Ch4 - Probability Theory - 3

(Short Chapter on Joint Distributions)

Statistics For Business and Economics - I

Shaikh Tanvir Hossain

East West University, Dhaka

Last Updated December 24, 2023

Outline

Outline

- In this chapter we will see a short overview of concepts related to joint distribution. The idea of joint probability distribution is just an application of the idea of joint probability table. Then automatically you can think about marginal probabilities and conditional probabilities. Almost all ideas are conceptually same as the ideas we saw in Chapter-2, but now this will be for multiple random variables.
- So let's start... 🚶 🚶 🚶 .

1. Introduction and Examples

2. Joint, Marginal and Conditional Distributions

- Joint PMF or PDF
- Marginal PMFs, Marginal Expectation and Marginal Variance
- Conditional PMFs, Conditional Expectation and Conditional Variance
- Independence of two discrete random variables
- Idea of Covariance for two ranadom variables

Introduction and Examples

Multiple Random Variables and Joint distributions

- So far we talked about a single random variable (whether it was discrete or continuous) and its distributions. So this a *univariate* setup, where we try to understand only one random variable and its distribution.
- But in statistics we can also think about more than a single random variable, we call this a *multivariate* setup.
- Multivariate setup captures how different random variables interact with each other.
- The most important object in this case is their *joint distribution*, which has the probabilities when multiple random variables take different values together.
- We will introduce multivariate analogs of the PMF, PDF and also CDF. Like the univariate case these objects will help us to think about the probabilities for multiple random variables.
- Three key concepts of this section are - *joint distribution, marginal distributions* and *conditional distributions*.
- If you have understood joint probability table, conditional probabilities and marginal probabilities*, this will be very easy for you. But if you are struggling with these concepts you should go back to Chapter-2, read and understand before you continue.
- Here are some real life examples,
  - 1. Suppose that we choose a *random family* of Bangladesh, and we would like to study following things,
    - The number of people in the family (this is a random variable X_1).
    - The household income of this family (this is a random variable X_2).
    - The household expenditure of this family (this is a random variable X_3).

    X_1, X_2 and X_3 are all random variables, and it is natural that they are all related to each other.

Multiple Random Variables and Joint distributions

  - 2. Two or more stock prices - Maybe we can represent a stock price of a company with X_1, and a stock price of another company with X_2, and often different stock prices move up and down together!
  - We will see more examples in the coming section.
- Next we will start talking about Joint Distributions. We will start from looking into only two random variables, X and Y and their distributions together. This is called *Bivariate setup*.
- Bivariate distributions are easy to understand and visualize, and this helps to understand situations when we have more than two random variables.

*This is explained in Chapter-2 and also look at PS-2

1. Introduction and Examples

2. Joint, Marginal and Conditional Distributions

- Joint PMF or PDF
- Marginal PMFs, Marginal Expectation and Marginal Variance
- Conditional PMFs, Conditional Expectation and Conditional Variance
- Independence of two discrete random variables
- Idea of Covariance for two ranadom variables

Joint Distribution

- Let's start with our familiar example from Chapter-2, recall,

Response	Treatment group				
	Imipramine	Lithium	Combination	Placebo	Total
Relapse	18	13	22	24	77
No relapse	22	25	16	10	73
Total	40	38	38	34	150

- From here we can easily calculate the joint probability table,

Response	Treatment group				
	Imipramine	Lithium	Combination	Placebo	Total
Relapse	0.12	0.08	0.15	0.16	0.51
No relapse	0.15	0.17	0.10	0.07	0.49
Total	0.27	0.25	0.25	0.23	1

- Important: We only took decimal places upto two digits, so there might be some rounding errors.

