ECE 353 Probability and Random Signals - Homework 9

Spring 2019
Instructor: Dr. Raviv Raich
School of Electrical Engineering and Computer Science
Oregon State Univeristy

Due: Jun. 6, 2019

Q1. For a random variable X, let $Y = aX + b$. Show that if $a > 0$ then $\rho_{X,Y} = 1$. Also show that if $a < 0$, then $\rho_{X,Y} = -1$.

Solution

First, we observe that Y has mean $\mu_Y = a\mu_X + b$ and variance $\text{Var}[Y] = a^2\text{Var}[X]$. The covariance of X and Y is

$$\text{Cov}[X, Y] = E[(X - \mu_X)(aX + b - a\mu_X - b)] = aE[X - \mu_X]^2 = a\text{Var}[X]$$

The correlation coefficient is

$$\rho_{X,Y} = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X]}\sqrt{\text{Var}[Y]}} = \frac{a\text{Var}[X]}{\sqrt{\text{Var}[X]}\sqrt{a^2\text{Var}[X]}} = \frac{a}{|a|}$$

So when $a > 0$, $\rho_{X,Y} = 1$ and when $a < 0$, $\rho_{X,Y} = -1$.

Q2. Let the random variable X which can be represented as a sum of random variables $X = \sum_{i=1}^{n} X_i$. Show that, if X_i and Y_i are uncorrelated, i.e. $\text{Cov}(X_i, Y_i) = 0$ or $E[X_iX_j] = E[X_i]E[X_j]$ for every pair of i and j with $1 \le i < j \le n$, then $\text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i]$.

Solution

From the definition of variance, we write:

$$
\begin{aligned}
\mathrm{Var}[X] &= E\left[\left(\sum_{i=1}^{n} X_i - E[\sum_{i=1}^{n} X_i]\right)^2\right] \\
&= E\left[\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} E[X_i]\right)^2\right] \\
&= E\left[\left(\sum_{i=1}^{n} (X_i - E[X_i])\right)^2\right] \\
&= E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} (X_i - E[X_i])(X_j - E[X_j])\right] \\
&= \left[\sum_{i=1}^{n}\sum_{j=1}^{n} E(X_i - E[X_i])(X_j - E[X_j])\right] \\
&= E\left[\sum_{i=1}^{n} (X_i - E[X_i])^2 + 2\sum_{i<j} (X_i - E[X_i])(X_j - E[X_j])\right] \\
&= \sum_{i=1}^{n} \mathrm{Var}[X_i] + 2\sum_{i<j} E\left[X_iX_j - E[X_i]X_j - X_iE[X_j] + E[X_i]E[X_j]\right] \\
&= \sum_{i=1}^{n} \mathrm{Var}[X_i]
\end{aligned}
$$

where we have $E\left[X_iX_j - E[X_i]X_j - X_iE[X_j] + E[X_i]E[X_j]\right] = 2E[X_i]E[X_j] - 2E[X_i]E[X_j] = 0$

Q3. Random variables X and Y have joint PDF:

$$
f_{X,Y}(x,y) = \begin{cases} cxy^2 & 0 \le x \le 1, \quad 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}
$$

(a) Find the constant c.

(b) Find $P[X > Y]$ and $P[Y < X^2]$.

(c) Find $P[\min(X,Y) \le 1/2]$.

