LECTURE: MULTIVARIATE DISTRIBUTIONS

1. Conditional Distribution

Another question we might ask is:

Question: What is the distribution of Y_1 given that $(Y_2 = y_2)$?

Example 1:

In the exam-coffee example, what is the distribution of the number of cups of coffee drunk today X_2 given that a student has a midterm this week ($X_1 = \texttt{yes}$)?

Recall: We can display the joint probability distribution for the pair (X_1, X_2) in a 2×3 table:

$$\begin{array}{cc|ccc|} & & & X_2 & \\ & & 0 & 1 & 2 \\ & \texttt{yes} & 2/20 & 3/20 & 3/20 \\ X_1 & \texttt{no} & 6/20 & 4/20 & 2/20 \end{array}$$

To answer the question, we look at the first row of the table, which corresponds to $(X_1 = \texttt{yes})$

$$2/20 \quad 3/20 \quad 3/20$$

This is not a valid probability mass function, because the elements do not sum to 1. To fix that, we just need to divide by the marginal probability

$$p_1(\texttt{yes}) = P(X_1 = \texttt{yes}) = 2/20 + 3/20 + 3/20 = 8/20$$

We then get the conditional probability for X_2 given $(X_1 = \texttt{yes})$, which we can write as $p(y_2|\texttt{yes})$ or $p(y_2|Y_1 = \texttt{yes})$:

y_2	$p(y_2\|\texttt{yes})$
0	$(2/20)/(8/20) = 2/8$
1	$(3/20)/(8/20) = 3/8$
2	$(3/20)/(8/20) = 3/8$

Definition:

Let Y_1 and Y_2 be discrete random variables with joint pmf $p(y_1, y_2)$ and let $p_2(y_2)$ be the marginal distribution of Y_2.

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

$$p(y_1|y_2) = P(Y_1 = y_1|Y_2 = y_2) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_2 = y_2)} = \frac{p(y_1, y_2)}{p_2(y_2)}$$

In words, the conditional distribution is the joint distribution divided by the marginal distribution.

We can similarly define the cond. distribution of Y_2 given $(Y_1 = y_1)$ as

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

Just like we did in the example above

2. Independence

The final question to settle is independence. Roughly speaking, two random variables are independent of if the probabilities of each one are not affected by the value of the other one. The following will serve as our definition for independence of two discrete random variables.

Definition:

Let Y_1 and Y_2 be discrete random variables with joint pmf $p(y_1, y_2)$ and let $p_1(y_1)$ and $p_2(y_2)$ the marginal distributions of Y_1 and Y_2.

Then Y_1 and Y_2 are **independent** if

$$p(y_1, y_2) = p_1(y_1)p_2(y_2) \qquad \text{for all } y_1, y_2$$

In other words, two random variables are independent if their joint distribution is the product of the two marginal distributions, compare this with $P(A \cap B) = P(A)P(B)$ for independent events.

Example 2:

In the exam-coffee example above, are Y_1 and Y_2 independent?

NO because, for example

$$p(\texttt{yes}, 1) = \frac{3}{20} \neq p_1(\texttt{yes})\, p_2(1) = \left(\frac{8}{20}\right)\left(\frac{7}{20}\right)$$

I mean, did we really expect coffee consumption and exams to be independent? ☺

3. Distribution of Two Continuous Variables

Video: What is a Double Integral

Video: Changing the Order of Integration

We will repeat the same discussion but for a pair of **continuous** random variables. This will require integration in two dimensions, cf. the videos above.

Joint Probability Density: For two continuous random variables, we have a joint density function, which is the continuous analogue of the joint pmf.

Definition:

Let Y_1 and Y_2 be two continuous random variables. Then the **joint density** of Y_1 and Y_2 is any function $f(y_1, y_2)$ such that

(1) $f(y_1, y_2) \geq 0$ for all y_1, y_2

(2)
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 dy_2 = 1$$

In other words, $f(y_1, y_2)$ is non-negative and integrates/sums up to 1.

In the single variable case, we had

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

Here we have the same thing except we integrate the joint density over a region:

Definition:

Let Y_1 and Y_2 be two continuous random variables with joint density $f(y_1, y_2)$. Let A be a region of the plane. Then

$$P((Y_1, Y_2) \in A) = \int \int_A f(y_1, y_2) dy_1 dy_2$$

Interpretation of f: Just as in the single-variable case, we have

$$f(a, b) \approx \frac{P((a \leq Y_1 \leq a + h) \text{ and } (b \leq Y_2 \leq b + k))}{hk}$$

So $f(a, b)$ is the probability that (Y_1, Y_2) lies in the rectangle $[a, a + h] \times [b, b + k]$ but scaled by the area hk of that rectangle.