Joint Distribution

- Let's introduce two random variables, X and Y....Random variables just take things to real numbers.
- X represents treatment group -
  - 1 for Imipramine
  - 2 for Lithium
  - 3 for Combination
  - 4 for Placebo
- Y represents response condition -
  - 1 means patient had a relapse
  - 0 means patient didn't have a relapse
- Now we can write following joint probability distribution

	Treatment group (X)				
Response (Y)	1	2	3	4	Total
1	0.12	0.08	0.15	0.16	0.51
0	0.15	0.17	0.10	0.07	0.49
Total	0.27	0.25	0.25	0.23	1

- Note in this case both of the variables are discrete random variables.
- Here $\mathbb{P}(X = 1, Y = 0) = 0.15$ means if we randomly select one individual from the *population of* 150, then there is a 15% chance that he/she had Imipramine and she didn't have a relapse (can you give a frequency interpretation, think about if you pcik randomly an individual 100 times from this population?)

Joint PMF

- The last example is an example of *joint probability distribution* of two discrete random variables X and Y.
- Note that in this case we have all 6 probabilities,
  - $\mathbb{P}(X = x, Y = y)$, where possible values of x are $1, 2, 3$ and possible values of y are $0, 1$.
  - You can check that these probabilities will sum to 1.
  - This means $\sum_x \sum_y \mathbb{P}(X = x, Y = y) = 1$
- What does double sum mean? Ans: This means we are summing over all x and y values.
- Now since these are discrete probabilities, we can also think about an underlying joint PMF or *joint probability mass function*.
- The idea of the joiny PMF is same as univariate PMF, it's just a PMF for all combinations of values.

Joint PMF

Definition 4.1: (Joint PMF of two random variables)

The joint PMF of discrete random variables X and Y is the function $f(x, y)$ given by

$$f(x, y) = \mathbb{P}(X = x, Y = y).$$

- As you can alreasy guess, for a valid joint PMFs, the values must be nonnegative and they should be summed to 1, so
$$\sum_x \sum_y f(x, y) = 1$$
- The joint PMF determines the distribution (or it is sufficient to know the Joint PMF if we want to know the joint distribution). This is because we can use it to find the probability of the event $(X, Y) \in A$ for any set A in the $x - y$ plane or $2D$ plane. All we have to do is sum the joint PMF over A :
$$\mathbb{P}((X, Y) \in A) = \sum_{(x,y) \in A}\sum \mathbb{P}(X = x, Y = y)$$

Joint PMF

Example 4.2:
The following table gives the joint PMF of the discrete variables X and Y.

	Treatment group (X)				
Response (Y)	1	2	3	4	Total
1	0.12	0.08	0.15	0.16	0.51
0	0.15	0.17	0.10	0.07	0.49
Total	0.27	0.25	0.25	0.23	1

From the table, we can find probabilities of all combination of values of (X, Y). To check you understood the concept, you should find

- $\mathbb{P}(X = 2, Y = 1) =?$
- $\mathbb{P}(X = 2, Y = 0) =?$
- $\mathbb{P}(X > 1, Y > 0) =?$

Joint PMF

Example 4.3:
Here is another example,

		X			
		-2	0	2	3
Y	3	0.27	0.08	0.16	0
	6	0	0.04	0.10	0.35

Try to find

- $\mathbb{P}(X = 2, Y = 3) =?$
- $\mathbb{P}(X = 3, Y = 3) =?$
- $\mathbb{P}(X > 1, Y > 3) =?$

Joint PMF

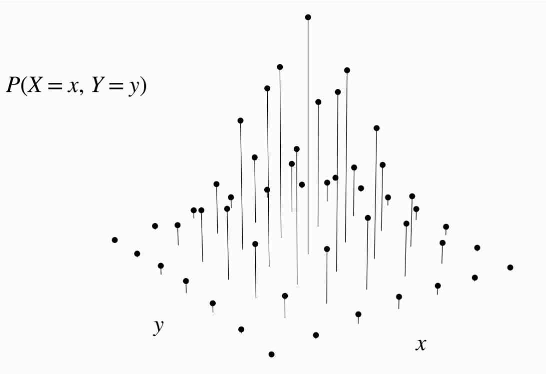


Figure 1: Figure above shows a sketch of what the joint PMF of two discrete random variables could look like. The height of a vertical bar at (x, y) represents the probability $\mathbb{P}(X = x, Y = y)$ or $f(x, y)$. For the joint PMF to be valid, the total height of the vertical bars must be 1 .

