AI5002 - Assignment 14

Tuhin Dutta
ai21mtech02002

Download code and LaTeX from below hyperlinks
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 us take a counter example to understand if option 1) being true implies X and Y to be independent random variables.

$$\begin{aligned} &X \sim \mathcal{N}(0,1) \\ &Y = 10X \; and \; Y \sim \mathcal{N}(0,100) \\ &a \in \mathbb{R}\{-10,-1\} \\ &\therefore \; \Pr(Y > a) = 1 \end{aligned} \tag{1}$$

Now irrespective of the fact that X and Y are dependent, we can write

$$\begin{aligned} \Pr(X > a|Y > a) &= \frac{\Pr(X > a, Y > a)}{\Pr(Y > a)} \\ &= \Pr(X > a) \end{aligned} \tag{2}$$

Although the condition in option 1) holds in our example but here X and Y are dependent random variables. Hence option 1) does not imply X and Y to be independent.

2) Let us denote the individual C.D.F.s of the continuous random variables X, Y and the joint C.D.F (X, Y), as below,

$$\begin{aligned} F_X(a) &= \Pr(X \le a) . = \Pr(X < a), \\ F_Y(b) &= \Pr(Y \le b) = \Pr(Y < b) \; and \\ F_{X,Y}(a,b) &= \Pr(X \le a, Y \le b) \\ &= \Pr(X < a, Y < b) \end{aligned} \tag{3}$$

To show independence, we want to prove that,

$$F_X(a)F_Y(b) = F_{X,Y}(a,b) \; \forall \; a, \; b \; \in \mathbb{R} \tag{4}$$

From conditional probability we know:

$$\Pr(X > a|Y < b) = \frac{\Pr(X > a, Y < b)}{\Pr(Y < b)} \tag{5}$$

and so using the given condition in option 2), we can write (4) as

$$\begin{aligned} \Pr(X > a) &= \frac{\Pr(X > a, Y < b)}{\Pr(Y < b)} \\ \implies \Pr(X > a)\Pr(Y < b) &= \Pr(X > a, Y < b) \end{aligned} \tag{6}$$

We can write the C.D.F as

$$\begin{aligned} &\Pr(X > a) = 1 - F_X(a) \; and, \\ &\Pr(Y < b) = F_Y(b) \end{aligned} \tag{7}$$

We may rewrite (5) using (6) as:

$$\begin{aligned} (1 - F_X(a))(F_Y(b)) &= \Pr(X > a, Y < b) \\ F_Y(b) - F_X(a)F_Y(b) &= \Pr(X > a, Y < b) \\ F_X(a)F_Y(b) &= F_Y(b) - \Pr(X > a, Y < b) \end{aligned} \tag{8}$$

Note that since X is continuous so we can write,

$$\Pr(X \le a) = \Pr(X < a) \tag{9}$$

Regardless of the value of X the marginal

C.D.F F_Y(b) is given by

$$F_Y(b) = \Pr(Y < b) \tag{10}$$

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

$$(Y < b) = (\text{Event A}) \cup (\text{Event B}) \tag{11}$$

So it implies

$$\Pr(A, B) = \Pr(Y < b) \tag{12}$$

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

$$\begin{aligned} \Pr(A, B) &= \Pr(A) + \Pr(B) \\ &= \Pr(Y < b, X < a) + \\ &\quad \Pr(Y < b, X > a) \end{aligned} \tag{13}$$

∴ We can write

$$\begin{aligned} F_Y(b) &= \Pr(X > a, Y < b) + \\ &\quad \Pr(X < a, Y < b) \end{aligned} \tag{14}$$

Now putting value of $F_Y(b)$ from (14) into (8) proves (4) ,

$$\begin{aligned} F_X(a)F_Y(b) &= \Pr(X < a, Y < b) \\ &= F_{X,Y}(a, b) \end{aligned} \tag{15}$$

Thus (15) shows that the joint C.D.F. is the product of the two individual C.D.F. Hence using the given the condition in option 2) we have proved that X and Y to be independent random variables.

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

$$\begin{aligned} Cov(X, Y) &= E(XY) - E(X)E(Y) \\ &[\because Cov(X, Y) = 0] \\ E(XY) &= E(X)E(Y) \end{aligned} \tag{16}$$

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

$$\begin{aligned} E[X] &= -1\left(\frac{1}{4}\right) + 0\left(\frac{1}{2}\right) + 1\left(\frac{1}{4}\right) = 0 = E[Y] \\ E[XY] &= -1\left(\frac{1}{4}\right) + 0\left(\frac{1}{2}\right) + 1\left(\frac{1}{4}\right) = 0 = E[X]E[Y] \end{aligned} \tag{17}$$

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

$$\Pr(X = x, Y = y) = \Pr(X = x)\Pr(Y = y) \tag{18}$$

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

$$\begin{aligned} \Pr(X = -1, Y = 1) &= \frac{1}{4} \\ \neq \frac{1}{16} &= \Pr(X = -1)\Pr(Y = 1) \end{aligned} \tag{19}$$

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

4) We extend L.H.S

$$\begin{aligned} E[(X - a)(Y - b)] &= E[XY] \\ &\quad -aE[Y] - bE[X] + ab \end{aligned} \tag{20}$$

and R.H.S to compare,

$$\begin{aligned} E(X - a)E(Y - b) &= E[X]E[Y] \\ &\quad -aE[Y] - bE[X] + ab \end{aligned} \tag{21}$$

We see (20) = (21), i.e. independent iff $E[XY] = E[X]E[Y]$ but in this case independence of X and Y cannot be inferred.

Let us take a counter example to understand further as to why this condition is not always true.

$$
\begin{aligned}
&Let\ X \sim \mathcal{N}(0,1)\\
&\quad Y = X^2\\
&\quad \text{Multiply X on both sides,}\\
&\quad XY = X^3\\
&\\
&\quad \text{Let us calculate the expected values,}\\
&\quad E(XY) = E(X^3) = E(X)E(Y)\\
&\quad \text{[By property of expectations]}\\
&\\
&\quad E(XY) = E(X^3) = 0.E(Y) = 0\\
&\quad [\because\ E(X) = 0]\\
&\\
&\therefore E(XY) = E(X)E(Y)
\end{aligned} \tag{22}
$$

From (22) we see they are uncorrelated but not independent because

a) We have a case when $X \in (0,1)$ and at the same time $Y = X^2 > 1$. This will never happen, because if $X^2 > 1$ then $X > 1$. So the probability is 0.

b) $\Pr(0 < X < 1)$
$= F_X(1) - F_X(0)$
$= 0.84134475 - 0.5$
$= 0.34134475$

c) $P(Y > 1)$
$= P(X^2 > 1)$
$= P(X > 1) + P(X < -1)$
$= 2 * (1 - P(X \leq 1))$
$= 2 * (1 - F_X(1))$
$= 2 * (1 - 0.84134475) = 0.31731$

Hence we can write,

$$
\begin{aligned}
\Pr(0 < X < 1, Y > 1) = 0 \neq\\
\Pr(0 < X < 1)\Pr(Y > 1)
\end{aligned} \tag{23}
$$

Thus condition in option 4) fail to imply X and Y are independent random variables.