Homework 1

George Duncan CX 4240

May 28, 2019

Problem 1) $\Sigma = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}$

We must have Σ positive definite. By the Cauchy-Schwartz inequality:
$\text{cov}(X_1, X_2)^2 \leq \text{cov}(X_1, X_1)\text{cov}(X_2, X_2)$
$\text{cov}(X_1, X_2)^2 \leq \text{var}(X_1)\text{var}(X_2)$
$r^2 \leq 1$ ... $r \in [-1, 1]$
For both $r = -1, r = 1$, the determinant of the matrix is 0, therefore the matrix is degenerate. Thus: $r \in (-1, 1)$. We can use the following expansion for the joint probability density function:

$$\Pr(X = x) = p(X) = \frac{1}{2\pi} \det(\Sigma)^{-1} \exp(-q(X - \mu))$$

, where q is defined by the following:

$$q(Z) = \frac{\frac{z_1^2}{\Sigma_{1,1}^2} - \frac{2\sqrt{1-\det(\Sigma)}z_1 z_2}{\text{tr}(\Sigma)} + \frac{z_2^2}{\Sigma_{2,2}^2}}{2\det(\Sigma)}$$

Substituting in the values form this problem:

$$p(X) = \frac{1}{2\pi(1 - r^2)} \exp(-\frac{(x_1 - \mu_1)^2 - 2r(x_1 - \mu_1)(x_2 - \mu_2) + (x_2 - \mu_2)^2}{2(1 - r^2)})$$

For the eigen values we must set the determinant of the matrix equal to 0:
$\det(\Sigma) = 0 = (1 - \lambda)^2 - r^2$, thus $\lambda = 1 \pm r$. Consider the positive case λ_1, and the negative case λ_2. For each λ we can find the corresponding eigen vectors accordingly:
$(\Sigma - \lambda I)v = 0$
$\begin{bmatrix} -r & r \\ r & -r \end{bmatrix} v_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, thus $v_1 = c(1, 1) : c \in \mathbb{R}$, after normalization: $v_1 = (\frac{\sqrt{(2)}}{2}, \frac{\sqrt{(2)}}{2})$
$\begin{bmatrix} r & r \\ r & r \end{bmatrix} v_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, thus $v_2 = c(1, -1) : c \in \mathbb{R}$, after normalization: $v_2 = (\frac{\sqrt{(2)}}{2}, \frac{-\sqrt{(2)}}{2})$

Problem 2) Suppose X is a random variable with the following density:

$$f_X(x) = \begin{cases} 0.5 & x = -1, 1 \\ 0 & x \neq -1, 1 \end{cases}$$

The expected value of X is given by $\mathbb{E}(X) = \sum\limits_{x_i \in \Omega(X)} x_i p(x_i) = 0.5(-1) + 0.5(1) = 0$

Suppose Y is a random variable with the standard normal distribution, with density $f_Y(y)$. Let $Z = XY$, we have Z is dependent on Y as the value of Z is restricted to $\pm Y$, from its original domain of $\mathbb{R}$

$p_Z(z) = \sum\limits_{x_i \in \Omega(X)} p_X(x_i) p_Y(\frac{z}{x_i}) = \frac{1}{2} p_Y(z) + \frac{1}{2} p_Y(-z)$. The standard normal distribution is symmetrical therefore: $= \frac{1}{2} p_Y(z) + \frac{1}{2} p_Y(z) = p_Y(z) = p_Z(z)$, the two variables have identical densities, therefore Z is a standard normal distribution: $\mathcal{N}(0, 1)$

$$\begin{aligned}
\operatorname{cov}(Y, Z) &= \mathbb{E}(YZ) - \mathbb{E}(Y)\mathbb{E}(Z) && (1) \\
&= \mathbb{E}(YZ) - (0)(0) && (2) \\
&= \mathbb{E}(YZ) && (3) \\
&= \mathbb{E}(XY^2) && (4) \\
&= \sum_{x_i \in \Omega(X)} \int\limits_{y \in \Omega(X)} p_{X,Y}(x_i, y) x y^2 dy && (5) \\
&= \sum_{x_i \in \Omega(X)} \int\limits_{y \in \Omega(X)} p_X(x_i) p_Y(y) x y^2 dy && (6) \\
&= \frac{-1}{2} \int\limits_{y \in \Omega(X)} p_Y(y) y^2 dy + \frac{1}{2} \int\limits_{y \in \Omega(X)} p_Y(y) y^2 dy && (7) \\
\operatorname{cov}(Y, Z) &= 0 && (8)
\end{aligned}$$

Y and Z are uncorrelated.

Problem 3.1) $p_1 = \begin{cases} 2\theta & x = 1 \\ 1 - 2\theta & x = 0 \end{cases}$ $p_2 = \begin{cases} \theta & x = 1 \\ 1 - \theta & x = 0 \end{cases}$
$p_{t1} = p_1(0) = 1 - 2\theta$ and $p_{t2} = p_2(0) = 1 - \theta$ We have for probability $p \geq 0$, thus $\theta \in [0, \frac{1}{2}]$
$l = p_1(1)p_2(1)p_1(0)p_2(0)p_1(0)p_2(1) = 4\theta^3(1 - 4\theta + 5\theta^2 - 2\theta^3)$

$$\text{mle}(\theta) = \theta : l(\theta) = \max(l(\theta) : \theta \in \{\theta : \frac{dl}{d\theta} = 0, 0, \frac{1}{2}\} \tag{9}$$