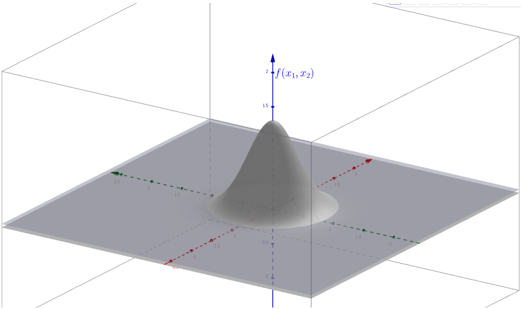
Joint PDF

- What about Joint PDF?
- The idea would extend in a similar way.
- For example if we have two continuous random variables X and Y, then there would be a joint density function $f(x, y)$ as follows and with that we can calculate the probabilities under the curve give a certain region

$$\mathbb{P}((X, Y) \in A) = \int\int_{(x,y)\in A} f(x, y)dxdy$$

- The double integral looks very complicated, but it is just a generalization of the sum for the discrete case.
- Here we are integrating over the region A in the $x - y$ plane, and the integrand is the joint PDF $f(x, y)$.
- Here is an example of bi-variate normal,

Joint PDF



- The functions looks a bit more scary, sorry,

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times$$

$$e^{\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2-2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\}}$$

- Here we have two random variables, X and Y which are jointly normal. Now we have 5 parameters, μ_X, μ_Y, σ_X, σ_Y and ρ.

Joint PDF

- We will not go into details of the joint distribution of two continuous random variables, rather we will understand using discrete random variables ... for joint the ideas extend in a similar way!

Joint, Marginal and Conditional Distributions

Marginal PMFs, Marginal Expectation and Marginal Variance

Notations for Joint, Marginal and Conditional

Be Careful with different notations

- Within few slides we will have lots of PMFs, Expectations and Variance. So we need to carefully denote different objects (I am sorry previosuly the notation was a bit sloppy!)
- The need for new notation is because now we have many objects.
- Throughout this lecture, we will use following notations when we have joint probabilities of two random variables X and Y.

Notations	*What does it mean*
$f(x, y)$	Joint PMF
$f_X(x)$, $\mathbb{E}(X)$ and $\mathrm{Var}(X)$	Marginal PMF of X, Marginal Expectation of X, and Marginal Variance of X
$f_Y(y)$, $\mathbb{E}(Y)$ and $\mathrm{Var}(Y)$	Marginal PMF of Y, Marginal Expectation of Y, and Marginal Variance of Y
$f_{Y\|X}(y\|x)$	Conditional PMF of Y, conditioning on a fixed value $X = x$
$\mathbb{E}(Y\|X = x)$	Conditional Expectation of Y, conditioning on a fixed value $X = x$
$\mathrm{Var}(Y\|X = x)$	Conditional Variance of Y, conditioning on a fixed value $X = x$
$f_{X\|Y}(x\|y)$	Conditional PMF of X, conditioning on a fixed value $Y = y$
$\mathbb{E}(X\|Y = y)$	Conditional Expectation of X, conditioning on a fixed value $Y = y$
$\mathrm{Var}(X\|Y = y)$	Conditional Variance of X, conditioning on a fixed value $Y = y$

- Note that depending on each conditioning value we will have a different PMF, different expectation and different variance, we will see examples of this.

