Q1.1

Given X and Y as two random variables sampled uniformly and at random from the interval [0,1]

On any sampling instance, the leftmost point will be the minimum of the two

Consider $Z = \min(X,Y)$, as the leftmost point

→ We need to find the expection of Z.

We know,

CDF of a random variable sampled uniformly at random from [0,1] is given by,

$$F_X(x) = P(X \leq x) \text{ [CDF general definition]}$$

In our case

$$F_X(x) = F_Y(y) = \begin{cases} 0 & x \leq 0 \\ (x/1) = x & 0 < x \leq 1 \\ 1 & x \geq 1 \end{cases}$$

↳ (trivially true)

Now for Z,

$$F_Z(z) = P(Z \leq z)$$

$$= 1 - P(Z > z)$$

$$= 1 - P(X > z \text{ and } Y > z)$$

(because even if one is less than z the min would be less than z but we want the probability that min is greater than z)

$$= 1 - P(X>z)P(Y>z) \quad \{X \text{ and } Y \text{ are independent}\}$$

$$= 1 - [1 - P(X \leq z)][1 - P(Y \leq z)]$$

$$= 1 - [1 - F_X(z)][1 - F_Y(z)]$$

$$= 1 - [1 - z]^2$$

$$= 2z - z^2$$

We know,

$f_Z(z)$ = PDF of $Z = \frac{d}{dz} F_Z(z) = \frac{d}{dz} 2z - z^2 = 2 - 2z$

Also

$E(Z) =$ $p_1 z_1 + p_2 z_2 + \ldots p_n z_n$ ~~$p_1 z_1 + p_2 z_2 + \ldots p_n z_n$~~ if Z was discrete

Here Z is continuous, hence

$$E(Z) = \int_0^1 z f_Z(z)\, dz$$

$$= \int_0^1 z(2-2z)\, dz$$

$$= \int_0^1 2z - 2z^2\, dz$$

$$= \left[z^2 - \frac{2z^3}{3} \right]_0^1$$

$$= 1 - \frac{2}{3} = \frac{1}{3} = 0.33$$

<u>Q1.2</u> Given

$$p(r) = k r^{(d-1)} e^{\left(-\frac{r^2}{2}\right)}$$

To find r, such that p(r) is maximum,

Hence solving for $\frac{dp(r)}{dr} = 0$

$$\frac{d\,p(r)}{dr} = K(d-1) r^{(d-2)} e^{\left(-\frac{r^2}{2}\right)} + K r^{(d-1)} e^{\left(-\frac{r^2}{2}\right)} \cdot -r = 0$$

$$(d-1) r^{d-2} - r^d = 0$$

$$r^2 = (d-1) \Rightarrow r = \sqrt{d-1}$$

$$r = \text{sqrt}(d-1)$$

Hence when $r = \sqrt{d-1}$, marginal density or p(r) will be maximized.

Q 1.3

Given,

X and Y as the random variables denoting age & height respectively.

Given a random sample of size 20 of both X & Y

let $X = [69, 74, 68 \text{ - - - - }, 76]^T$

$Y = [153, 175, 155 \text{ - - - - }, 220]^T$

For representation

$X = [x_1, x_2 \text{ - - - } x_{20}]^T$

$Y = [y_1, y_2 \text{ - - - } y_{20}]^T$

Mean of $X = \frac{1}{n} \sum_{i=1}^{n} x_i$

$\bar{X} = 71.45$

Median ~~Mode~~ of X = after sorting in increasing order then (10th position number + 11th position) no ~~by~~ divided by 2

$= \frac{71+72}{2} = 71.5$

Mode of X = highest frequency observation

= 74 (3 times)

Variance of $Y = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{Y})^2$

$= 1441.27$

Given sample mean $\bar{X} = 71.45$

Sample variance $\sigma_X^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2$

$= 14.57$

Assuming X follows a normal distribution with mean $= \bar{X} = \mu$

& Variance $= \sigma_X^2$

$X \sim N(71.45, 14.57)$

$P(X > 80)$

$$z = \frac{X-\mu}{\sigma} = \frac{80 - 71.45}{\sqrt{14.57}} = 2.23$$

$P(X>80) = P(Z>2.23) = 1 - P(Z \le 2.23)$

Now following the standard normal distribution table

$\cancel{P(Z \le 2.23) = (0.4788 + 0.5) = (0.9788)}$

$\cancel{1 - 0.9788 = 0.}$

$P(Z \le 2.23) = (0.4871 + 0.5) = 0.9871$

$P(X>80) = 1 - P(Z \le 2.23) = 0.0129$

→ 2 dimension mean $= [\bar{X} \ \bar{Y}] = [71.45 \ 164.7]$

→ 2 dimensional covariance matrix $= \begin{pmatrix} \sigma(x,x) & \sigma(x,y) \\ \sigma(y,x) & \sigma(y,y) \end{pmatrix}$

$$\sigma(x,y) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

Hence $\sigma(x,x) = \sigma_x^2 = 14.57$

$\sigma(y,y) = \sigma_y^2 = 1441.27$

$\sigma(x,y) = 128.87$

Hence Covar matrix $= \begin{pmatrix} 14.57 & 128.87 \\ 128.87 & 1441.27 \end{pmatrix}$