Jointly Distributed Random Variables

Suppose we have two random variables X and Y defined on a sample space S. Then we call (X, Y) a random vector.

Discrete joint distribution: Let X and Y be two discrete random variables. The **joint probability mass function** (also called the joint distribution) of X and Y is

$$p_{XY}(x, y) = P(X = x, Y = y) = P(\{X = x\} \cap \{Y = y\}).$$

The **range** of the random vector (X, Y) is

$$R_{XY} = \{(x, y) \ : \ p_{XY}(x, y) \neq 0\}.$$

Properties of the joint distribution:

1. (non-negative probability)
$$p_{XY}(x, y) \geq 0$$
2. (total mass =1)
$$\sum_{(x,y)\in R_{XY}} p_{XY}(x, y) = 1$$
3. (addition property)
$$P((X, Y) \in A) = \sum_{(x,y)\in A\cap R_{XY}} p_{XY}(x, y)$$

Definition: If X and Y are jointly distributed random variables with joint probability mass function p_{XY}, then the **marginal probability mass functions** of X and Y are respectively

$$p_X(x) = P(X = x) = \sum_y p_{XY}(x, y) \ \text{ and } \ p_Y(y) = P(Y = y) = \sum_x p_{XY}(x, y)$$

Independence: We say that X and Y are **independent** if

$$p_{XY}(x,y) = p_X(x)\,p_Y(y), \quad \text{for all } (x,y) \in R_X \times R_Y,$$

where

$$R_X \times R_Y = \{(x,y)\ :\ p_X(x) \neq 0 \text{ et } p_Y(y) \neq 0\}.$$

Example 26: Consider the following joint probability mass function:

x	y	$p_{XY}(x,y)$
1	1	1/30
1	2	5/30
1	3	6/30
2	1	2/30
2	2	10/30
3	1	6/30

Determine the following probabilities:
(a) $P(X < 2, Y < 2)$

Solution: $P(X < 2, Y < 2) = p_{XY}(1,1) = \frac{1}{30}$
(b) $P(X = 1)$

Solution: $P(X = 1) = p_{XY}(1,1) + p_{XY}(1,2) + p_{XY}(1,3) = \frac{1}{30} + \frac{5}{30} + \frac{6}{30} = \frac{12}{30}$
(c) $P(X > 1, Y \leq 2)$

Solution: $P(X > 1, Y \leq 2) = p_{XY}(2,1) + p_{XY}(2,2) + p_{XY}(3,1) = \frac{2}{30} + \frac{10}{30} + \frac{6}{30} = \frac{18}{30}$
(d) $P(Y > 2|X = 1)$

$P(Y > 2|X = 1) = \frac{P(\{Y>2\}\cap\{X=1\})}{P(X=1)} = \frac{p_{XY}(1,3)}{12/30} = \frac{6/30}{12/30} = \frac{6}{12}$

(e) Determine the marginal distribution of X.

Solution:

$$p_X(x) = \begin{cases} 1/30 + 5/30 + 6/30; & x = 1 \\ 2/30 + 10/30; & x = 2 \\ 6/30; & x = 3 \end{cases}$$

(f) Determine the marginal distribution of Y.

Solution:

$$p_Y(y) = \begin{cases} 1/30 + 2/30 + 6/30; & y = 1 \\ 5/30 + 10/30; & y = 2 \\ 6/30; & y = 3 \end{cases}$$

(g) Are X and Y independent? **Solution:**

$$p_X(3) \times p_Y(3) = \frac{6}{30} \times \frac{6}{30}$$

but

$$P_{XY}(3,3) = 0,$$

so $p_X(3) \times p_Y(3) \neq P_{XY}(3,3)$. Therefore, X and Y are **not** independent.

Expectation of a random vector

Definition: Let (X, Y) be a discrete random vector. The expectation of $h(X, Y)$ is

$$E[h(X,Y)] = \sum_{(x,y)\in R_{XY}} h(x,y)\, p_{XY}(x,y).$$

Interpretation: If we take the average of $h(X, Y)$ over a large number of trials of a random experiment, we expect to get a value close to $E[h(X, Y)]$.

Here are some examples of expectation of a random vector.

1. The expectation of the the sum is equal to the sum of the expectations:

$$E[X + Y] = E[X] + E[Y].$$

Proof:

$$\begin{aligned} E[X+Y] &= \sum_{(x,y)\in R_{XY}} (x+y) p_{XY}(x,y) \\ &= \sum_{(x,y)\in R_{XY}} x\, p_{XY}(x,y) + \sum_{(x,y)\in R_{XY}} x\, p_{XY}(x,y) \\ &= E[X] + E[Y] \end{aligned}$$

2. The **covariance** of X and Y is defined as

$$\begin{aligned} \sigma_{XY} = \mathrm{Cov}[X,Y] &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= \sum_{(x,y)\in R_{XY}} (x-\mu_X)(y-\mu_Y) p_{XY}(x,y) \end{aligned}$$

Remarks :

- There is an alternative formula to compute the covariance that gives the same answer as the above definition:

$$\begin{aligned} \sigma_{XY} = \mathrm{Cov}[X,Y] &= E[X\,Y] - \mu_X\,\mu_Y \\ &= \left(\sum_{(x,y)\in R_{XY}} x\,y\,p_{XY}(x,y) \right) - \mu_X\,\mu_Y \end{aligned}$$

- If X and Y are independent, then $E[XY] = E[X]\,E[Y]$.
- From the previous result we get that

  X and Y are independent $\Longrightarrow \mathrm{Cov}[X,Y] = E[X]E[Y] - \mu_X\,\mu_Y = 0$.

  and so,

  $$\mathrm{Cov}[X,Y] \neq 0 \Longrightarrow X \text{ and } Y \text{ are } \textbf{not} \text{ independant}$$

- $V[X] = \mathrm{Cov}[X,X]$. So we can use the properties of the covariance to study the variance.