Example 3:

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

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

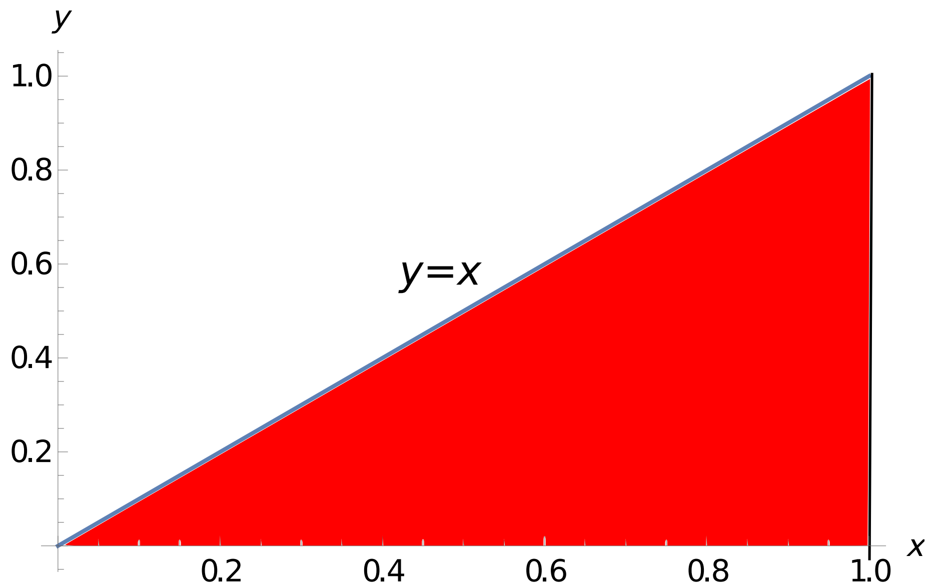
(a) Find the value of k such that $f(x, y)$ is a valid joint density function.

$$\text{We need } \int \int_A f(x, y) dx\, dy = 1$$

The bounds of the density function are the following:

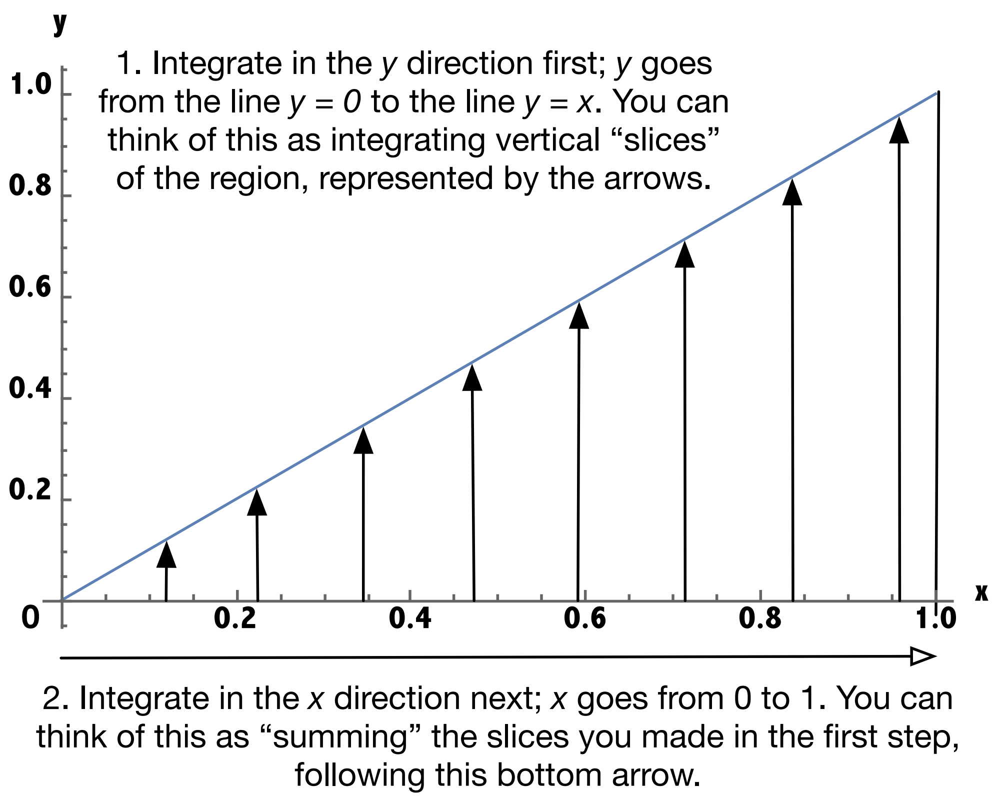
$$A = \{y \geq 0, x \leq 1, y \leq x\}$$

This describes the triangular region illustrated below.



