LECTURE: MARGINAL AND CONDITIONAL DISTRIBUTION

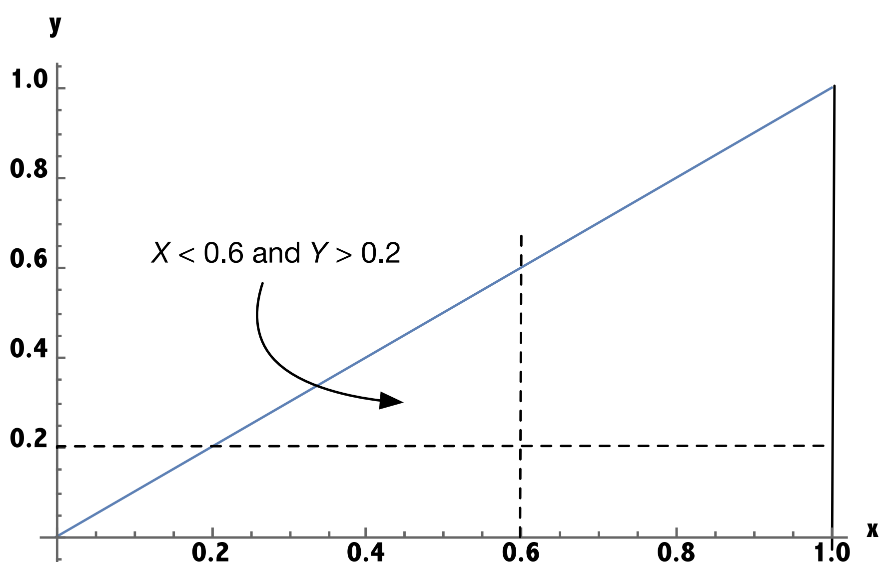
1. Multivariate Distributions

Let X and Y be random variables with joint density $f(x, y)$ where

$$f(x,y) = \begin{cases} 8xy & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

(b) Find $P((X < 0.6) \cap (Y > 0.2))$

Here we are finding the probability that the pair (X, Y) falls in a specific region of the plane. The first step is to draw the region.



From this picture, we get the limits of integration. Let's integrate in the y direction first. Recall that we found that $k = 8$ above. This gives

$$\begin{aligned} P(X < 0.6 \cap Y > 0.2) &= \int_{0.2}^{0.6} \int_{0.2}^{x} 8xy \, dydx \\ &= 8 \int_{0.2}^{0.6} \left[x\frac{y^2}{2} \right]_{y=0.2}^{y=x} dx \\ &= 4 \int_{0.2}^{0.6} (x^3 - 0.04x) dx \\ &= 4 \left[\frac{x^4}{4} - 0.04\frac{x^2}{2} \right]_{0.2}^{0.6} \\ &= \left(0.6^4 - 0.2^4\right) - 0.08 \left(0.6^2 - 0.2^2\right) \\ &= 0.1024 \end{aligned}$$

Note: We could also have integrated in the x direction first.

2. Marginal Distribution

Suppose Y_1 and Y_2 are continuous random variables with joint density $f(y_1, y_2)$

Then Y_1 and Y_2 are themselves random variables, and their densities are called the **marginal densities** of Y_1 and Y_2.

How do we find the marginal densities? Recall that for the discrete case, we had

$$p_1(y_1) = \sum_{y_2} p(y_1, y_2)$$

Here we do the exact same thing, except we replace summation with integration:

Definition:

Let Y_1 and Y_2 be cont. random var with joint density $f(y_1, y_2)$

Then the **marginal densities** of Y_1 and Y_2 are given by:

$$f_1(y_1) = \int f(y_1, y_2)dy_2 \qquad f_2(y_2) = \int f(y_1, y_2)dy_1$$

In other words, we "integrate out" the other random variable.

Example 1:

