Question 6

RVs X and Y have ranges $S_X = S_Y = \{-3, -1, 1, 3\}$ with joint pmf

$$f(x, y) = \begin{cases} \frac{1}{8} & x \in \{-3, -1\}, y \in \{-3, -1\} \\ \frac{1}{8} & x \in \{1, 3\}, y \in \{1, 3\} \end{cases} \tag{1}$$

Find the correlation coefficient of X and Y. Start by justifying that X and Y are uniform RVs. Then justify that $E(X) = E(Y) = 0$.

Solution

The joint probability mass function of (X,Y) have 8 possible outcomes given in the following pairs (-3,-3), (-3,-1), (-1,-3), (-1,-1), (3,3), (3,1), (1,3) and (1,1) all with equal probability of $\frac{1}{8}$. Thus, X and Y follow a uniform RV.

The correlation coefficient of X and Y is found using

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

$E(X) = \sum_x x\, g(x) = -3(\frac{1}{4}) - 1(\frac{1}{4}) + 1(\frac{1}{4}) + 3(\frac{1}{4}) = 0.$

$E(Y) = \sum_y y\, h(y) = -3(\frac{1}{4}) - 1(\frac{1}{4}) + 1(\frac{1}{4}) + 3(\frac{1}{4}) = 0.$

$\sigma_{XY} = E(XY) - E(X)E(Y) = \sum_x \sum_y xyf(x, y)$

$= -3(-3)(\frac{1}{8}) + -1(-3)(\frac{1}{8}) + -3(-1)(\frac{1}{8}) + -1(-1)(\frac{1}{8}) + 1(1)(\frac{1}{8}) + 3(1)(\frac{1}{8}) + 1(3)(\frac{1}{8}) + 3(3)(\frac{1}{8}) = 4$

$\sigma_X^2 = \sum_x (x - E(X))^2 g(x) = (-3)^2(\frac{1}{4}) + (-1)^2(\frac{1}{4}) + (1)^2(\frac{1}{4}) + (3)^2(\frac{1}{4}) = 5$

$\sigma_Y^2 = \sum_y (y - E(Y))^2 h(y) = (-3)^2(\frac{1}{4}) + (-1)^2(\frac{1}{4}) + (1)^2(\frac{1}{4}) + (3)^2(\frac{1}{4}) = 5$

Finally,

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{4}{\sqrt{25}} = 0.8.$$

Question 7

This question is related to the Poisson RV.

(a) Show that a Poisson RV, X, with mean α has MGF $M_X(t) = e^{\alpha(e^t - 1)}$.

Solution

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} \frac{e^{tx} e^{-\alpha} \alpha^x}{x!} = e^{-\alpha} \sum_{x=0}^{\infty} \frac{(\alpha e^t)^x}{x!} = e^{-\alpha} e^{\alpha e^t} = e^{\alpha(e^t - 1)}.$$

(b) X_1 and X_2 are two independent Poisson RVs with means α_1 and α_2 respectively. What is $h(y)$, the pmf of $Y = X_1 + X_2$.

Solution

Since we are dealing with independent Poisson RVs, then there is a simple way to find the MGF of $Y = X_1 + X_2$, where we just multiply the separate, individual MGFs of X_1 and X_2.

$$M_Y(t) = E(e^{tY}) = E(e^{t(X_1+X_2)}) = E(e^{tX_1})E(e^{tX_2}) = e^{\alpha_1(e^t-1)}e^{\alpha_2(e^t-1)} = e^{(\alpha_1+\alpha_1)(e^t-1)}.$$

Thus, we conclude that $X_1 + X_2 \sim Poisson(\alpha_1 + \alpha_2)$.

Question 8
X_1 and X_2 are i.i.d. zero-mean Gaussian RVs with variance σ^2. Define $Y_1 = X_1^2 + X_2^2$ and $Y_2 = \tan^{-1}(X_2/X_1)$. Find, $h(y_1, y_2)$, the joint pdf of Y_1 and Y_2.

Solution
$X_1 = \sqrt{Y_1}\cos(Y_2)$
$X_2 = \sqrt{Y_1}\sin(Y_2)$
To solve the question, we need to use the following transformation: $h(y_1, y_2) = f(x_1, x_2)|J|$
To find $|J|$, we proceed as

$$|J| = \begin{vmatrix} \frac{1}{2}Y_1^{-1/2}\cos(Y_2) & -\sqrt{Y_1}\sin(Y_2) \\ \frac{1}{2}Y_1^{-1/2}\sin(Y_2) & \sqrt{Y_1}\cos(Y_2) \end{vmatrix} = |\frac{1}{2}(\cos^2(Y_2) + \sin^2(Y_2))| = |\frac{1}{2}(1)| = \frac{1}{2}.$$

X_1 and X_2 are iid with zero mean and variance σ^2.
$h(y_1, y_2) = f(\sqrt{Y_1}\cos(Y_2), \sqrt{Y_1}\sin(Y_2)) * \frac{1}{2} =$

$$\frac{1}{2}\left[\frac{1}{2\pi\sqrt{\sigma^4}}\exp\left\{\frac{-1}{2}\left[\left(\frac{\sqrt{Y_1}\cos(Y_2)}{\sigma}\right)^2 + \left(\frac{\sqrt{Y_1}\sin(Y_2)}{\sigma}\right)^2\right]\right\}\right] =$$

$$\frac{1}{2}\left[\frac{1}{2\pi\sqrt{\sigma^4}}\exp\left\{\frac{-1}{2}\left[\left(\frac{(Y_1\cos(Y_2))^2}{\sigma^2}\right) + \left(\frac{(Y_1\sin(Y_2))^2}{\sigma^2}\right)\right]\right\}\right] =$$

$$\frac{1}{2}\left[\frac{1}{2\pi\sqrt{\sigma^4}}\exp\left\{\frac{-1}{2}\left[\left(\frac{(Y_1(\cos^2(Y_2) + \sin^2(Y_2))}{\sigma^2}\right)\right\}\right] =$$

$$\frac{1}{4\pi\sqrt{\sigma^4}}\exp\left\{\left(\frac{-Y_1}{2\sigma^2}\right)\right\}.$$