Other properties: (These properties are a consequence of the linearity of expectation and the commutativity property of multiplication)

- $\mathrm{Cov}[X,Y] = \mathrm{Cov}[Y,X]$
- $\mathrm{Cov}[a\,X+b, c\,Y+d] = a\,c\,\mathrm{Cov}[X,Y]$
- $\mathrm{Cov}[X+Y,W] = \mathrm{Cov}[X,W] + \mathrm{Cov}[Y,W]$

Theorem:

$$V[X+Y] = V[X] + V[Y] + 2\,\mathrm{Cov}[X,Y]$$

If X and Y are independent, then $V[X+Y] = V[X] + V[Y]$.

Proof:

$$\begin{aligned} V[X+Y] &= \mathrm{Cov}[X+Y,X+Y] \\ &= \mathrm{Cov}[X+Y,X] + \mathrm{Cov}[X+Y,Y] \\ &= \mathrm{Cov}[X,X] + \mathrm{Cov}[Y,X] + \mathrm{Cov}[X,Y] + \mathrm{Cov}[Y,Y] \\ &= V[X] + V[Y] + 2\,\mathrm{Cov}[X,Y] \end{aligned}$$

If X and Y are independant, then $\mathrm{Cov}[X,Y] = 0$ and so $V[X+Y] = V[X] + V[Y]$.

Example 27: Consider the discrete random variables X and Y with the following joint distribution:

x	y	$p_{XY}(x,y)$
1	1	1/30
1	2	5/30
1	3	6/30
2	1	2/30
2	2	10/30
3	1	6/30

The marginal distributions of X and Y are

$$p_X(x) = \begin{cases} 12/30, & x = 1 \\ 12/30, & x = 2 \\ 6/30, & x = 3 \end{cases} \quad \text{et} \quad p_Y(y) = \begin{cases} 9/30, & y = 1 \\ 15/30, & y = 2 \\ 6/30, & y = 3 \end{cases}$$

repsectively. With these marginal distributions, we can compute

$$E[X] = 1.8; \quad E[Y] = 1.9; \quad V[X] = 0.56; \quad V[Y] = 0.49.$$

(a) Compute the expectation of $X+Y$ and the expectation of $X-Y$.
Solution: $E[X+Y] = E[X] + E[Y] = 1.8 + 1.9 = 3.7$

$E[X-Y] = E[X] - E[Y] = 1.8 - 1.9 = -0.1$

(b) Compute the covariance of X and Y.

Solution: $E[XY] = 1 \times 1 \times 1/30 + 1 \times 2 \times 5/30 + 1 \times 3 \times 6/30 + 2 \times 1 \times 2/30 + 2 \times 2 \times 10/30 + 3 \times 1 \times 6/30 = \frac{91}{30}$ $\mathrm{Cov}[X,Y] = E[XY] - E[X]E[Y] = \frac{91}{30} - 1.8 \times 1.9 = -0.387$

(c) Compute the variance of $X+Y$ and the variance of $X-Y$.
Solution:

$V[X+Y] = V[X] + V[Y] + 2\mathrm{Cov}[X,Y] = 0.56 + 0.49 + 2 \times (-0.387) = 0.276$

$$V[X+Y] = V[X]+V[-Y]+2\mathrm{Cov}[X,-Y] = V[X]+(-1)^2\times V[Y]+2\mathrm{Cov}[X,-Y]$$
$$= 0.56 + 0.49 - 2 \times (-0.387) = 1.824$$

Example 28: Consider the following joint distribution:

x	-1	0	0	1
y	0	-1	1	0
$p_{XY}(x,y)$	1/4	1/4	1/4	1/4

Prove that the covariance of X and Y is zero, but that X and Y are nevertheless dependent.
Solution: $\mathrm{E}[X] = -1\times\frac{1}{4}+0+0+1\times\frac{1}{4} = 0$ Also, $E[XY] = 0+0+0+0 = 0$
Therefore, $\mathrm{Cov}[XY] = E[XY] - E[X]E[Y] = 0 - 0 \times E[Y] = 0$.

However, $p_X(1) = \frac{1}{4}$ and $p_Y(1) = \frac{1}{4}$.

But $p_{XY}(1,1) = 0$.

Thereofre, $p_X(1) \times p_X(1) = \frac{1}{4} \times \frac{1}{4} \neq p_{XY}(1,1) = 0$, so no $X$and Y are not independent.

Remarks:

- If we know that two random variables are independent, then their covariance must be zero. However, if instead we know that the covariance is zero this does not neccessarily imply that the random variables are independent.
- Covariance is a measure of a particular kind of dependence. We will discuss this type of dependence later in the context of statistics.
- **Terminology:**
  If $\mathrm{Cov}[X,Y] = 0$, we say that X and Y are **uncorrelated**.