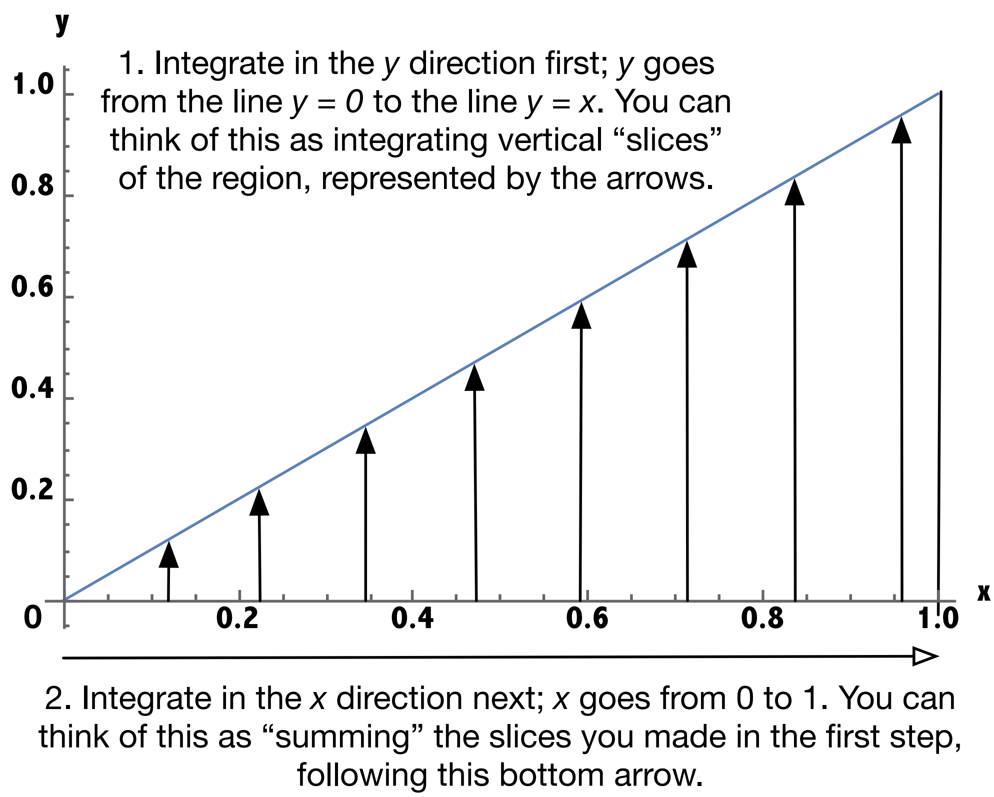
Let X and Y be random variables with joint density $f(x, y)$ where

$$f(x,y) = \begin{cases} 8xy & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the marginal densities for X and Y

We will denote them by $f_X(x)$ and $f_Y(y)$

STEP 1: Let's find the marginal density for X by integrating over y



To integrate in y, we start at $y = 0$ and integrate until we reach the line $y = x$. Thus the limits of integration are 0 and x.

$$f_X(x) = \int_0^x 8xy\,dy = 8x \left[\frac{y^2}{2}\right]_{y=0}^{y=x} = 4x^3$$

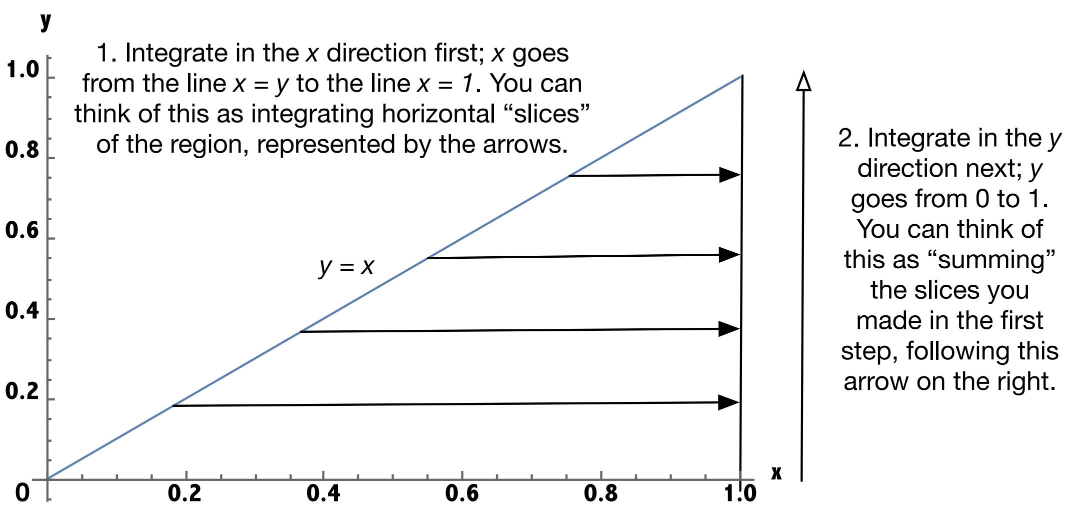
Note: The above expression is only valid for $0 \le x \le 1$. Outside that range, the marginal density is 0. Therefore

$$f_X(x) = \begin{cases} 4x^3 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

You can check that $f_X(x)$ in fact integrates to 1 and that f_X is a function of x alone; y does not appear anywhere since we integrated it out!

