May 21 Lec—Multivariate Distributions Intro, jp(d)fs, Marginal Distributions

Examples related to beta distributions

1. Let X be a continuous random variable such that

$$f(x) = \begin{cases} cx^2(1-x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find C and $\mathbb{E}\left(\min\left(X, \frac{1}{2}\right)\right)$.
We note that $X \sim B(\alpha = 3, \beta = 2)$. So $C = \frac{\Gamma(3+2)}{\Gamma(3)\Gamma(2)} = \frac{4!}{2!1!} = 12$.

$$\mathbb{E}\left(\min\left(X, \frac{1}{2}\right)\right) = \int_0^1 12x^2(1-x) \cdot \min\left(x, \frac{1}{2}\right) dx$$

$$= \int_0^{1/2} 12x^2(1-x) \cdot x \, dx + \int_{\frac{1}{2}}^1 12x^2(1-x) \cdot \frac{1}{2} \, dx$$

$$\ldots$$

2. $m_X(t) = \frac{e^{3t}}{1-2t}$ for $t < \frac{1}{2}$
Compute $\mathbb{E}(X)$, $V(X)$, and $P(X > 5)$.
Let $Y \sim G(\alpha = 1, \beta = 2)$ (i.e. exponential, $\beta = 2$; or chi-square, degree of freedom $p = 2$). We have $X = Y + 3$.
$\mathbb{E}(X) = \mathbb{E}(Y) + 3 = 5$
$V(X) = V(Y) = 4$
$P(X > 5) = P(Y > 2) = e^{-\frac{2}{2}} = e^{-1}$ (using the survival function of an exponential distribution)

The quiz tonight will cover Chapter 4.

Multivariate distributions

We'll mostly talk about bivariate distributions.
This chapter will use Calc 3 knowledge, but the professor thinks that you

should still understand even if you haven't taken it before.

Motivation for some discussion about double integrals:
Given a uniform probability distribution with nontrivial support on $(0, 2)$, you know that to find the constant λ of the pdf we would need to do
$\int_0^2 \lambda\, dx = 1 \implies \lambda = \frac{1}{2}$

Given an isosceles right triangle with leg lengths 1, suppose it has uniform density C and total mass 1. Then $\iint_{\text{triangle}} C dA = 1 \implies C = 2$.

Definition: Joint probability function

Let X and Y be two *discrete* random variables. The joint probability function (jpf) of X and Y is the function

