Collections of Random Variables: Theory

Contents

Joint probability mass function

Consider two random variables X and Y. The *joint probability mass function* of the pair (X, Y) is the function $f_{X,Y}(x, y)$ giving the probability that $X = x$ and $Y = y$. Mathematically (and introducing a simplified notation), we have:

$$p(x, y) \equiv p(X = x, Y = y) \equiv f_{X,Y}(x, y) := \mathbb{P}\left(\{\omega : X(\omega) = x, Y(\omega) = y\}\right).$$

Properties of the joint probability mass function

- It is nonnegative:

$$p(x, y) \geq 0.$$

- If you sum over all the possible values of all random variables, you should get one:

$$\sum_x \sum_y p(x, y) = 1.$$

- If you *marginalize* over the values of one of the random variables you get the pmf of the other. For example:

$$p(x) = \sum_y p(x, y),$$

and

$$p(y) = \sum_x p(x, y).$$

Joint probability mass function of many random variables

Take N random variables $X_1, \dots, X_N$. We can define their joint probability mass function in the same way we did it for two:

$$p(x_1, \dots, x_N) \equiv p(X_1 = x_1, \dots, X_N = x_N) \equiv f_{X_1,\dots,X_N}(x_1, \dots, X_N) := \mathbb{P}\left(\{\omega : X_1(\omega) = x_1, \dots, X_N(\omega) = x_N\}\right).$$

Just like before, we can marginalize over any subset of random variables to get the pmf of the remaining ones. For example:

$$p(x_i) = \sum_{x_j, j \neq i} p(x_1, \dots, x_N).$$

Joint probability density function

Let X and Y be two random variables. There joint probability density $f_{X,Y}(x,y)$ is the function that can give us the probability that the pair (X,Y) belongs to any "good" subset A of $\mathbb{R}^2$ as follows:

$$p\left((X,Y)\in A\right) = \int\int_A f_{X,Y}(x,y)dxdy.$$

Of course, we will be writing:

$$p(x,y) := f_{X,Y}(x,y),$$

when there is no ambiguity.

If you integrate one of the variables out of the joint, you get the PDF of the other variable. For example:

$$p(x) = \int_{-\infty}^{\infty} p(x,y)dy,$$

and

$$p(y) = \int_{-\infty}^{\infty} p(x,y)dx.$$

Conditioning a random variable on another

Consider two random variables X and Y. If we had observed that $Y=y$, how would this change the PDF of X? The answer is given via Bayes' rule. The PDF of X conditioned on $Y=y$ is:

$$p(x|y) = \frac{p(x,y)}{p(y)}.$$

The covariance operator

The covariance operator measures how correlated two random variables X and Y are. Its definition is:

$$\mathbb{C}[X,Y] = \mathbb{E}\left[(X-\mathbb{E}[X])\left(Y-\mathbb{E}[Y]\right)\right].$$

If $\mathbb{C}[X,Y]$ is positive, then we say that the two random variables are correlated. If it is negative, then we say that the two random variables are anti-correlated. If it is zero, then we say that the two random variables are not correlated. We will talk more about this in a later lecture.

A usefull property of the covariance operator is that it can give tell you something about the variance of the sum of two random variables. It is:

$$\mathbb{V}[X+Y] = \mathbb{V}[X] + \mathbb{V}[Y] + 2\mathbb{C}[X,Y].$$

Independent random variables

Take two random variables X and Y. We say that the two random variables are independent given the background information I, and we write:

$$X\perp Y|I,$$

if and only if conditioning on one does not tell you anything about the other, i.e.,

$$p(x|y,I) = p(x|I).$$

It is easy to show using Bayes' rule that the definition is consistent, i.e., you also get:

$$p(y|x,I) = p(y|I).$$

When there is no ambiguity, we can drop I.

Independent random variables

Repeat section

Independent random variables

Take two random variables X and Y. We say that the two random variables are independent given the background information I, and we write:

$$X \perp Y | I,$$

if and only if conditioning on one does not tell you anything about the other, i.e.,

$$p(x|y, I) = p(x|I).$$

It is easy to show using Bayes' rule that the definition is consistent, i.e., you also get:

$$p(y|x, I) = p(y|I).$$

When there is no ambiguity, we can drop I.

Properties of independent random variables

- The joint pmf factorizes:

$$p(x, y) = p(x)p(y).$$

- The expectation of the product is the product of the expectation:

$$\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y].$$

- The covariance of two independent random variables is zero:

$$\mathbb{C}[X, Y] = 0.$$

Be careful **the reverse is not true!**

- A consequence of the above property is that the variance of the sum of two independent random variables is the sum of the variables:

$$\mathbb{V}[X + Y] = \mathbb{V}[X] + \mathbb{V}[Y].$$

By Ilias Bilionis (ibilion[at]purdue.edu)