) $\phi_\rho(0,0)$

40. Let the joint distribution of (X,Y) be bivariate normal with the mean vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and variance-covariance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, where -1 $< \rho <$ 1. Let $\phi_\rho(0,0) = P(X \leq 0, Y \leq 0)$. Then the Kendall's τ coefficient between X and Y equals

a) $4\phi_\rho(0,0) - 1$

b) $4\phi_\rho(0,0)$

c) $4\phi_\rho(0,0) + 1$

d) $\phi_\rho(0,0)$

41. Consider a markov chain with 5 states 1,2,3,4,5 and transition matrix P =

$$\begin{pmatrix} 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 1/7 & 0 & 0 & 6/7 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/3 & 0 & 0 & 2/3 & 0 \\ 0 & 5/8 & 0 & 0 & 3/8 \end{pmatrix} \tag{0.0.2}$$

Which of the following is true?

a) 3 and 1 communicating class

b) 1 and 4 communicating class

c) 4 and 2 communicating class

d) 2 and 5 communicating class