STEP 2: Find the marginal density for Y by integrating over x

The limits for x are $x = y$ and $x = 1$



$$f_Y(y) = \int_y^1 8xy\,dx = 8y\left[\frac{x^2}{2}\right]_{x=y}^{x=1} = 4y(1-y^2)$$

y can take values from 0 to 1, so the marginal density of Y is

$$f_Y(y) = \begin{cases} 4y(1-y^2) & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(b) Find the expected values for X and Y

Upshot: Since X is a one-dimensional random variable with density f_X, we can just use the formula for expectation in 1D

Definition:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x f_X(x) dx = \int_0^1 x \left(4x^3\right) dx = 4 \int_0^1 x^4 dx = 4/5$$

Here we used that f_X is 0 outside of $[0, 1]$

Similarly, we find the expected value of Y.

$$\begin{aligned} E(Y) = & \int_{-\infty}^{\infty} y f_Y(y) dy = \int_0^1 y f_Y(y)\, dy = \int_0^1 y \, \left[4y(1 - y^2)\right] dy \\ = & 4 \int_0^1 (y^2 - y^4) dy = 8/15 \end{aligned}$$

3. Conditional Distribution

Just as in the discrete case, we can talk about conditional distributions, for instance the distribution of Y_1 given that $(Y_2 = y_2)$.

Definition:

Let Y_1 and Y_2 be continuous random variables with joint density $f(y_1, y_2)$ and let $f_2(y_2)$ be the marginal density of Y_2

Then the **conditional density** of Y_1 given $(Y_2 = y_2)$ is:

$$f(y_1|y_2) = \frac{f(y_1, y_2)}{f_2(y_2)}$$

In other words, the conditional density is the joint density divided by the marginal density. Similarly, we can define define the conditional density of Y_2 given $(Y_1 = y_1)$

Mnemonic: Although this notation makes no sense, one way to remember this is

$$f(y_1|y_2) = \frac{f(Y_1 = y_1 \text{ and } Y_2 = y_2)}{f(Y_2 = y_2)} = \frac{f(y_1, y_2)}{f_2(y_2)}$$

(c) Find the conditional density of X given $(Y = y)$

Application: For concreteness, think of X as the temperature of a random spot in the world and Y as a random time and $y = 12$ pm. Then $f(x|y)$ gives you the density of the temperature in the world at noon.

$$f(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{8xy}{4y(1 - y^2)} = \frac{2x}{1 - y^2}$$

Warning: We are not done! We need to figure out for which x and y this is valid. Note that if $(Y = y)$, then, in the picture above, X can only range from the diagonal line $y = x$ to 1, i.e. X must be between

y and 1. Therefore we get:

$$f(x|y) = \begin{cases} \frac{2x}{1-y^2} & y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Note: here y is fixed, so we only care about the range of x

(d) Find the conditional density of Y given $(X = x)$

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{8xy}{4x^3} = \frac{2y}{x^2}$$

Moreover, if $(X = x)$, then y can only range from 0 to the diagonal line $y = x$, i.e. Y must be between 0 and x

$$f(y|x) = \begin{cases} \frac{2y}{x^2} & 0 \leq y \leq x \\ 0 & \text{otherwise} \end{cases}$$

(e) Find the (conditional) expected value of X given $(Y = y)$

The conditional density $f(x|y)$ is just a probability density of a continuous random variable in terms of x, so we can find its expected value using the standard expected value formula

Definition:

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f(x|y) dx$$

$$E[X|Y=y] = \int_{-\infty}^{\infty} xf(x|y)dx = \int_y^1 x\left(\frac{2x}{1-y^2}\right)dx = \frac{2}{1-y^2}\int_y^1 x^2$$
$$= \frac{2}{1-y^2}\left[\frac{x^3}{3}\right]_{x=y}^{x=1} = \frac{2(1-y^3)}{3(1-y^2)}$$

Note that we used the bounds on the conditional density in the second line above. Unsurprisingly, this depends on y.

Application: In terms of our previous example, $E[X|Y=y]$ gives you the average temperature of earth at $y=$12 pm.

4. Independence

Definition:

Let Y_1 and Y_2 have joint density $f(y_1, y_2)$ and $f_1(y_1)$ and $f_2(y_2)$ be the marginal densities of Y_1 and Y_2.

Then Y_1 and Y_2 are **independent** if

$$f(y_1, y_2) = f_1(y_1)f_2(y_2) \qquad \text{for all } y_1, y_2$$

In other words, two continuous random variables are independent if their joint density is the product of the two marginal densities.

(f) Are X and Y independent?

$f(x,y) = 8xy \neq f_X(x)f_Y(y) = 4x^3 4y(1-y^2)$ hence the answer is **NO**