Probability Questions and solutions.

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

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

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.

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

   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

   Solution:
   Definition of Bivariate Gaussian and its independency Bi-variate random variables are distribution of normal distribution to two coordinates. are said to be bivariate normal or jointly normal, if $aX + bY$ has normal distribution $\forall$ $a, b \in R$.
   Random normal vector

$$\mathbf{z} = \begin{pmatrix} X \\ Y \end{pmatrix} \quad (1)$$

   is Bi-variate when it is jointly normal
   Joint PDF of Z is given as

$$f_z(Z) = \frac{1}{2\pi\sqrt{det\Sigma}} \quad \exp\left\{\frac{-1}{2}(z-m)^T\Sigma^{-1}(z-m)\right\} \quad (2)$$

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

Where

$$\mathbf{m} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \tag{3}$$

$$\Sigma = [[\mathbf{Z} - \mathbf{E}(\mathbf{Z})][\mathbf{Z} - \mathbf{E}(\mathbf{Z})]^{\mathbf{T}}] \tag{4}$$

$$\Sigma = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_Y\sigma_X & \sigma_Y^2 \end{bmatrix} \tag{5}$$

If X, Y, which are independent, then they are un-correlated or their co-variances are $\rho\sigma_Y\sigma_X = 0$ then co-variance matrix becomes a diagonal matrix

$$\Sigma = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_Y\sigma_X & \sigma_Y^2 \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix} \tag{6}$$

0.1 Co-variance Matrix

$$\Sigma = [[\mathbf{Z} - E(\mathbf{Z})][\mathbf{Z} - E(\mathbf{Z})]^T] \tag{7}$$

When $\mathbf{Z}' = \mathbf{TZ}$, Where T is a transformation

$$\Sigma_{TZ} = [[\mathbf{TZ} - E(\mathbf{TZ})][\mathbf{TZ} - E(\mathbf{TZ})]^T] \tag{8}$$
$$= [[T\mathbf{Z} - TE(\mathbf{Z})][T\mathbf{Z} - TE(\mathbf{Z})]^T] \tag{9}$$
$$= [T[\mathbf{Z} - E(\mathbf{Z})][\mathbf{Z} - E(\mathbf{Z})]^T T^T] \tag{10}$$
$$= [T[\mathbf{Z} - E(\mathbf{Z})][\mathbf{Z} - E(\mathbf{Z})]^T T^T] \tag{11}$$
$$= [T\Sigma T^T] \tag{12}$$

Where , $[[\mathbf{Z} - E(\mathbf{Z})][\mathbf{Z} - E(\mathbf{Z})]^T] = \Sigma$

$$\Sigma_{TZ} = [T\Sigma T^T] \tag{13}$$

0.2 Few observations:

a) **T** is diagonal matrix
b) σ_{TZ} is diagonal matrix.
c) σ is symmetric.
d) coloums of **T** are eigen vectors.

Now if condition 1 and 4 satisfies then we can say a transformation is independent.

0.3 Evaluating option 1

: Given $\sigma_x = \sigma_y$

$$\Sigma = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_X \\ \rho\sigma_X\sigma_X & \sigma_Y^2 \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X^2 \\ \rho\sigma_X^2 & \sigma_X^2 \end{bmatrix} = \sigma_X^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

Σ is not a diagonal matrix so components of Z in option 1 are not independent.

0.4 Evaluating option 2

X, X-Y can be written as $\begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$

Where T = $\begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$, Z = $\begin{bmatrix} X \\ Y \end{bmatrix}$

Co-variance matrix Σ for X, X-Y From Eq. 1 $\Sigma_{TZ} = [T\Sigma T^T]$

$$\Sigma_{TZ} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \sigma_X^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$$

7. (52) Suppose X∼ Cauchy(0,1). Then the distribution of $\frac{1-X}{1+X}$ is? **Solution**
A continuous random variable X follows **Cauchy distribution** with parameters μ and λ if its pdf is given by

$$f(x) = \begin{cases} \frac{\lambda}{\pi} \cdot \frac{1}{\lambda^2 + (x-\mu)^2}, & -\infty < x < \infty; \\ & -\infty < \mu < \infty,\ \lambda > 0; \\ 0, & \text{Otherwise.} \end{cases}$$

The parameter μ and λ are location and scale parameters respectively.
When μ=0 and λ=1, then the distribution is called **Standard Cauchy Distribution**. The pdf of standard Cauchy distribution is

$$f(x) = \begin{cases} \frac{1}{\pi} \cdot \frac{1}{1+x^2}, & -\infty < x < \infty; \\ 0, & \text{Otherwise.} \end{cases}$$