$$p(x, y) = P(X = x, Y = y)$$
$$= P(\{X = x\} \cap \{Y = y\})$$

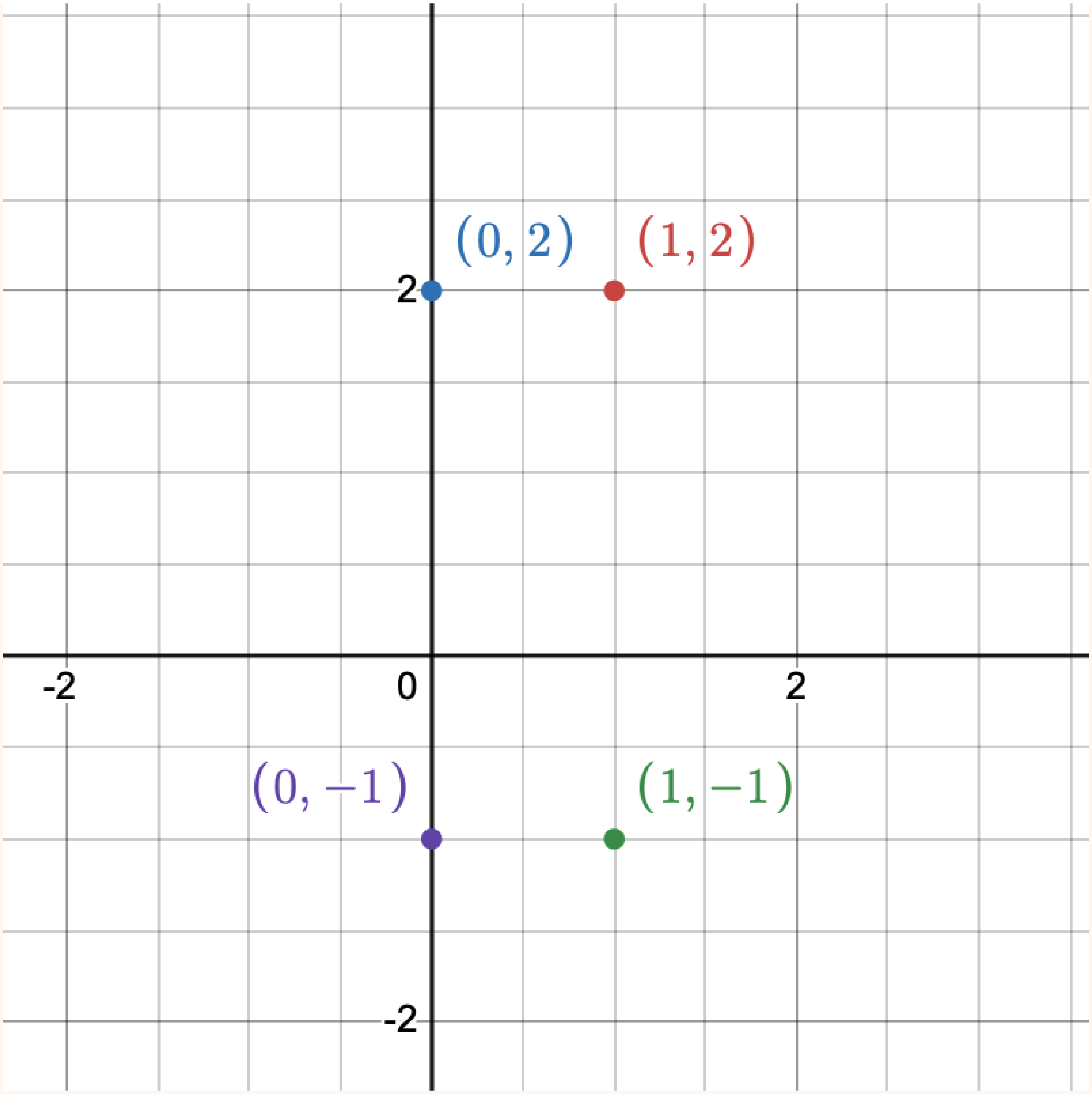
Properties of the jpf

1. $0 \leq p(x, y) \leq 1$
2. The set $\{(x, y) \mid p(x, y) > 0\}$, called the **support** of (X, Y), is discrete.
3. $\sum_{x,y} p(x, y) = 1$

Example

Urn A contains 3 Red and 7 Blue. Urn B contains 6 Red and 4 Blue.
An (unfair) coin is such that $P(H) = \frac{1}{3}$.
Flip the coin. If the result is H, set $X = 1$, and draw a ball from Urn A. Otherwise, set $X = 0$ and draw a ball from Urn B.
If the ball is Red, set $Y = 2$. Otherwise, set $Y = -1$.

Picture of the support:



We calculate $P(X=0, Y=-1)$

$$= P(Y=-1|X=0) \cdot P(X=0)$$

$$= \frac{4}{10} \cdot \frac{2}{3} = \frac{8}{30}$$

Similarly, $P(X=0, Y=2)$

$$= P(Y=2|X=0) \cdot P(X=0)$$

$$= \frac{6}{10} \cdot \frac{2}{3} = \frac{12}{30}$$

We continue to calculate the other ones:

$$P(X=1, Y=-1) = P(Y=-1|X=1) \cdot P(X=1)$$
$$= \frac{7}{10} \cdot \frac{1}{3} = \frac{7}{30}$$
$$P(X=1, Y=2) = P(Y=2|X=1) \cdot P(X=1)$$
$$= \frac{3}{10} \cdot \frac{1}{3} = \frac{3}{30}$$

We can summarize the above in a table:

y-value \ x-value	0	1
-1	$\frac{8}{30}$	$\frac{7}{30}$
-2	$\frac{12}{30}$	$\frac{3}{30}$

The total weight over all the cells is 1, as expected.

Now let's see the analog of all this stuff in terms of continuous random variables.

Definition: Jointly continuous random variable

Two continuous random variables X and Y are jointly described by the **joint probability density function** (jpdf) which is a function $f(x, y)$ such that

1. $f(x, y) \geq 0$
2. $\iint_{\mathbb{R}^2} f(x, y)\, dA = 1$ (i.e. the volume under the surface formed by $f(x, y)$ over the entire xy-plane must total to 1)

