Independent and identical distributed (IID) random variables example explained: Theses example come from Example 4.33 Lecture Note 1[1]

Consider the case with two random variables X and Y. Remember random variables is a formalization of a random experiment in a way that the structure of events is preserved. Let's say that we want to check the probabilistic relationship between two phenomena of interest: 'Virginia Tech football team wins' and 'it rains' on this Saturday. These two phenomena are formalized into two random variables X and Y: the football winning chance represented by X, and the weather condition by Y.

As we learned, the randomness structure (or, precisely, chance regularity) of any random variables can be expressed in terms of their probability function, in our case, $f(x)$ and $f(y)$. We need the joint distribution of $f(x)$ and $f(y)$ as $f(x,y)$ to evaluate whether X and Y are independent and identical distributed (IID), or not. Three examples will be presented next.

Example 1

Numerical values of events for each random variable are assigned as follow:

Phenomenon	**Events**	**Numerical value of each event**
Virginia Tech footfall team wins - X	Win	$x = 1$
	Lose	$x = 2$
It rains - Y	Rain	$y = 0$
	Not rain	$y = 2$

Assume that we know *in advance* the joint probability $f(x,y)$ as follow:

$x\backslash y$	$y = 0$	$y = 2$
$x = 1$	$f(1,0) = 0.18$	$f(1,2) = 0.12$
$x = 2$	$f(2,0) = 0.42$	$f(2,2) = 0.28$

In term of events, $f(1,0) = 0.18$ means that the probability of both a win for the Virginia Tech team ($x = 1$) and rain to happen ($y = 0$) is 0.18, and so on.

From this table, we can calculate the marginal probability of each events.

$x\backslash y$	$y = 0$	$y = 2$	
$x = 1$	$f(1,0) = 0.18$	$f(1,2) = 0.12$	$f_x(x = 1) = 0.18 + 0.12 = 0.3$
$x = 2$	$f(2,0) = 0.42$	$f(2,2) = 0.28$	$f_x(x = 2) = 0.42 + 0.28 = 0.7$
	$f_y(y = 0) = 0.18 + 0.4 = 0.6$	$f_y(y = 2) = 0.12 + 0.28 = 0.4$	

[1] These examples are also Example 4.39 in chapter 4 of Dr. Spanos' textbook.

The meaning of $f_x(x = 1) = 0.3$ is that the probability of the event of Virginia Tech's victory (averaging over information about the weather condition) is 0.3. Now, we will check the IID relationship between X and Y.

First, two random variables X and Y are independent if and only if:

$$f(x, y) = f_x(x). f_y(y), \ \forall x \in \mathbb{R}_X, y \in \mathbb{R}_Y \quad \textbf{(1)}$$

Now, we check whether the condition **(1)** applies for all values of X and Y (all events related to X and Y):

$$(X, Y) = (1,0): f(1,0) = f_x(1). f_y(0) = 0.18 = 0.6 * 0.3$$

$$(X, Y) = (1,2): f(1,2) = f_x(1). f_y(2) = 0.12 = 0.4 * 0.3$$

$$(X, Y) = (2,0): f(2,0) = f_x(2). f_y(0) = 0.42 = 0.7 * 0.6$$

$$(X, Y) = (2,2): f(2,2) = f_x(2). f_y(2) = 0.28 = 0.7 * 0.4$$

These results suggest that X and Y are independent. Now, we need to check if X and Y are identical distributed (ID). Two random variables X and y are defined as ID if and only if:

$$f_x(x) = f_y(y), \ \forall x \in \mathbb{R}_X, y \in \mathbb{R}_Y \quad \textbf{(2)}$$

First, we have $f_x(2) = 0.6 \neq f_y(2) = 0.4$. Secondly, for each of $f_x(1)$ and $f_y(0)$, there are not existence of corresponding $f_y(1)$ and $f_x(0)$. In other words, the domain $\mathbb{R}_X = \{1,2\}$ of $f_x(x)$ and $\mathbb{R}_Y = \{0,2\}$ of $f_y(y)$ are not identical. Thus, the condition **(2)** does not apply. And, we can conclude X and Y are not ID.

In summary, the two random variables X and Y are independent but not ID.

<u>Example 2</u>

In this example, the joint and marginal probability distributions between X and Y are changed as:

$x \backslash y$	$y = 0$ (raining)	$y = 1$ (not raining)	
$x = 0$ (win)	$f(0,0) = 0.18$	$f(0,1) = 0.12$	$f_x(0) = 0.18 + 0.12 = 0.3$
$x = 1$ (lose)	$f(2,0) = 0.42$	$f(1,1) = 0.28$	$f_x(1) = 0.42 + 0.28 = 0.7$
	$f_y(0) = 0.18 + 0.42 = 0.6$	$f_y(1) = 0.12 + 0.28 = 0.4$	

Compared to the joint and marginal probability in example 1, the probability of all events in X and Y are the same. The only difference is that the numbers assigned to each of the events are modified. Now, the domain of X is $R_X = \{0,1\}$ (rather than $\{1,2\}$), and of Y is $R_Y = \{0,1\}$ (rather than $\{0,2\}$).

Because the joint probability distribution $f(x, y)$ is the same, X and Y are also independent. How about the identical distributed condition? The probability distribution of X and Y are not the same:

$$f_x(0) = 0.6 \neq f_y(0) = 0.3$$

$$f_x(1) = 0.4 \neq f_y(0) = 0.7$$

Thus, the ID condition of X and Y are not satisfied. In conclusion, the two random variables X and Y are independent but still *not* IID.

Example 3

In example 3, the joint and marginal probability distributions between X and Y are modified as:

$x\backslash y$	$y = 0$ (raining)	$y = 1$ (not raining)	
$x = 0$ (win)	$f(0,0) = 0.36$	$f(0,1) = 0.24$	$f_x(0) = 0.36 + 0.24 = 0.6$
$x = 1$ (lose)	$f(1,0) = 0.24$	$f(1,1) = 0.16$	$f_x(1) = 0.24 + 0.28 = 0.4$
	$f_y(0) = 0.36 + 0.24 = 0.6$	$f_y(1) = 0.24 + 0.16 = 0.4$	

In this example, the assigned numerical value of each event and its probability are chosen to X and Y to be ID:

$$f_x(0) = 0.6 = f_y(0) = 0.6$$

$$f_x(1) = 0.4 = f_y(1) = 0.4$$

Now, we need to evaluate whether these two variables are independent:

$$(X, Y) = (0,0): f(0,0) = f_x(0). f_y(0) = 0.36 = 0.6 * 0.6$$

$$(X, Y) = (0,1): f(0,1) = f_x(0). f_y(1) = 0.24 = 0.6 * 0.4$$

$$(X, Y) = (1,0): f(1,0) = f_x(1). f_y(0) = 0.24 = 0.4 * 0.6$$

$$(X, Y) = (1,1): f(1,1) = f_x(1). f_y(1) = 0.16 = 0.4 * 0.4$$

This check confirms that X and Y are independent. Therefore, we finally have the two random variables X and Y to be truly IID.