Marginal PMF

- Now let's start talking about marginal objects first.
- Question is - if we have a joint PMF $f(x, y)$ (for example look at Example 3.3), can we get PMF $f_X(x)$ and PMF $f_Y(y)$? which are simply the univariate PMF of X and Y? The answer is YES!
- These are called *Marginal PMFs*.
- For the discrete random variables X, we can obtain $f_X(x)$ by summing the joint PMF $f(x, y)$ over the possible values of Y (similarly we can do it for Y)
- Here is the formal definition, don't worry we will look at some examples now.

Definition 4.4: (Marginal PMF)

For discrete random variables X and Y

- The marginal PMF of X is $f_X(x) = \sum_y f(x, y)$.
- The marginal PMF of Y is $f_X(y) = \sum_x f(x, y)$.

- It is important to note that the marginal PMF of X is the PMF of X, viewing X individually rather than jointly with Y (similarly think about marginal PMF of Y) .
- The operation of summing over the possible values of Y in order to convert the joint PMF into the marginal PMF of X is known as *marginalization of Y* or *marginalizing out Y*.

Marginal PMF

Example 4.5:
Consider the table in Example 3.3. The marginal PDF of X, denoted by $f_X(x)$ is obtained as follows:

$$f_X(-2) = \sum_y f(x,y) = 0.27 + 0 = 0.27$$

$$f_X(0) = \sum_y f(x,y) = 0.08 + 0.04 = 0.12$$

$$f_X(2) = \sum_y f(x,y) = 0.16 + 0.10 = 0.26$$

$$f_X(3) = \sum_y f(x,y) = 0 + 0.35 = 0.35$$

So we get $f_X(x)$

x	-2	0	2	3
$f_X(x)$	0.27	0.12	0.26	0.35

Marginal PMF

Example 4.6:
Likewise, the marginal PDF of Y, denoted by $f_Y(y)$ is obtained as

$$f_Y(3) = \sum_x f(x, y) = 0.27 + 0.08 + 0.16 + 0 = 0.51$$
$$f_Y(6) = \sum_x f(x, y) = 0 + 0.04 + 0.10 + 0.35 = 0.49$$

So we get $f_Y(y)$

y	3	6
$f_Y(y)$	0.51	0.49

Marginal PMF

We can also write the marginal PMFs in following ways

$$f_X(x) = \begin{cases} 0.27 & \text{if } x = -2 \\ 0.12 & \text{if } x = 0 \\ 0.26 & \text{if } x = 2 \\ 0.35 & \text{if } x = 3 \\ 0 & \text{otherwise} \end{cases}$$

Try to write the marginal PMF of Y in this way.

$$f_Y(y) = \begin{cases} 0.51 & \text{if } y = 3 \\ 0.49 & \text{if } y = 6 \\ 0 & \text{otherwise} \end{cases}$$

Note that now you should understand the idea of PMF clearly.

Marginal PMF

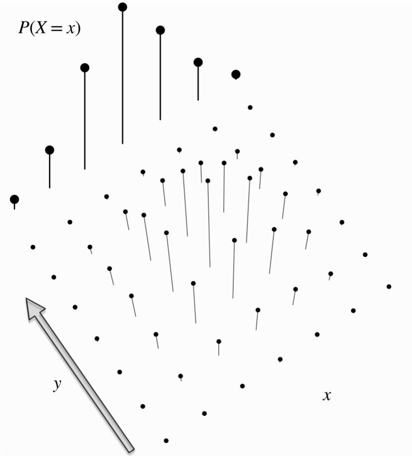


Figure 2: Figure above shows the process of obtaining the marginal PMF from the joint PMF. Here we take a bird's-eye view of the joint PMF for a clearer perspective; each column of the joint PMF corresponds to a fixed x and each row corresponds to a fixed y. For any x, the probability $\mathbb{P}(X = x)$ is the total height of the bars in the corresponding column of the joint PMF: we can imagine taking all the bars in that column and stacking them on top of each other to get the marginal probability. Repeating this for all x, we arrive at the marginal PMF, depicted in bold.