Example

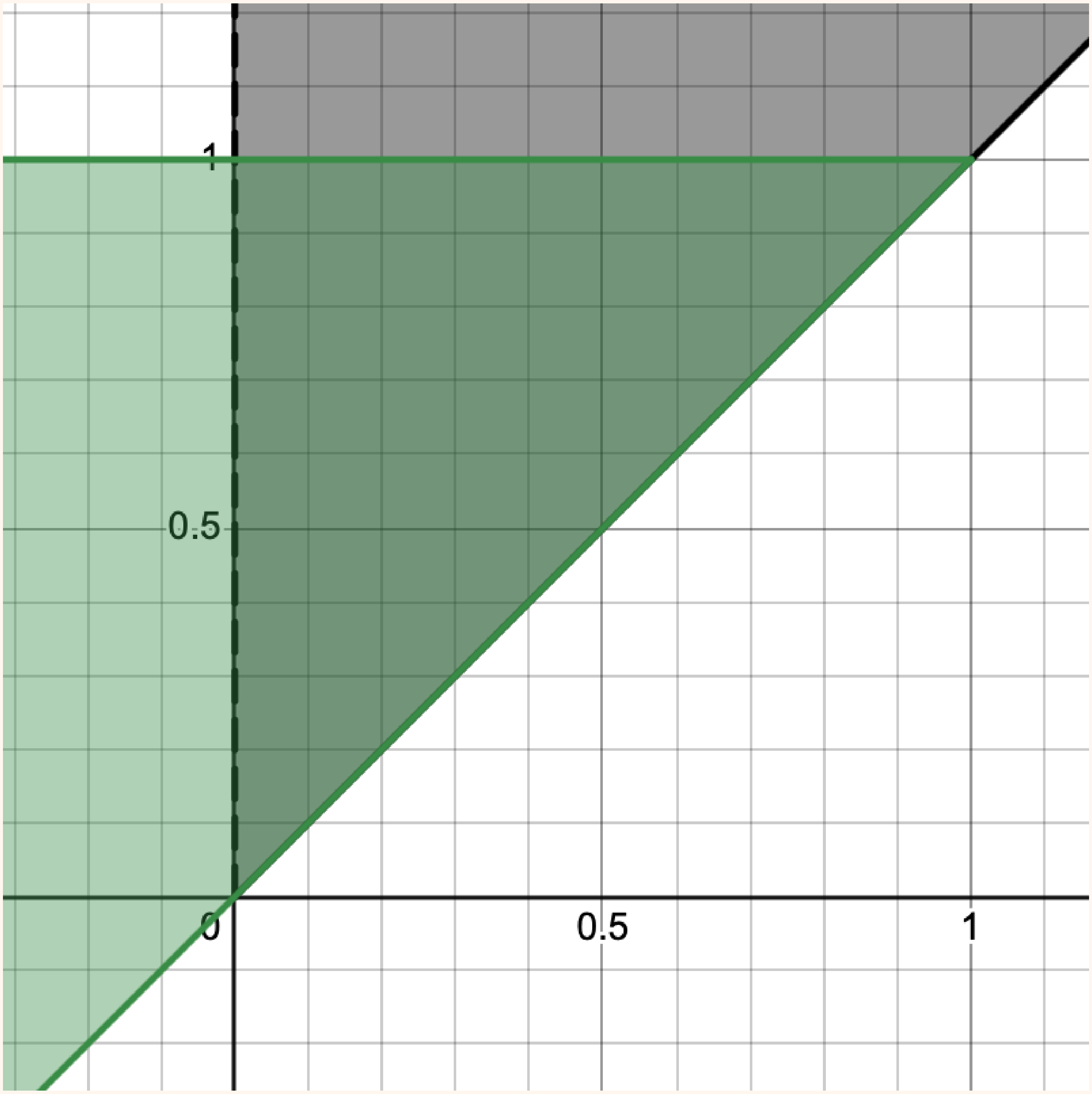
$$f(x, y) = \begin{cases} Cx & 0 < x \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find C such that f is a jpdf ($C > 0$).

Before you do any work with double integrals, *draw the domain (support)!!!!!!!!*

As always, x is the kernel.

The domain is the overlap of the black area $(0 < x \leq y)$ and the green area $(x \leq y \leq 1)$.



You can also sketch the area by thinking, for any fixed y in $(0, 1]$, what are all the permitted values of x. (or for any fixed x in $(0, 1]$, what are all the permitted values of y)

We now need to compute $\iint_D Cx\, dA$. To do this, we need to parametrize D, by making one variable in terms of the other.

The first way is to parametrize x in terms of y. That means, if we let y be the free variable (not depending on anything), what are all the possible values of x given fixed y?