Question 9 (a) Show that the joint pdf of all the samples is given by

$$f(\mathbf{x}; A) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\sum_{i=1}^{n}\frac{(x_i - A\alpha^i)^2}{2\sigma^2}\right);$$

where $\mathbf{x} = [x_1, x_2, ..., x_n]$ denotes the vector covering the variables x_i.

Solution

Since we are dealing with independent RVs, we have

$$P((X_1 = x_1, X_2 = x_2, ..., X_n = x_n); A) =$$

$$P((X_1 = x_1); A)P((X_2 = x_2); A)...P((X_n = x_n); A) =$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_1 - A\alpha)^2}{2\sigma^2}\right) * \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_2 - A\alpha^2)^2}{2\sigma^2}\right) * ... * \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_n - A\alpha^n)^2}{2\sigma^2}\right) =$$

$$\prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \exp\left(-\frac{(x_i - A\alpha^i)^2}{2\sigma^2}\right) =$$

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \prod_{i=1}^{n} \exp\left(-\frac{(x_i - A\alpha^i)^2}{2\sigma^2}\right) =$$

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\sum_{i=1}^{n} \frac{(x_i - A\alpha^i)^2}{2\sigma^2}\right).$$

(b) Show that this pdf satisfies the regularity condition.

Solution
The pdf satisfies the regularity condition if

$$E\left[\frac{\partial}{\partial A} \ln f(\mathbf{x}; A)\right] = 0$$

In our case, it yields

$$E\left[\frac{\partial \ln(f(\mathbf{x}; A))}{\partial A}\right] = E\left[\frac{-2}{2\sigma^2} \sum_{i=1}^{n} (x_i - A\alpha^i)(-\alpha^i)\right] = \frac{-2}{2\sigma^2} \sum_{i=1}^{n} (E\left[x_i\right] - A\alpha^i)(-\alpha^i) = \frac{-2}{2\sigma^2} \sum_{i=1}^{n} (A\alpha^i - A\alpha^i)(-\alpha^i) = 0.$$

Thus, the pdf satisfies the regularity condition.

(c) Show that an efficient estimator exists for A.

Solution

$$f(\mathbf{x}; A) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - A\alpha^i)^2\right)$$

$$\frac{\partial \ln(f(\mathbf{x}; A))}{\partial A} = \frac{-2}{2\sigma^2} \sum_{i=1}^{n} (x_i - A\alpha^i)(-\alpha^i)$$

$$= \frac{1}{\sigma^2} (\sum_{i=1}^{n} x_i\alpha^i - A \sum_{i=1}^{n} \alpha^{2i})$$

$$= \frac{1}{\sigma^2} \sum_{i=1}^{n} \alpha^{2i} \left(\frac{\sum_{i=1}^{n} x_i\alpha^i}{\sum_{i=1}^{n} \alpha^{2i}} - A\right)$$

An efficient estimator is found that attains the bounds for all A since $\frac{\partial ln(f(\mathbf{x};A))}{\partial A} = I(A)(g(x) - A)$, where $I(A) = \frac{1}{\sigma^2}(\sum_{i=1}^{n}(\alpha^{2i})$.

The MVU estimator is $\hat{A} = g(x) = \frac{\sum_{i=1}^{n}(x_i)(\alpha^i)}{\sum_{i=1}^{n}\alpha^{2i}}$ and the minimum variance is $\frac{1}{I(A)}$.

(d) Find the variance of this estimator.

Solution

From part (c), the minimum variance is

$$\frac{1}{I(A)} = \frac{1}{\frac{1}{\sigma^2}\sum_{i=1}^{n}\alpha^{2i}} = \frac{\sigma^2}{\sum_{i=1}^{n}\alpha^{2i}}.$$

Question 10
(a) What is the significance of this test? Justify any approximations you make.

Solution
$\mu = np_c = 25$; $\sigma^2 = np_c(1 - p_c) = 18.75$.
Using Gaussian approximation, $X_H - X_C$ is zero mean ($\mu_{X_H} - \mu_{X_C} = 0$) with variance $\sigma^2 = \sigma^2_{X_H} + \sigma^2_{X_C} = 37.5$.
The significance of the test is found by computing $\alpha = p(X_H - X_C \geq 10 | p_c = 0.25)$. Knowing that

$$Z = \frac{(X + 0.5 - \mu)}{\sigma} = \frac{10.5}{\sqrt{37.5}} = 1.71.$$

$\alpha = p(Z > 1.71) = 1 - p(Z < 1.71) = 1 - 0.9564 = 0.0436$.

The results indicate a small Type I error.

(b) On running a trial with $n = 100$ patients in each group, we get $X_C = 25$ and $X_H = 33$. What is the corresponding p-value?

Solution
We now have $X_H - X_C = 8$. The value of Z is

$$\frac{8 + 0.5}{\sqrt{37.5}} = 1.39. \tag{3}$$

This yields $p(X > 8|H_0) = p(Z > 1.39) = 1 - p(Z < 1.39) = 1 - 0.9117 = 0.0823$. Hence, the p-value is 0.0823, which is the minimum value of α that makes X in the critical region.