Marginal Expectations and Variance

- So from joint PMF of two random variables X and Y, we can calculate marginal PMFs, $f_X(x)$ and $f_X(y)$ and using these marginal PMFs we can calculate marginal expectations, $\mathbb{E}(X)$ and $\mathbb{E}(Y)$. And also marginal variance $\mathrm{Var}(X)$ and $\mathrm{Var}(Y)$.
- Let's calculate the marginal expectation and variance using the PMF of page 16, 17 and 18.

$$\mathbb{E}(X) = \sum_x x f_X(x) = (-2)(0.27) + (0)(0.12) + (2)(0.26) + (3)(0.35) = 1.03$$

- Similarly,

$$\mathbb{E}(Y) = \sum_y y f_Y(y) = (3)(0.51) + (6)(0.49) = 4.47$$

- Calculate $\mathrm{Var}(X)$ $\mathrm{Var}(Y)$.

Conditional PMF

- Now suppose that we observe the value of X and want to update our distribution of Y to reflect this information. Instead of using the marginal PMF $f_Y(y)$, which does not take into account any information about X, we should use a PMF that conditions on the event $X = x$, where x is the value we observed for X. This naturally leads us to consider conditional PMFs.

Definition 4.7: (Conditional PMFs)

For discrete random variables X and Y,

- The conditional PMF of of Y given $X = x$ is

$$f_{Y|X}(y \mid x) = \frac{f(x, y)}{f_X(x)}$$

- The conditional PMF of of X given $Y = y$ is

$$f_{X|Y}(x \mid y) = \frac{f(x, y)}{f_Y(y)}$$

- Note that depending on conditioning values, we will have different conditional PMFs.
- Note that the conditional PMF of Y (for fixed x) is a valid PMF. So we can define the conditional expectation of Y given $X = x$, denoted by $\mathbb{E}(Y \mid X = x)$, in the same way that we defined $\mathbb{E}(Y)$ except that we replace the PMF of Y with the conditional PMF of Y given $X = x$. Like any PMF, the sum of the conditional PMF of of Y given $X = x$ over all possible values of Y is 1.

Conditional PMF

Example 4.8: Continuing with previous example (Example 3.3), let us compute the conditional PMF of X when $Y = 3$, so we need $f_{X|Y}(x|3)$. This means we need to calculate $f_{X|Y}(-2|3), f_{X|Y}(0|3), f_{X|Y}(2|3)$ and $f_{X|Y}(3|3)$.

$$f_{X|Y}(-2 \mid 3) = \frac{f(-2,3)}{f_Y(3)} = 0.27/0.51 = 0.53$$

We can also calculate

$$f_{X|Y}(0|3), f_{X|Y}(2|3) \text{ and } f_{X|Y}(3|3).$$

- Now because conditional PMF is a valid PMF, we will have $f_{X|Y}(-2 \mid 3) + f_{X|Y}(0|3) + f_{X|Y}(2|3) + f_{X|Y}(3|3) = 1$ (Try this at your home!)
- Notice that the unconditional (or marginal) probability $f_X(-2)$ is 0.27, but if Y has assumed the value of 3 , the probability that X takes the value of -2 is 0.53.

Conditional PMF

Homework: Using Example 3.3., calculate the conditional PMF $f_{Y|X}(y|2)$, this means you need $f_{Y|X}(3|2)$ and $f_{Y|X}(6|2)$. Then calculate conditional expectation $\mathbb{E}(Y|X=2)$, and finally calculate conditional variance $\mathrm{Var}(Y|X=2)$.