We have that y can range from 0 to 1, and for any y, x can vary from 0 to y.
We then conclude that with this parametrization, we need to "sum" with respect to x first because it depends on y, and then we "sum" with respect to y. This means, we make x the inner integral and y the outer one.

$$\iint_D Cx\,dA = \int_0^1 \left(\int_0^y Cx\,dx\right) dy$$
$$= \int_0^1 C \cdot \frac{1}{2}y^2\,dy = C \cdot \frac{1}{6}$$

We can also parametrize D in another way, by letting x be the free variable and let y depend on x. x can range from 0 to 1, and for any fixed x, y can range from x to 1.

$$\iint_D Cx\,dA = \int_0^1 \left(\int_x^1 Cx\,dy\right) dx$$
$$= C\int_0^1 \left(x\int_x^1 dy\right) dx$$
$$= C \cdot \int_0^1 x(1-x)\,dx = C \cdot \frac{\Gamma(2)\Gamma(2)}{\Gamma(4)} = C \cdot \frac{1}{6}$$

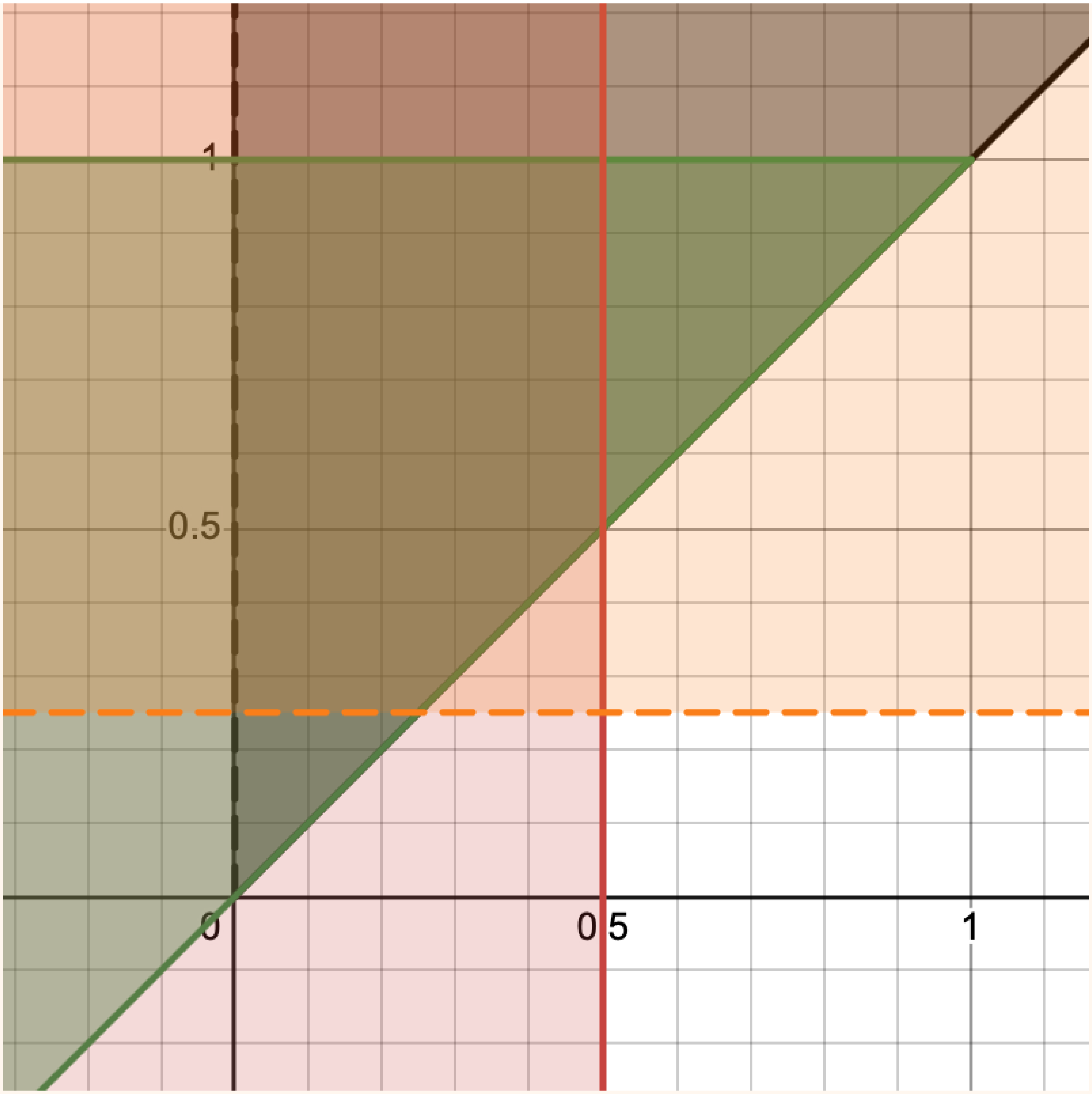
Note that the two integrals are the same (Fubini's Theorem).
Regardless of the parametrization we chose, we deduce $C = 6$.

