Three pairwise i.i.d. random variables that are not i.i.d.

SML Introduction to Stochastic Calculus

Rasmus Peter Skadborg

We provide an example of three random variables $X_1, X_2, X_3 : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Lambda, \mathcal{E})$ with the following properties:

- They are *identically distributed*, i.e.

$$\mu_{X_i} = \mu_{X_j}, \qquad \text{for all } i, j \in \{1, 2, 3\}.$$

- They are *pairwise independent*, i.e.

$$\mu_{(X_i, X_j)}(A \times B) = \mu_{X_i}(A)\mu_{X_j}(B) \qquad \text{for all } i, j \in \{1, 2, 3\} \text{ with } i \neq j, \text{ and all } A, B \in \mathcal{E}.$$

- They are not all independent, i.e. there exist $A_1, A_2, A_3 \in \mathcal{E}$ such that

$$\mu_{(X_1, X_2, X_3)}(A_1 \times A_2 \times A_3) \neq \mu_{X_1}(A_1)\mu_{X_2}(A_2)\mu_{X_3}(A_3).$$

In other words, X_1, X_2, X_3 are pairwise i.i.d., but the three of them together are not independent, hence not i.i.d. This shows that pairwise independence is not quite a strong enough condition to define i.i.d.

The idea for the construction is pretty simple: We can think of X_1, X_2 as the outcomes of two independent fair coinflips (1 for heads, 0 for tails), and then let $X_3 = X_1 + X_2 \mod 2$. X_3 is equally likely to take on the values 0 and 1, so it has the same distribution as X_1 and X_2. If one knows the outcome of the two coinflips, then one also knows the value of X_3, so the three are not all independent. However, if one knows only the value of, say, the first coinflip X_1, then X_3 acts as if it were a fair coin independent from X_1 (it will be either a copy of X_2 or the opposite of X_2, depending on whether the first coin landed tails or heads), so we see that X_1 and X_3 are still independent (and similarly X_2 and X_3 are independent).

Let us formalize this idea. Let $\Omega = \{0, 1\}^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ which we will equip with the powerset σ-algebra $\mathcal{F} = \mathcal{P}(\Omega)$ and the uniform probability measure

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{4},$$

where $|A|$ denotes the number of elements in the set $A \in \mathcal{F}$. We set $\Lambda = \{0, 1\}$, which we also equip with the powerset σ-algebra $\mathcal{E} = \mathcal{P}(\Lambda)$. Note that all functions $\Omega \to \Lambda$ are measurable, since Ω is equipped with the powerset σ-algebra.

Now define the random variables $X_1, X_2, X_3 : \Omega \to \Lambda$ by

$$X_1(\omega_1, \omega_2) = \omega_1, \qquad X_2(\omega_1, \omega_2) = \omega_2, \qquad X_3(\omega_1, \omega_2) = \begin{cases} 0 & \text{if } \omega_1 = \omega_2, \\ 1 & \text{if } \omega_1 \neq \omega_2, \end{cases}$$

for $\omega = (\omega_1, \omega_2) \in \{0,1\}^2$. The induced measures μ_{X_i} of these discrete random variables are completely determined by the PMF's $p_{X_i}(x) = \mathbb{P}(X_i = x)$. Since Ω is quite small, and since the probability measure $\mathbb{P}$ is uniform, we can compute $\mathbb{P}(X_i = x)$ simply by counting. Let us summarize the values of X_1, X_2, X_3 on all points in Ω with a table:

ω_1	ω_2	$X_1(\omega_1, \omega_2)$	$X_2(\omega_1, \omega_2)$	$X_3(\omega_1, \omega_2)$
0	0	0	0	0
0	1	0	1	1
1	0	1	0	1
1	1	1	1	0

We see that each X_i is equally likely to take on the values 0 and 1, i.e.

$$p_{X_i}(0) = p_{X_i}(1) = \frac{1}{2}, \qquad \text{for all } i \in \{1,2,3\}.$$

Thus, since X_1, X_2, X_3 have identical PMF's, their induced measures also coincide, i.e. they are identically distributed.

From the above table, we can also see that for each $i, j \in \{1,2,3\}$ with $i \neq j$, the pair (X_i, X_j) takes on the four values $(0,0), (0,1), (1,0)$ and $(1,1)$ each at precisely one point in Ω. Therefore

$$p_{(X_i,X_j)}(x,y) = \frac{1}{4} = p_{X_i}(x) p_{X_j}(y), \qquad \text{for all } i,j \in \{1,2,3\} \text{ with } i \neq j, \text{ and all } \; x,y \in \{0,1\}.$$

This implies that X_i, X_j are independent, for if $A, B \in \mathcal{E}$, then

$$\mu_{(X_i,X_j)}(A \times B) = \sum_{(a,b) \in A \times B} p_{(X_i,X_j)}(a,b) = \sum_{(a,b) \in A \times B} p_{X_i}(a) p_{X_j}(b) = \sum_{a \in A} p_{X_i}(a) \sum_{b \in B} p_{X_j}(b) = \mu_{X_i}(A) \mu_{X_j}(B).$$

We have thus shown that X_1, X_2, X_3 are pairwise independent. On the other hand, one also sees from the table above that some values of the triple (X_1, X_2, X_3) are impossible; for example one cannot simultaneously have $X_1 = X_2 = 0$ and $X_3 = 1$. Therefore,

$$p_{(X_1,X_2,X_3)}(0,0,1) = 0 \neq \frac{1}{8} = p_{X_1}(0) p_{X_2}(0) p_{X_3}(1),$$

or in terms of the induced measures,

$$\mu_{(X_1,X_2,X_3)}(\{0\} \times \{0\} \times \{1\}) = 0 \neq \frac{1}{8} = \mu_{X_1}(\{0\}) \mu_{X_2}(\{0\}) \mu_{X_3}(\{1\}),$$

so the three random variables X_1, X_2, X_3 are not independent.