Solution

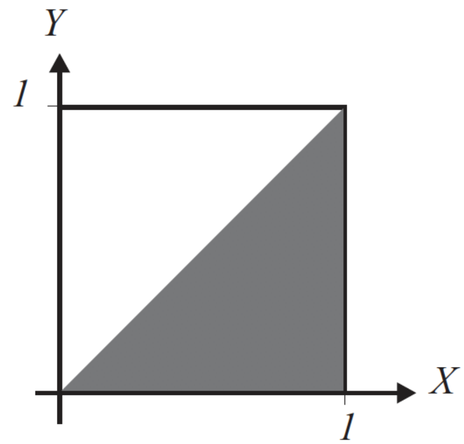

Figure 1: region of $P[X \geq Y]$

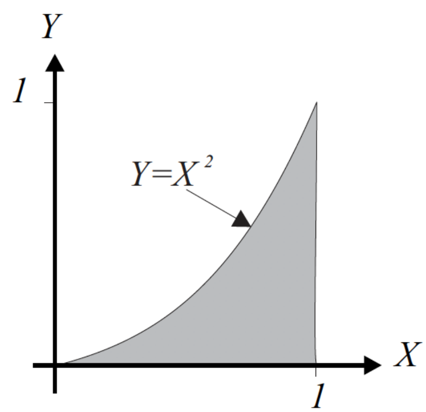

Figure 2: region of $P[Y \leq X^2]$

(a) To find the constant c integrate $f_{X,Y}(x,y)$ over the all possible values of X and Y to get

$$1 = \int_0^1 \int_0^1 cxy^2 dxdy = c/6 \quad (1)$$

Therefore $c = 6$.

(b) The probability $P[X \geq Y]$ is the integral of the joint PDF $f_{X,Y}(x,y)$ over the indicated shaded region.

$$P[X \geq Y] = \int_0^1 \int_0^x 6xy^2 dxdy = \int_0^1 6x \int_0^x y^2 dydx = \int_0^1 2x^4 dx = 2/5 \quad (2)$$

Similarly, to find $P[Y \leq X^2]$ we can integrate over the region shown in the figure.

$$P[Y \geq X^2] = \int_0^1 \int_0^{x^2} 6xy^2 dxdy = \int_0^1 6x \int_0^{x^2} y^2 dydx = \int_0^1 2x^7 dx = 1/4 \quad (3)$$

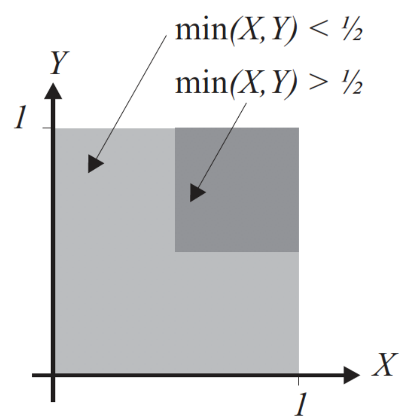

Figure 3: region of $PP[\min(X,Y) \leq 1/2]$

(c) Here we can choose to either integrate $f_{X,Y}(x,y)$ over the lighter shaded region, which would require the evaluation of two integrals, or we can perform one integral over the darker region by recognizing

$$P[\min(X,Y) \leq 1/2] = 1 - P[\min(X,Y) > 1/2] \tag{4}$$
$$= 1 - P[X > 1/2, Y > 1/2] \tag{5}$$
$$= 1 - \int_{1/2}^{1}\int_{1/2}^{1} 6xy^2 dxdy = 11/32 \tag{6}$$

Q4. Let$X$ and Y be two independent normal random variables $N(0,1)$. Show that $aX + bY$ and $bX - aY$ are independent random variables.

Solution

Because X and Y are independent, we have:

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}} = \frac{1}{2\pi}e^{-\frac{\left\|\binom{x}{y}\right\|^2}{2}} \tag{7}$$

we also have $Z = aX + bY$ and $W = bX - aY$:

$$\begin{pmatrix} Z \\ W \end{pmatrix} = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}\begin{pmatrix} X \\ Y \end{pmatrix} \tag{8}$$

so we obtain:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{\begin{bmatrix} -a & -b \\ -b & a \end{bmatrix} \binom{Z}{W}}{a^2 + b^2} \tag{9}$$

hence

$$f_{Z,W}(z,w) = f_{X,Y}(x(z,w), y(w,z))\,|\mathbf{J}| \tag{10}$$

$$|\mathbf{J}| = \left| \begin{bmatrix} \frac{\partial x}{\partial w} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial w} & \frac{\partial y}{\partial z} \end{bmatrix} \right| = \left| \begin{bmatrix} \frac{a}{a^2+b^2} & \frac{b}{a^2+b^2} \\ \frac{b}{a^2+b^2} & \frac{-a}{a^2+b^2} \end{bmatrix} \right| = \frac{1}{a^2 + b^2} \tag{11}$$