Another continuous example

From this point forward, we will use this particular jpdf a lot.
For

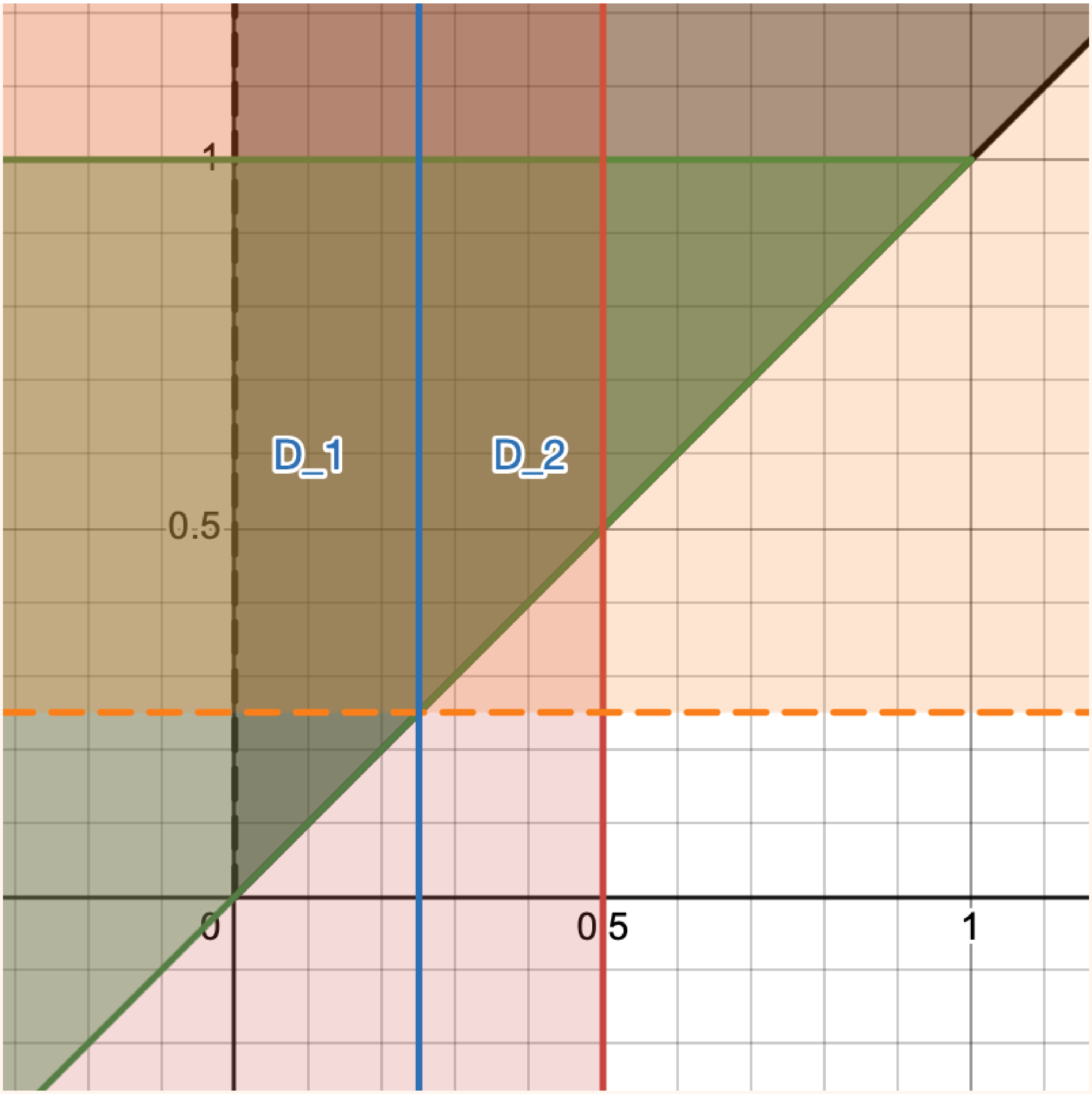
$$f(x,y) = \begin{cases} 6x & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

find $P\left(X \le \frac{1}{2}, Y > \frac{1}{4}\right)$.
This area is represented below as the area where all four colors (green, black, light orange, light red) overlap, sorry for the messy picture :(



Note that the function is 0 on the non-green, non-black areas of the graph.