Conditional PMF

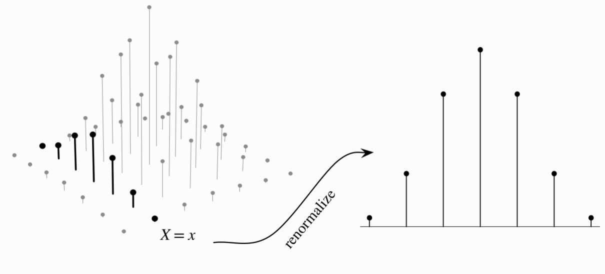


Figure 3: Figure above illustrates the definition of conditional PMF. To condition on the event $X = x$, we first take the joint PMF and focus in on the vertical bars where X takes on the value x; in the figure, these are shown in bold. All of the other vertical bars are irrelevant because they are inconsistent with the knowledge that $X = x$ occurred. Since the total height of the bold bars is the marginal probability $\mathbb{P}(X = x)$, we then renormalize the conditional PMF by dividing by $\mathbb{P}(X = x)$; this ensures that the conditional PMF will sum to 1. Therefore conditional PMFs are PMFs, just as conditional probabilities are probabilities. Notice that there is a different conditional PMF of Y for every possible value of X; Figure above highlights just one of these conditional PMFs.

Independence of discrete random variables

We earlier discussed independence of events. Armed with an understanding of joint, marginal, and conditional distributions, we can revisit the definition of independence for random variables.

Definition 4.9: (Independence of discrete random variables)

If X and Y are discrete random variables, then X and Y are *independent* if for all x and y,

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

for all x and y. Note that, this means

$$f(x, y) = f_X(x) f_Y(y)$$

for all x and y

- The definition says that for independent r.v.s, the joint PMF factors into the product of the marginal PMFs.

Independence of discrete random variables

- Remember! in general, the marginal distributions do not determine the joint distribution.
- This means joint distribution comes first, and this is the entire reason why we wanted to study joint distributions in the first place!
- But in the case of independence, the marginal distributions are all we need to specify the joint distribution.
- Try to construct a joint PMF of X and Y after constructing Marginal PMFs of X and Y such that the two random variables are independnt.
- Another way of looking at independence is that all the conditional PMFs are the same as the marginal PMF (why is that?), becuase if two random variables are independent conditional probabilities are same as marginal probabilities.
- In other words, starting with the marginal PMF of Y, no updating is necessary when we condition on $X = x$, regardless of what x is.

Independence of discrete random variables

Example 4.10: Show that, X and Y variables whose joint given in Example 3.3 are NOT independent (Note: Finding even one pair of values x and y such that $f(x, y) \neq f_X(x)f_X(y)$ is enough to rule out independence.)

Covariance and Correlation

Definition 4.11: (Covariance and Correlation)

The covariance between two random variables X and Y is

$$\text{Cov}(X, Y) = \mathbb{E}\left[\left(X - \mathbb{E}(X)\right)\left(Y - \mathbb{E}(Y)\right)\right] = \mathbb{E}\left[\left(X - \mu_X\right)\left(Y - \mu_Y\right)\right]$$

And the Correlation between two random variables X and Y is

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\left(\sqrt{\text{Var}(X)}\right)\left(\sqrt{\text{Var}(Y)}\right)} = \frac{\text{Cov}(X, Y)}{\sigma_X \times \sigma_Y}$$

- where μ_X and μ_Y are the marginal Expected values of X and Y, and σ_X and σ_Y are the standard deviations of X and Y.
- Notice if the covariance is 0, then correlation is also 0.
- Sometimes $\rho_{X,Y}$ is used for correlation.
- Notice $-1 \leq \text{Cor}(X, Y) \leq 1$, this is the benefit of correlation that it gives a value between -1 and 1, but what does this mean.
- Let's understand Covariance first, then we will understand Correlation.