$$\begin{aligned} f_{Z,W}(z,w) &= f_{X,Y}(x(z,w), y(w,z)) && (12) \\ &= \frac{1}{2\pi} e^{-\frac{\left\|\binom{x}{y}\right\|^2}{2}} = \frac{1}{2\pi} e^{-\frac{\left\|\mathbf{A}\binom{z}{w}\right\|^2}{2}} && (13) \\ &= \frac{1}{2\pi} e^{-\frac{(z,w)\mathbf{A}^T\mathbf{A}\binom{z}{w}}{2}} && (14) \end{aligned}$$

we also have

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \begin{bmatrix} a & b \\ b & -a \end{bmatrix} = \mathbf{I}/(a^2 + b^2) \tag{15}$$

hence

$$f_{Z,W}(z,w) = \frac{1}{2\pi(a^2 + b^2)} e^{-\frac{z^2+w^2}{2(a^2+b^2)}} \tag{16}$$

so we conclude: $\binom{Z}{W} \sim N(\mathbf{0}, \mathbf{I}/(a^2 + b^2))$, so they are independent.

Q5. Consider the two independent RVs $X \sim U[-1,1]$ and $Y \sim U[-1,1]$. Let $Z = X^2 Y$.

(a) Find the mean of Z, $E[Z]$.

(b) Find $\mathrm{Corr}(X,Z)$ and $\mathrm{Corr}(Y,Z)$.

(c) Determine if Z and Y are uncorrelated.

Solution

(a)

$$E[Z] = E[X^2Y] = E[X^2]E[Y] = 0 \tag{17}$$

(b)

$$\text{Corr}(X, Z) = E[X, Z] = E[X^3Y] = E[X^3]E[Y] = 0 \tag{18}$$

$$\text{Corr}(Y, Z) = E[Y, Z] = E[X^2Y^2] = E[X^2]E[Y^2] = \frac{1}{3} \times \frac{1}{3} = 0.111 \tag{19}$$

(c)

$$E[Y]E[Z] = 0 \tag{20}$$

Hence we have $\text{Cov}(Y, Z) = E[Y, Z] = 1/9 \neq 0$, so Z and Y are not uncorrelated.

Q6. You are interested in estimating the probability of A, denoted as $P[A]$, where

$$A = \{\text{a student in EECS at OSU wants to go to graduate school}\}.$$

You start asking your fellow students "do you want to go to graduate school?" so that you can make an estimation. Suppose every fellow student who you've asked answers your question with 'yes' or 'no'. Assume that the students' answers are independent and identically distributed. Denote $\hat{P}_n(A) = \frac{\text{number of 'yes'}}{n}$ when you have answers of n fellow students. If you wish to make

$$|\hat{P}_n(A) - P[A]| < 0.1$$

with a confidence level 0.95, how many fellow students do you need to ask, in the worst case?

Solution

Let us define a random variable X_i as

$$X_i = \begin{cases} 1, & \text{answer 'yes'} \\ 0, & \text{answer 'no'} \end{cases}$$

$$\hat{P}_n(A) = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{21}$$

Apply Chebyshevs inequality

$$P[|\hat{P}_n(A) - P[A]| < 0.1] > 1 - \frac{Var(X_i)}{nc^2} \tag{22}$$

we also have

$$Var(X_i) = P[A](1 - P[A]) \tag{23}$$

While $P[A] = 1/2$ which has the largest $P[A](1 - P[A])$, hence we have

$$P[|\hat{P}_n(A) - P[A]| < 0.1] > 1 - \frac{1}{4nc^2} \tag{24}$$

while $c = 0.1$ and confidence level 0.95, we need $n \geq 500$.