We can split this area into two areas D_1, D_2 to parametrize it better:



For D_1, x ranges from 0 to $\frac{1}{4}$ and for any fixed x, y ranges from $\frac{1}{4}$ to 1.

For D_2, x ranges from $\frac{1}{4}$ to $\frac{1}{2}$ and for any fixed x, y ranges from x to 1.

So

$$P\left(X \leq \frac{1}{2}, y > \frac{1}{4}\right) = \iint_{D_1} 6x\, dA + \iint_{D_2} 6x\, dA$$

$$= \int_0^{\frac{1}{4}} \int_{\frac{1}{4}}^1 6x\, dy\, dx + \int_{\frac{1}{4}}^{\frac{1}{2}} \int_x^1 6x\, dy\, dx$$

For the remainder of today, we will talk about different X and Y distributions and their resulting joint distribution.

Definition: Marginial distribution

1. Let $p(x, y)$ be the jpf of two discrete random variables X, Y. The marginal probability function of X is

$$p_X(x) = P(X = x) = \sum_y p(x, y)$$

   Similarly, the marginal probability function of Y is

$$p_Y(y) = P(Y = y) = \sum_x p(x, y)$$

2. Let $f(x, y)$ be the jpdf of two continuous random variables X, Y. The marginal pdf of X is

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

   and the marginal pdf of Y is

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

Example of discrete marginal pfs (see above)

y-value\ x-value	-1	0	1	$p_Y(y)$
0	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{2}$
1	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
$p_X(x)$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$	

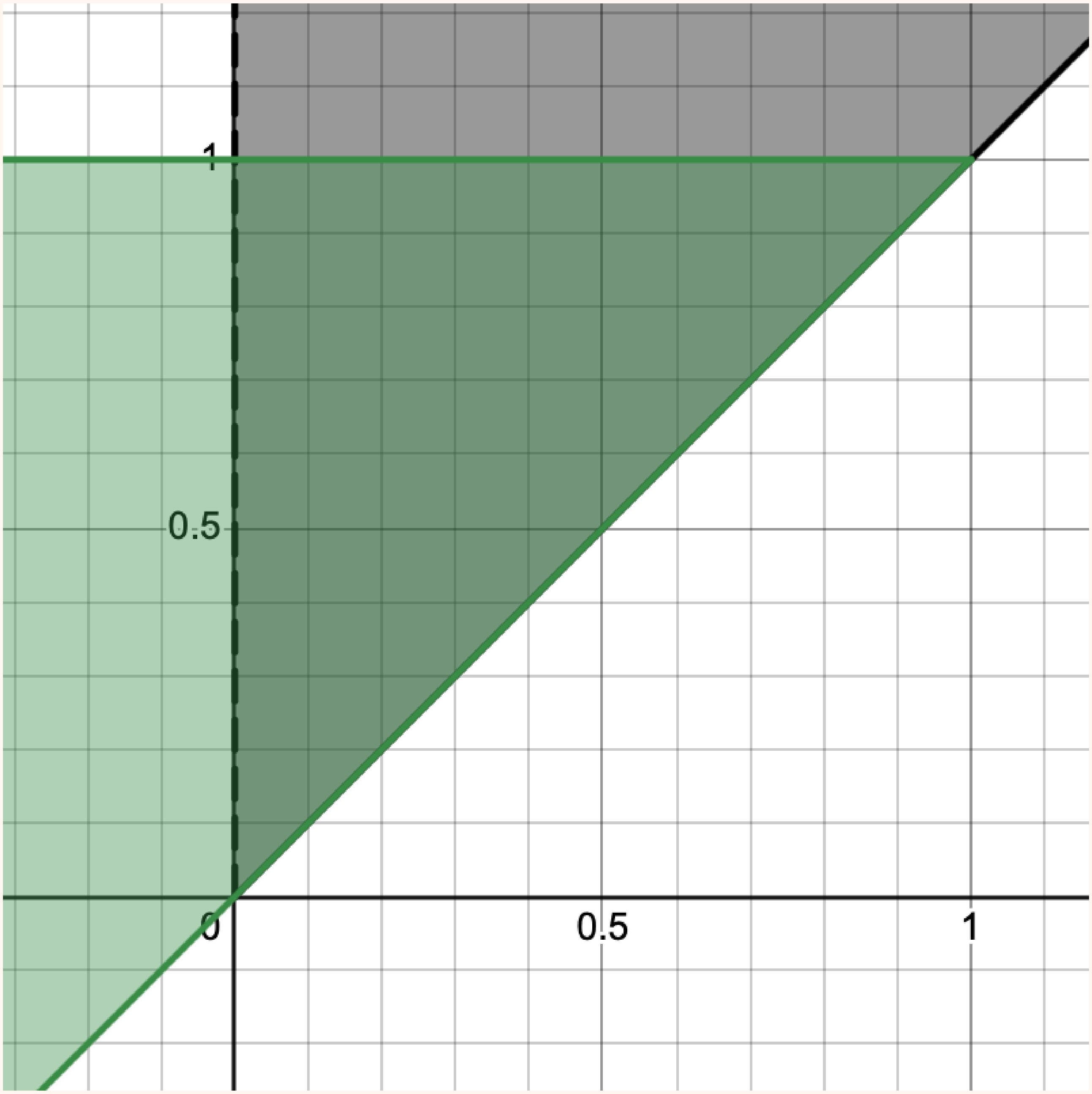
Chanllenge yourself with making up a randomized experiment that has this distribution. (It might help to note that Y is a Bernoulli variable, $p = \frac{1}{2}$.)