Covariance and Correlation

- Let's think about the definition intuitively. If X and Y tend to move in the same direction, then $X - \mathbb{E}(X)$ and $Y - \mathbb{E}(Y)$ will tend to be either both positive or both negative, so $(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))$ will be positive on average, giving a positive covariance.
- If X and Y tend to move in opposite directions, then $X - \mathbb{E}(X)$ and $Y - \mathbb{E}(Y)$ will tend to have opposite signs, giving a negative covariance.
- If X and Y are independent, we can show that their covariance is zero (we will see why later!).
- We say that random variables with zero covariance are *uncorrelated*. We will come back to this again after we see an example.
- For the covariance , we have an equivalent expression:
$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Example 4.12:
Let us find out the covariance between discrete random variables X and Y whose joint PMF is as shown in Example 3.3. This is a tedious calculation, but you should use the following formula,

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \mathbb{E}\left[(X - \mathbb{E}(X))\,(Y - \mathbb{E}(Y))\right] \\ &= \sum_x \sum_y \left[(x - \mathbb{E}(X))\,(y - \mathbb{E}(Y))\right] \times f(x, y)\end{aligned}$$

It should be 2.24

Expectation of function of X and Y

The two-dimensional version of LOTUS lets us calculate the expectation of a random variable that is a function of two random variables X and Y, using the joint distribution of X and Y.

Theorem 4.13: (2D LOTUS Discrete)

Let g be a function from $\mathbb{R}^2$ to $\mathbb{R}$. If X and Y are discrete, then

$$\mathbb{E}(g(X,Y)) = \sum\sum g(x,y)f(x,y)$$

Expectation of function of X and Y

Example 4.14: Let us find $\mathbb{E}(X, Y)$ from the joint PMF given in Example 3.3.

$$\begin{aligned}\mathbb{E}(XY) = &\textstyle\sum_x \sum_y xyf(x, y)\\ =&(-2)(3)(0.27) + (0)(3)(0.08) + (2)(3)(0.16) + (3)(3)(0)\\ &+ (-2)(6)(0) + (0)(6)(0.04) + (2)(6)(0.10) + (3)(6)(0.35)\\ =&6.84\end{aligned}$$

We can also verify previous calculation

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\\ &= 6.84 - (1.03)(4.47)\\ &= 2.24\end{aligned}$$

Covariance and Independence

Theorem 4.15:

If X and Y are independent, then they are uncorrelated.

- This can be easily show for two discrete random variables X and Y. If X and Y are independent, their joint PDF is the product of the marginal PDFs $f(x, y) = f_X(x)f_X(y)$. By 2D LOTUS,

$$\mathbb{E}(XY) = \textstyle\sum_x \sum_y xyf_X(x)f_X(y) = \displaystyle\sum_x xf_X(x) \sum_y yf_X(y) = \mathbb{E}(X)\mathbb{E}(Y).$$

  which gives
$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0 - 0 = 0$$

- CAREFUL! The converse of this theorem is false: just because X and Y are uncorrelated does not mean they are independent.
- Covariance is a measure of linear association, so r.v.s can be dependent in nonlinear ways and still have zero covariance. We see an example of that below

Example 4.16: Suppose that the random variable X has probability distribution
$$\mathbb{P}(-1) = 1/4 \quad \mathbb{P}(0) = 1/2 \quad \mathbb{P}(1) = 1/4$$

Let the random variable Y be defined as follows:
$$Y = X^2$$

Thus, knowledge of the value taken by X implies knowledge of the value taken by Y, and, therefore, these two random variables are certainly not independent. Whenever $X = 0$, then $Y = 0$, and if X is either -1 or 1 , then $Y = 1$. The joint probability distribution of X and Y is
$$\mathbb{P}(-1, 1) = 1/4 \quad \mathbb{P}(0, 0) = 1/2 \quad \mathbb{P}(1, 1) = 1/4$$

with the probability of any other combination of values being equal to 0 . It is then straightforward to verify that
$$E[X] = 0 \quad E[Y] = 1/2 \quad E[XY] = 0$$

The covariance between X and Y is 0 . Thus we see that random variables that are not independent can have a covariance equal to 0 .

References