Note: Whenever we have a double integral, we have two choices when we do our integration. We can integrate in x direction first, or we can integrate in the y direction first. Both ways give the same answer, but sometimes one is easier than the other. We will show both of them here.

Method 1: Let's start by integrating in the y direction first.



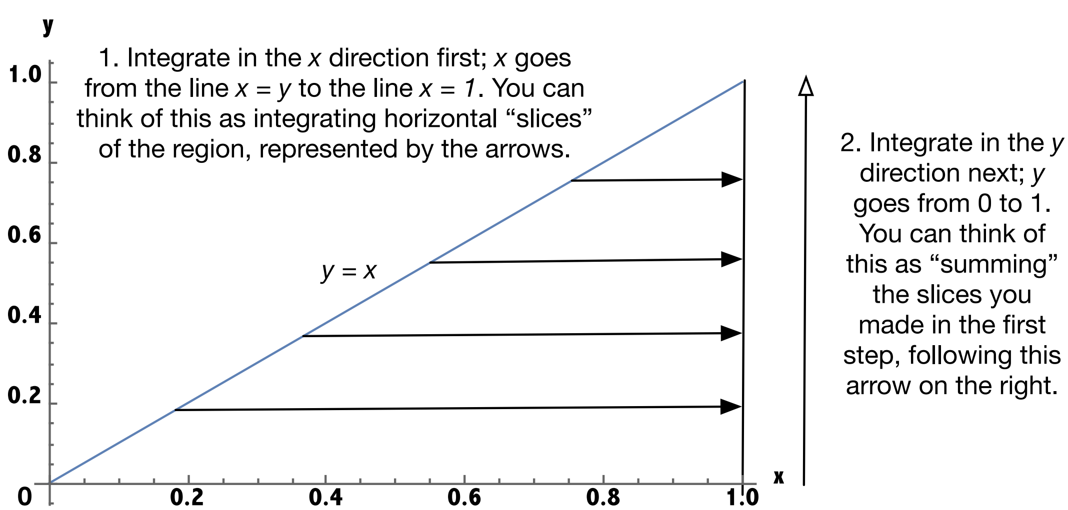
y goes from 0 to the diagonal line $y = x$, so those are the limits for the integral with respect to y (inner integral). Then x goes from 0 to 1, so those are the limits for the integral with respect to x (outer integral).

Putting this together, we have:

$$\begin{aligned}1 &= \int_0^1 \int_0^x kxy\,dydx \\ &= k\int_0^1 x\left[\frac{y^2}{2}\right]_{y=0}^{y=x} dx \\ &= \frac{k}{2}\int_0^1 x^3 dx \\ &= \frac{k}{2}\left[\frac{x^4}{4}\right]_0^1 = \frac{k}{8}\end{aligned}$$

Therefore $k = 8$

Method 2: This time we integrate in the x direction first.



x starts at the diagonal line $y = x$ and goes to $x = 1$. The diagonal line has the equation $y = x$, which we solve for x to get $x = y$. Thus the lower limit is the function $x = y$. The upper limit is 1, so the limits of the integral with respect to x (inner integral) are y and 1. Then y goes from 0 to 1, so those are the limits for the integral with respect to y (outer integral). Putting this together, we have:

$$\begin{aligned}
1 &= \int_0^1 \int_y^1 kxy \, dx dy \\
&= k \int_0^1 y \left[\frac{x^2}{2} \right]_{x=y}^{x=1} dy \\
&= \frac{k}{2} \int_0^1 y(1 - y^2) dy \\
&= \frac{k}{2} \int_0^1 (y - y^3) dy \\
&= \frac{k}{2} \left[\frac{y^2}{2} - \frac{y^4}{4} \right]_0^1 = \frac{k}{8}
\end{aligned}$$

We get the same answer! Which one was easier?

Aside: Fubini's theorem says that for "nice" functions f, both ways will always give the same answer.