$$\frac{dl}{d\theta} = (12\theta^2 - 64\theta^3 + 100\theta^4 - 48\theta^5) = -4\theta^2(\theta - 1)(3\theta - 1)(4\theta - 3) \tag{10}$$

$$\text{mle}(\theta) = \theta : l(\theta) = \max(l(\theta) : \theta \in \{0, \frac{1}{3}, \frac{1}{2}, 1\}, \text{ restricting to the domain of } \theta : \tag{11}$$

$$= \theta : l(\theta) = \max(l(\theta) : \theta \in \{0, \frac{1}{3}, \frac{1}{2}\} \tag{12}$$

$$l(0) = 0, l(\frac{1}{3}) = \frac{16}{3^6}, l(\frac{1}{2}) = 0 \tag{13}$$

$$\text{mle}(\theta) = \frac{1}{3} \tag{14}$$

Problem 3.2) Suppose $\theta < \max(x_1, ..., x_n)$ then $\exists i \in \{1, 2, ..., n\} \;:\; x_i > \theta$, therefore $f(x_i) = 0$, therefore $l(\theta) = 0$.
Suppose $\theta > \max(x_1, ..., x_n)$ and $\theta' = \max(x_1, ..., x_n)$. $\forall i \in \{1, 2, ..., n\}$ $f(x_i) = \frac{1}{\theta}$ thus $l(\theta) = \frac{1}{\theta^n}$ and $l(\theta') = \frac{1}{(\theta')^n}$, because $l'(\theta) < 0 \forall \theta$: we have $\theta' < \theta \implies l(\theta') > l(\theta)$, thus $\max(l(\theta)) = l(\max(x_1, ..., x_n))$, therefore $\text{mle}(\theta) = \max(x_1, ..., x_n)$

Problem 4) (Note: logs are all in base 2) $p_X(x) = \sum_{y_i \in \Omega(Y)} p_{X,Y}(x, y_i) = \begin{cases} \frac{1}{2} & x = 0 \\ \frac{1}{2} & x = 1 \end{cases}$

$H(X) = \sum_{x_i \in \Omega(X)} -\log(p_X(x_i))$

$H(X) = -\log(\frac{1}{2}) = 1$

$H(X|Y) = \sum_{(x_i,y_i) \in (\Omega(X),\Omega(Y))} p_{X,Y}(x_i, y_i)(-\log(p_{X|Y}(x_i|y_i)))$

$H(X|Y) = \frac{1}{4}(-\log(\frac{1}{3})) + \frac{1}{2}(-\log(\frac{2}{3})) + \frac{1}{4}(-\log(1))$

$...H(X|Y) = \frac{3\log(3)-2}{4}$

$H(Y|X) = \sum_{(x_i,y_i) \in (\Omega(X),\Omega(Y))} p_{X,Y}(x_i, y_i)(-\log(p_{Y|X}(y_i|x_i)))$

$H(Y|X) = \frac{1}{4}(-\log(\frac{1}{2})) + \frac{1}{4}(-\log(\frac{1}{2})) + \frac{1}{2}(-\log(1))$

$...H(Y|X) = \frac{1}{2}$

$I(X, Y) = \sum_{(x_i,y_i) \in (\Omega(X),\Omega(Y))} p_{X,Y}(x_i, y_i) \log(\frac{p_{X,Y}(x_i,y_i))}{p_X(x_i)p_Y(y_i)}$

$I(X, Y) = \sum_{(x_i,y_i) \in (\Omega(X),\Omega(Y))} p_{X,Y}(x_i, y_i) \log(\frac{p_{X,Y}(x_i,y_i))}{p_X(x_i)p_Y(y_i)}$

$I(X, Y) = \frac{1}{4}\log(\frac{2}{3}) + \frac{1}{4}\log(2) + \frac{1}{2}\log(\frac{4}{3})$

$...I(X, Y) = \frac{3}{2} - \frac{3}{4}\log(3)$

$H(X, Y) = H(X) + H(Y|X) = 1 + \frac{1}{2} =$

$...H(X, Y) = \frac{3}{2}$

Theorem .. Show X, Y are independent $\implies H(X|Y) = H(X)$

Proof. Suppose X, Y are independent, then:

$H(X|Y) = \sum_{(x_i,y_i) \in (\Omega(X),\Omega(Y))} p_{X,Y}(x_i, y_i)(-\log(p_{X|Y}(x_i|y_i)))$

By independence we have $p_{X|Y} = p_X$, and $p_{X,Y} = p_X p_Y$, thus:

$H(X|Y) = \sum_{(x_i,y_i) \in (\Omega(X),\Omega(Y))} p_{X,Y}(x_i, y_i)(-\log(p_X(x_i)))$

$H(X|Y) = \mathbb{E}(-\log(p_X)) = H(X)$ □

The same holds for $H(Y|X)$

Theorem .. Show X, Y are independent $\implies H(X, Y) = H(X) + H(Y)$

Proof. Suppose X, Y are independent, then:

$H(X, Y) = H(X) + H(Y|X) = H(X) + H(Y)$, as follows from the above proof □

Theorem .. Show $I(X; X) = H(X)$:

Proof. $I(X; X) = \sum_{x_i \in \Omega(X)} p_X(x_i) \log(\frac{p_X(x_i)}{p_X(x_i)p_X(x_i)})$

$I(X; X) = \sum_{x_i \in \Omega(X)} p_X(x_i) \log(\frac{1}{p_X(x_i)})$

$I(X; X) = \sum_{x_i \in \Omega(X)} p_X(x_i)(-\log(p_X(x_i)))$

$I(X; X) = H(X)$

□