AI5002 - Assignment 14

Tuhin Dutta
ai21mtech02002

Download code and LaTeX from below hyper-links
1. LaTeX

Problem NET_JUNE_2012_Q104

Which of the following conditions imply independence of the random variables X and Y ?

1) $\Pr(X > a|Y > a) = \Pr(X > a)\ \forall\ a \in \mathbb{R}$
2) $\Pr(X > a|Y < b) = \Pr(X > a)\ \forall\ a,\ b \in \mathbb{R}$
3) X and Y are uncorrelated.
4) $E[(X-a)(Y-b)] = E(X-a)\ E(Y-b)\ \forall\ a,\ b \in \mathbb{R}$

Solution

We analyze the options one by one and see which option best implies that random variables X and Y are independent.

1) Let's assume continuous r.v.s X and Y are not independent and,

$$\begin{aligned} X &\in \{0, 1\} \\ Y &= X + 2 \\ \Pr(X = 0) = \Pr(X = 1) &= \frac{1}{2} \end{aligned} \tag{1}$$

Now since Y is always greater than X therefore $\Pr(X > a\ |Y > a)$ equals $\Pr(X > a)\ \ \forall\ a\ \ \in \mathbb{R}$ and thus in spite of the option (1) being true, it fails to imply that X and Y are independent random variables and hence option (1) is false.

2) Let us denote the cumulative distribution functions of X, Y and (X, Y), as below,

$$\begin{aligned} F_X(a) &= \Pr(X \leq a), \\ F_Y(b) &= \Pr(Y \leq b)\ \ and \\ F_{X,Y}(a, b) &= \Pr(X \leq a, Y \leq b) \end{aligned} \tag{2}$$

To show independence, we want to prove that,

$$F_X(a)F_Y(b) = F_{X,Y}(a, b)\ \forall\ a,\ b\ \ \in \mathbb{R} \tag{3}$$

Conditional probability tells us that:

$$\Pr(X > a|Y < b) = \frac{\Pr(X > a, Y < b)}{\Pr(Y < b)} \tag{4}$$

and so by the assumptions of the option (1),

$$\begin{aligned} \Pr(X > a) &= \frac{\Pr(X > a, Y < b)}{\Pr(Y < b)} \\ \Pr(X > a)\Pr(Y < b) &= \Pr(X > a, Y < b) \end{aligned} \tag{5}$$

Now since

$$\begin{aligned} &\Pr(X > a) = 1 - F_X(a)\ \ and, \\ &\Pr(Y < b) = F_Y(b) \end{aligned} \tag{6}$$

We may rewrite the above equation as:

$$\begin{aligned} F_Y(b) - F_X(a)F_Y(b) &= \Pr(X > a, Y < b) \\ F_X(a)F_Y(b) &= F_Y(b) - \Pr(X > a, Y < b) \end{aligned} \tag{7}$$

Note that since X is continuous so we can write,

$$P(X \leq a) = P(X < a)$$

Regardless of the value of X the marginal C.D.F F_Y(b) is given by

$$F_Y(b) = \Pr(Y < b)$$

Now let us define two events

$$Event\ A : (Y < b \cap X < a)$$

$$Event\ B : (Y < b \cap X > a)$$

We can also think of the event (Y < b) as

$$(Y < b) = (\text{Event A}) \cup (\text{Event B})$$

So

$$\Pr(A, B) = \Pr(Y < b)$$

Since X cannot both be less than a and greater than a, we have

$$\begin{aligned} \Pr(A, B) &= \Pr(A) + \Pr(B) \\ &= \Pr(Y < b, X < a) + \\ &\quad \Pr(Y < b, X > a) \end{aligned} \tag{8}$$

$\therefore$ We can write

$$\begin{aligned} F_Y(b) = \Pr(X > a, Y < b) + \\ \Pr(X < a, Y < b) \end{aligned} \tag{9}$$

Thus putting value of $F_Y(b)$ from (9) into (7) proves (2) ,

$$F_X(a)F_Y(b) = \Pr(X < a \text{ and } Y < b) \tag{10}$$

Thus option (2) seems to be always true.

3) Given random variables X and Y are uncorrelated which means that their correlation is 0, or, equivalently, Cov(X, Y) = 0.

$$\begin{aligned} Cov(X, Y) &= E[XY] - E[X]E[Y] \\ &[\because Cov(X, Y) = 0] \\ E[XY] &= E[X]E[Y] \end{aligned} \tag{11}$$

We have to prove that uncorrelated implies independence.
Let's take X and Y to exist as an ordered pair at the points (-1,1), (0,0), and (1,1) with probabilities $\frac{1}{4}, \frac{1}{2}$, and $\frac{1}{4}$. The expected values of X and Y is

$$\begin{aligned} E[X] = -1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0 = E[Y] \\ E[XY] = -1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0 = E[X]E[Y] \end{aligned} \tag{11}$$

Now let's look at the marginal distributions of X and Y. X and Y both take on the values $-1, 0, 1$ and the probability it takes for each of those are given by $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$. Then looping through the possibilities, we have to check if

$$\Pr(X = x, Y = y) = \Pr(X = x)\Pr(Y = y)$$

Let's take the first point (-1, 1) and examine,

$$\begin{aligned} \Pr(X = -1, Y = 1) = \frac{1}{4} \\ \neq \frac{1}{16} = \Pr(X = -1)\Pr(Y = 1) \end{aligned} \tag{12}$$

We loop through the other two points, and see that X and Y do not meet the definition of independent.
Thus it proves that uncorrelated random variables are not always independent. Hence option (3) is false.

4) We extend L.H.S and R.H.S to compare,

$$\begin{aligned} E[(X - a)(Y - b)] = E[XY] \\ -aE[Y] - bE[X] + ab \end{aligned} \tag{13}$$

Also,

$$\begin{aligned} E(X - a)E(Y - b) = E[X]E[Y] \\ -aE[Y] - bE[X] + ab \end{aligned} \tag{14}$$

We see (13) = (14), i.e. independent iff $E[XY] = E[X]E[Y]$ but in this case independence of X and Y cannot be inferred. Hence option (4) is false.