Ferhat Abbas-Sétif 1 University
Faculty of Sciences
Department Informatics

Chapter 3 : Probability distributions for combined random variables : Joint distributions of discrete random variables

1. Joint distributions of discrete random variables

Définition 1. (**joint probability mass function**) Let X and Y, discrete random variables that are defined on the same sample space S, then their joint probability mass function is defined as

$$P_{XY}(x,y) = P(X=x, Y=y) = p(x,y)$$

where (x,y) is a pair of possible values for the pair of random variables (X,Y), and $p(x,y)$ satisfies the following conditions :
- $0 \leq p(x,y) \leq 1$
- $\sum_{x\in A}\sum_{y\in A} p(x,y) = 1$
- $P((X,Y)\in A)) = \sum_{x\in A}\sum_{y\in A} p(x,y)$

Définition 2. (**joint cumulative probability function**) Let X and Y are two discrete random variables that are defined on the same sample space S, then their joint cumulative distribution function is obtained by summing the joint PMF's, i.e :

$$F(x,y) = P(X\leq x, Y\leq y) = \sum_{x_i\leq x}\sum_{y_j\leq y} p\left(x_i, y_j\right)$$

where x_i and y_j denote respectively the possible values of X and Y.

From the above definitions, we can obtain the PMF of X from its joint PMF with Y. Indeed,

$$\begin{aligned} P_X(x) &= P(X=x) \\ &= \sum_{y_j\in S_Y} P(X=x, Y=y_j) \\ &= \sum_{y_j\in S_Y} P_{XY}(x,y_j). \end{aligned}$$

Here, $P_X(x)$ is called the marginal probability mass function of X. Similarly, we can derive the marginal PMF of Y from its joint PMF with X.

Définition 3. Let X and Y are discrete random variable and let $p(x,y)$ their joint PMF. The marginal probability mass functions of X and Y are respectively defined as follows :

$$P_X(x) = \sum_{y_j\in S_Y} p(x,y_j), \qquad \text{for any } x\in S_X$$
$$P_Y(y) = \sum_{x_i\in S_X} p(x_i,y), \qquad \text{for any } y\in S_Y$$

Exemple 1. We toss a fair coin three times and record the sequence of heads (H) and tails (T). Let X the random variable of the number of obtained heads and Y denotes the winnings earned in a single play of a game with the following rules, based on the outcomes of the probability experiment :
- player wins 1$ if first heads occurs on the first toss
- player wins 2$ if first heads occurs on the second toss
- player wins 3$ if first heads occurs on the third toss

TABLE 1: Joint PMF of X and Y

$p(x,y)$	X			
Y	0	1	2	3
-1	1/8	0	0	0
1	0	1/8	2/8	1/8
2	0	1/8	1/8	0
3	0	1/8	0	0

TABLE 2: Marginal PMF's for X and Y

x	$p_X(x)$	y	$p_Y(y)$
0	1/8	-1	1/8
1	3/8	1	1/2
2	3/8	2	1/4
3	1/8	3	1/8

- player loses 1$ if no heads occur

The possible values of X are $S_x = \{0, 1, 2, 3\}$, and the possible values of Y are $S_y = \{-1, 1, 2, 3\}$. We represent the joint PMF in Table 1 below. Given the joint PMF, the marginal PMF's of both X and Y is represented in Table 2, below. Also, the joint CDF for X and Y is represented in Table 1.3 below, by summing over values of the joint frequency function. For example, consider $F(1,1)$:

TABLE 3: Joint CDF of X and Y

$F(x,y)$	X			
Y	0	1	2	3
-1	1/8	1/8	1/8	1/8
1	1/8	1/4	1/2	5/8
2	1/8	3/8	3/4	7/8
3	1/8	1/2	7/8	1

$$F(1,1) = P(X \leq 1 \text{ and } Y \leq 1) = \sum_{x \leq 1} \sum_{y \leq 1} p(x,y) = p(0,-1) + p(0,1) + p(-1,1) + p(1,1) = \frac{1}{4}$$

Définition 4. (Independent Random Variables)

Discrete random variables $X_1, X_2, \ldots, X_n$ are considered independent if the joint probability mass function factors into a product of the marginal PMF's i.e,

$$p\left(x_1, x_2, \ldots, x_n\right) = p_{X_1}\left(x_1\right) \cdot p_{X_2}\left(x_2\right) \cdots p_{X_n}\left(x_n\right)$$

This condition holds also for the cumulative distribution functions.

For example in Table 2 of the above example, in the case where $(x,y) = (0,-1)$, we have

$$p(0,-1) = \frac{1}{8}, p_X(0) = \frac{1}{8}, p_Y(-1) = \frac{1}{8},$$

this mean that $p(0,-1) \neq p_X(0) \cdot p_Y(-1)$, so the variables X and Y are not independent (in other word they are dependent).

1.1. Expectations of joint discrete distributions

Théorème 1. *Let X and Y are discrete random variables and let $p(x, y)$ their joint PMF. If $g(X, Y)$ is a function of these two random variables, then its mean value is given as follows :*

$$\mathrm{E}[g(X,Y)] = \sum_{x\in\S_X}\sum_{y\in\S_Y} g(x,y)p(x,y)$$

Exemple 2. Lets consider Example 1. Find $E(XY), E(X)$ and $E(Y)$.

Solution

1- For the first case, we define $g(x, y) = xy$. Therefore, the expected value of XY is :

$$\begin{aligned}
\mathrm{E}[XY] = \sum_{x\in S_X}\sum_{y\in S_Y} xy\cdot p(x,y) &= (0)(-1)\left(\frac{1}{8}\right) \\
&+ (1)(1)\left(\frac{1}{8}\right) + (2)(1)\left(\frac{2}{8}\right) + (3)(1)\left(\frac{1}{8}\right) \\
&+ (1)(2)\left(\frac{1}{8}\right) + (2)(2)\left(\frac{1}{8}\right) \\
&+ (1)(3)\left(\frac{1}{8}\right) \\
&= \frac{17}{8} = 2.125
\end{aligned}$$

2- For the second case, we define $g(x) = x$, so the expected value of X :

$$\begin{aligned}
\mathrm{E}[X] = \sum_{x\in S_X} x\cdot p(x,y) &= (0)\left(\frac{1}{8}\right) \\
&+ (1)\left(\frac{1}{8}\right) + (2)\left(\frac{2}{8}\right) + (3)\left(\frac{1}{8}\right) \\
&+ (1)\left(\frac{1}{8}\right) + (2)\left(\frac{1}{8}\right) \\
&+ (1)\left(\frac{1}{8}\right) \\
&= \frac{12}{8} = 1.5
\end{aligned}$$

3. For the last case, we define $g(x, y) = y$, and the expected value of Y is :

$$\begin{aligned}
\mathrm{E}[Y] = \sum_{y\in S_Y} y\cdot p(x,y) &= (-1)\left(\frac{1}{8}\right) \\
&+ (1)\left(\frac{1}{8}\right) + (1)\left(\frac{2}{8}\right) + (1)\left(\frac{1}{8}\right) \\
&+ (2)\left(\frac{1}{8}\right) + (2)\left(\frac{1}{8}\right) \\
&+ (3)\left(\frac{1}{8}\right) \\
&= \frac{10}{8} = 1.25
\end{aligned}$$

Théorème 2. *If X and Y are independent random variables, then* $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$.