A continuous example

Take again

$$f(x,y) = \begin{cases} 6x & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Picture of support:



We have

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

$$= \begin{cases} \int_x^1 6x\,dy & x \in (0,1) \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} 6x(1-x) & x \in (0,1) \\ 0 & \text{otherwise} \end{cases}$$

so $X \sim Beta(\alpha = 2, \beta = 2)$.

Similarly, we have

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$= \begin{cases} \int_0^y 6x\,dx & y \in (0,1) \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} 3y^2 & y \in (0,1) \\ 0 & \text{otherwise} \end{cases}$$

so $Y \sim Beta(\alpha = 3, \beta = 1)$.

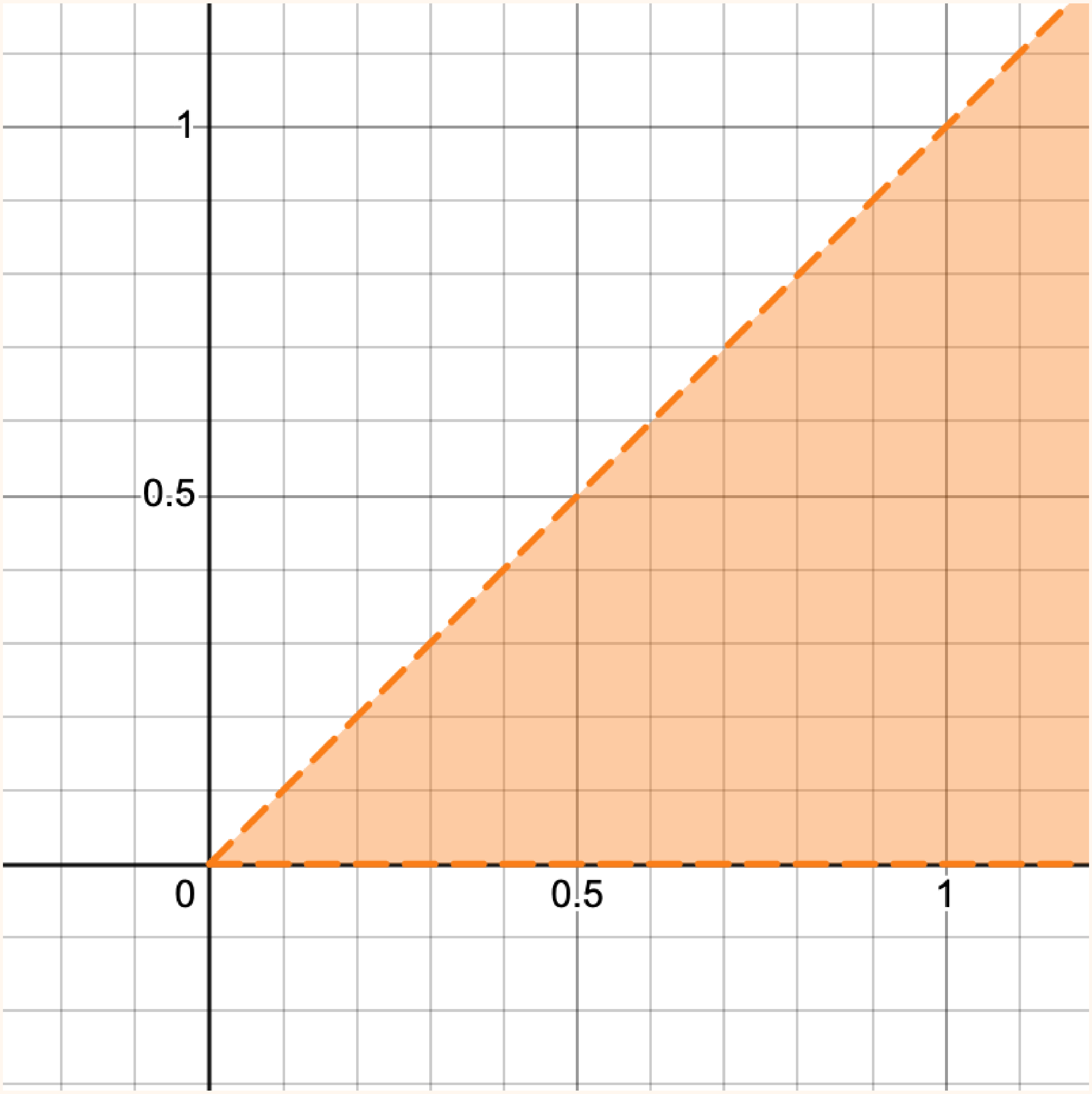
Yet another example

Let

$$f(x,y) = \begin{cases} Ce^{-x} & 0 < y < x \\ 0 & \text{otherwise} \end{cases}$$

Find C such that f is a jpdf, and find f_X, f_Y.

Picture of support:



Parametrizing, we have that $x \in (0, \infty)$ and $y \in (0, x)$.
We need to find C such that $C \cdot \iint_D e^{-x} \, dA = 1$.

$$\iint_D e^{-x} dx = \int_0^\infty \int_0^x e^{-x} \, dy \, dx$$

$$= \int_0^\infty x e^{-x} \, dx = \Gamma(2) = 1$$

So $C = 1$, and

$$f(x, y) = \begin{cases} e^{-x} & 0 < y < x \\ 0 & \text{otherwise} \end{cases}$$

$$f_X(x) = \begin{cases} \int_0^x e^{-x}\, dy & x \in (0,\infty) \\ 0 & \text{otherwise} \end{cases} = \begin{cases} xe^{-x} & x \in (0,\infty) \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \int_y^\infty e^{-x}\, dx & y \in (0,\infty) \\ 0 & \text{otherwise} \end{cases} = \begin{cases} e^{-y} & y \in (0,\infty) \\ 0 & \text{otherwise} \end{cases}$$

