3.1.4 Independent Random Variables

In real life, we usually need to deal with more than one random variable. For example, if you study physical characteristics of people in a certain area, you might pick a person at random and then look at his/her weight, height, etc. The weight of the randomly chosen person is one random variable, while his/her height is another one. Not only do we need to study each random variable separately, but also we need to consider if there is *dependence* (i.e., correlation) between them. Is it true that a taller person is more likely to be heavier or not? The issues of dependence between several random variables will be studied in detail later on, but here we would like to talk about a special scenario where two random variables are *independent*.

The concept of independent random variables is very similar to independent events. Remember, two events A and B are independent if we have $P(A,B) = P(A)P(B)$ (remember comma means *and*, i.e., $P(A,B) = P(A \text{ and } B) = P(A \cap B)$). Similarly, we have the following definition for independent discrete random variables.

Definition 3.2
Consider two discrete random variables X and Y. We say that X and Y are independent if

$$P\Big(X = x, Y = y\Big) = P(X = x)P(Y = y), \qquad \text{for all } x, y.$$

In general, if two random variables are independent, then you can write

$$P\Big(X \in A, Y \in B\Big) = P(X \in A)P(Y \in B), \qquad \text{for all sets } A \text{ and } B.$$

Intuitively, two random variables X and Y are independent if knowing the value of one of them does not change the probabilities for the other one. In other words, if X and Y are independent, we can write

$$P(Y = y|X = x) = P(Y = y), \text{ for all } x, y.$$

Similar to independent events, it is sometimes easy to argue that two random variables are independent simply because they do not have any physical interactions with each other. Here is a simple example: I toss a coin $2N$ times. Let X be the number of heads that I observe in the first N coin tosses and let Y be the number of heads that I observe in the second N coin tosses. Since X and Y are the result of independent coin tosses, the two random variables X and Y are independent. On the other hand, in other scenarios, it might be more complicated to show whether two random variables are independent.

Example 3.6

I toss a coin twice and define X to be the number of heads I observe. Then, I toss the coin two more times and define Y to be the number of heads that I observe this time. Find $P\bigg((X<2) \textrm{ and } (Y>1)\bigg)$.

Solution

Since X and Y are the result of different independent coin tosses, the two random variables X and Y are independent. Also, note that both random variables have the distribution we found in Example 3.3. We can write

$$
\begin{aligned}
P\bigg((X<2) \textrm{ and } (Y>1)\bigg) & \\
&= \big(P_X(0)+P_X(1)\big)P_Y(2) \\
&= \left(\frac{1}{4}+\frac{1}{2}\right)\frac{1}{4} \\
&= \frac{3}{16}.
\end{aligned}
$$

We can extend the definition of independence to n random variables.

Definition 3.3

Consider n discrete random variables $X_1, X_2, X_3, \ldots, X_n$. We say that $X_1, X_2, X_3, \ldots, X_n$ are independent if

$$P\Big(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\Big)$$

$$= P(X_1 = x_1)P(X_2 = x_2) \ldots P(X_n = x_n), \qquad \text{for all } x_1, x_2, \ldots, x_n.$$