Let,

$$Y = \frac{1-X}{1+X} \tag{14}$$

Then, cdf of Y is

$$F_Y(y) = P(Y \leq y) \tag{15}$$

$$= P\left(\frac{1-X}{1+X} \le y\right) \tag{16}$$
$$= P\left((1-X) \le y.(1+X)\right) \tag{17}$$
$$= P\left((1-X) \le (y+y.X)\right) \tag{18}$$
$$= P\left((1-y) \le (X+y.X)\right) \tag{19}$$
$$= P\left(\frac{1-y}{1+y} \le X\right) \tag{20}$$
$$= 1 - P\left(X < \frac{1-y}{1+y}\right) \tag{21}$$
$$= 1 - \int_{-\infty}^{\frac{1-y}{1+y}} f(x)\,dx \tag{22}$$
$$= 1 - \int_{-\infty}^{\frac{1-y}{1+y}} \frac{1}{\pi}\cdot\frac{1}{1+x^2}\,dx \tag{23}$$
$$= 1 - \frac{1}{\pi}\cdot\left[tan^{-1}x\right]_{-\infty}^{\frac{1-y}{1+y}} \tag{24}$$
$$= 1 - \frac{1}{\pi}\cdot\left[\left[tan^{-1}x\right]_{-\infty}^{0} + \left[tan^{-1}x\right]_{0}^{\frac{1-y}{1+y}}\right] \tag{25}$$
$$= 1 - \frac{1}{\pi}\cdot\left[-\frac{\pi}{2} + \left[tan^{-1}x\right]_{0}^{\frac{1-y}{1+y}}\right] \tag{26}$$
$$\tag{27}$$

$$F_Y(y) = 1 - \frac{1}{\pi}\cdot\left[-\frac{\pi}{2} + tan^{-1}\left(\frac{1-y}{1+y}\right)\right] \tag{28}$$

The pdf of Y is

$$f_Y(y) = \frac{dF_Y(y)}{dy} = -\frac{1}{\pi}\cdot\frac{d\left(tan^{-1}\left(\frac{1-y}{1+y}\right)\right)}{dy} \tag{29}$$
$$= -\frac{1}{\pi}\cdot\frac{1}{1+\left(\frac{1-y}{1+y}\right)^2}\cdot\frac{d\left(\frac{1-y}{1+y}\right)}{dy} \tag{30}$$
$$= -\frac{1}{\pi}\cdot\frac{1}{1+\left(\frac{1-y}{1+y}\right)^2}\cdot\left(-\frac{2}{(y+1)^2}\right) \tag{31}$$
$$= \frac{1}{\pi}\cdot\frac{1}{1+y^2} \tag{32}$$

Hence, Y ~ Cauchy(0,1).

8. Question 8

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.

1. 12	2. 24
3. 26	4. 27

Solution

a) Maximum likelihood estimation: The goal of maximum likelihood estimation is to make inferences about the population that is most likely to have generated the sample, specifically the joint probability distribution of the random variables

b) : Associated with each probability distribution is a unique vector $\theta = [\theta_1, \theta_2, \ldots, \theta_k]^\mathsf{T}$ of parameters that index the probability distribution within a parametric family $\{f(\cdot\,;\theta) \mid \theta \in \Theta\}$, where Θ is called the parameter space, a finite-dimensional subset of Euclidean space.

Evaluating the joint density at the observed data sample $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ gives a real-valued function,

$L_n(\theta) = L_n(\theta; \mathbf{y}) = f_n(\mathbf{y}; \theta)$

c) : which is called the likelihood function. For independent and identically distributed random variables, $f_n(\mathbf{y}; \theta)$ will be the product of univariate density functions.

The goal of maximum likelihood estimation is to find the values of the model parameters that maximize the likelihood function over the parameter space, that is

$\hat{\theta} = \arg\max_{\theta\in\Theta} \widehat{L}_n(\theta\,;\mathbf{y})$

d) : Intuitively, this selects the parameter values that make the observed data most probable. The specific value $\hat{\theta} = \hat{\theta}_n(\mathbf{y}) \in \Theta$ that maximizes the likelihood function L_n is called the maximum likelihood estimate

e) Solution: Here, p.d.f of distribution is

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

, so for data point $x_i = 12$, p.d.f. $f_\theta(x_i = 12)$

$$= \begin{cases} \frac{1+12^2}{3\theta(1+\theta^2)}, & -2\theta \le x \le \theta, \theta > 0 \\ 0, & otherwise \end{cases}$$

$$= \begin{cases} \frac{145}{3\theta(1+\theta^2)}, & -2\theta \leq x \leq \theta, \theta > 0 \\ 0, & otherwise \end{cases}$$

for every x_i we can can calculate p.d.f in the same way.
Now, for all random variables joint p.d.f. will be

$$f_\theta(x_1, x_2, ..., x_7|\theta) = \begin{cases} \frac{1+x_1}{3\theta(1+\theta^2)}, & -2\theta \leq x \leq \theta, \theta > 0 \\ 0, & otherwise \end{cases} * \begin{cases} \frac{1+x_2}{3\theta(1+\theta^2)}, & -2\theta \leq x \leq \theta, \theta > 0 \\ 0, & otherwise \end{cases} ....$$

Where, $x_1 = 12, X_2 = -54$ and so on.
Here from looking at the options given in question, 27(option 4) is the correct answer because it's the only option that satisfy condition $-2\theta \leq x \leq \theta, \theta > 0$. In all other cases we'll get joint p.d.f. 0.

f) : **Hence answer is 27(option 4)**