(so $Y \sim Exponential(mean = 1)$.)

Definition: Conditional distribution

1. Let p be the jpf of a pair (X, Y) of discrete random variables. The conditional pf of Y given $X = x$ is defined as

$$p(y|X = x) = P(Y = y|X = x) = \frac{p(x, y)}{p_X(x)}$$

   for $p_X(x) \neq 0$.
   Similarly, the conditional pf of X given $Y = y$ is defined as

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

2. Let f be the jpdf of a pair (X, Y) of continuous random variables. The conditional pdf of Y given $X = x$ is

$$f(y|X = x) = \frac{f(x, y)}{f_X(x)}$$

   and the conditional pdf of X given $Y = y$ is

$$f(x|Y = y) = \frac{f(x, y)}{f_Y(y)}$$

Examples

y**-value** x**-value**	-1	0	1	$p_Y(y)$
0	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{2}$
1	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
$p_X(x)$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$	

Refer to the table above.

Let us find $p(x|Y=y)$ for $y=0,1$.

$p(-1|Y=0)=\frac{\frac{1}{4}}{\frac{1}{2}}=\frac{1}{2}$

$p(0|Y=0)=\frac{\frac{1}{6}}{\frac{1}{2}}=\frac{1}{6}$

$p(1|Y=0)=\frac{\frac{1}{12}}{\frac{1}{2}}=\frac{1}{6}$

Similarly, for $y=1$:

$p(-1|Y=1)=\frac{\frac{1}{4}}{\frac{1}{2}}=\frac{1}{2}$

$p(0|Y=1)=\frac{0}{\frac{1}{2}}=0$

$p(1|Y=1)=\frac{\frac{1}{4}}{\frac{1}{2}}=\frac{1}{2}$