Joint pdfs of multiple random variables

Just as with discrete random variables, we can describe everything there is to know about a **pair** of random variables X, Y associated with the same experiment using their **joint pdf** $f_{X,Y}(x, y)$.

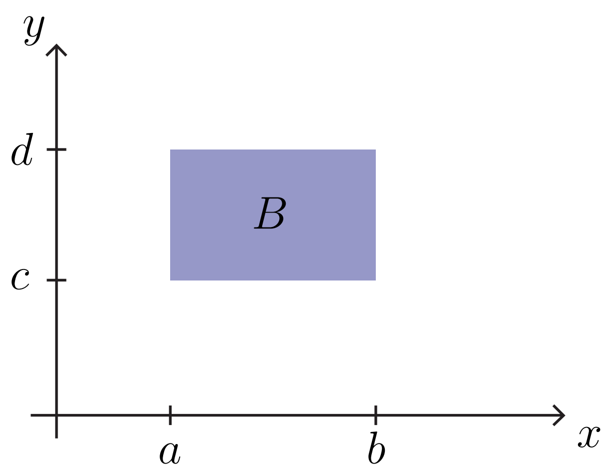
The joint pdf $f_{X,Y}(x, y)$ is a positive function of two variables from which we can calculate the probability of any event B (i.e., "any" subset of the plane $\mathbb{R}^2$):

$$\mathrm{P}\left(B\right) = \int\int_{(x,y)\in B} f_{X,Y}(x, y)\ \mathrm{d}x\,\mathrm{d}y$$

For example, the event

$$B = \{a \le X \le b \text{ and } c \le Y \le d\}$$

corresponds to a rectangle in the plane:



and

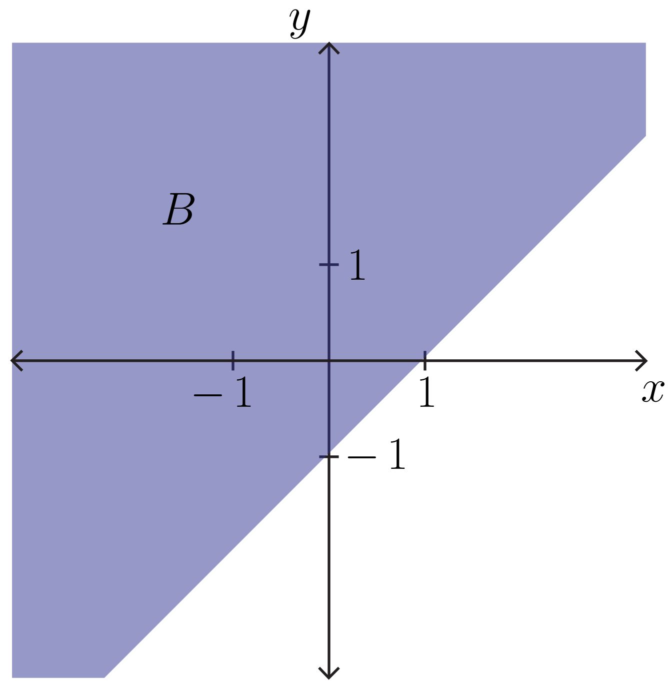
$$\begin{aligned}\mathrm{P}\left(B\right) &= \mathrm{P}\left(a \le X \le b,\ c \le Y \le d\right) \\ &= \int_{y=c}^{d}\int_{x=a}^{b} f_{X,Y}(x, y)\ \mathrm{d}x\ \mathrm{d}y\end{aligned}$$

Another example: Let B be the event

$$B = \{X \leq Y + 1\}$$

We can specify this as a subset of the plane with

$$B = \{(x, y) \mid x \leq y + 1\}$$



Then

$$\begin{aligned}
\mathrm{P}(B) &= \mathrm{P}(X \leq Y + 1) \\
&= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{y+1} f_{X,Y}(x, y)\, \mathrm{d}x\, \mathrm{d}y
\end{aligned}$$

or equivalently

$$= \int_{x=-\infty}^{\infty} \int_{y=x-1}^{\infty} f_{X,Y}(x, y)\, \mathrm{d}y\, \mathrm{d}x$$

Example. Han and Chewbacca plan to meet at a certain place and a certain time. Both will be delayed for some random amount of time between 0 and 1 hour; let

$$X = \text{amount of time (in hours) Han is delayed}$$
$$Y = \text{amount of time (in hours) Chewbacca is delayed}$$

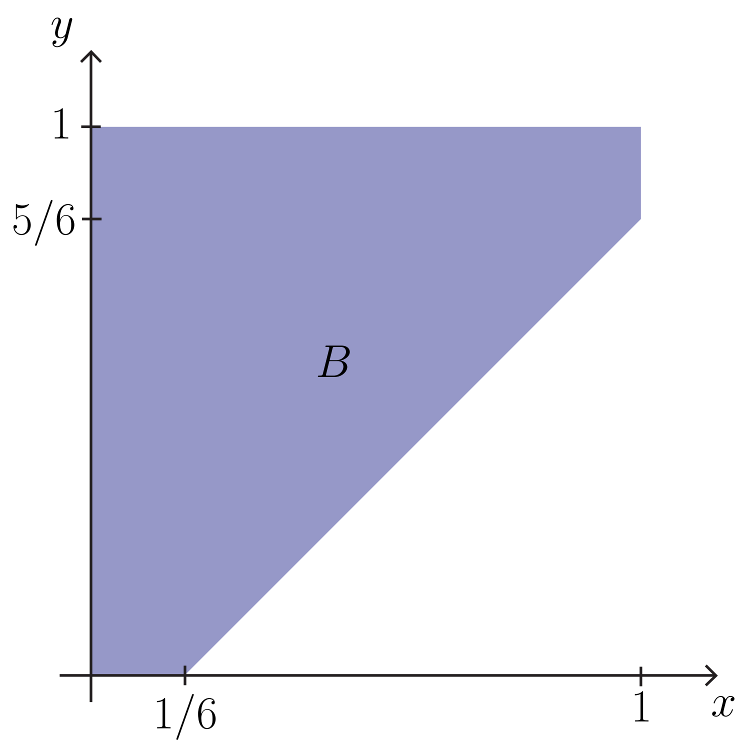
We model X and Y using the following joint pdf

$$f_{X,Y}(x,y) = \begin{cases} 1 & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{else} \end{cases}$$

What is the probability that Han arrives before or no more than 10 minutes after Chewbacca?

The event of interest is

$$B = \{(x,y) \mid x \le y + 1/6,\ 0 \le x \le 1,\ 0 \le y \le 1\}$$



and we see that

$$
\begin{aligned}
\mathrm{P}(B) &= \int_{x=0}^{1} \int_{y=\max(0,x-1/6)}^{1} 1 \, \mathrm{d}y \; \mathrm{d}x \\
&= 1 - \frac{1}{2}\left(\frac{5}{6}\right)^2 = 0.6528
\end{aligned}
$$

Exercise:
You are waiting for Alice and Bob to arrive. Let X be the time it takes Alice to arrive (in minutes) and Y be the time it takes Bob to arrive, and suppose X and Y have the joint pdf

$$
f_{X,Y}(x,y) = \begin{cases} 2\mathrm{e}^{-(x+2y)} & x, y \geq 0 \\ 0 & \text{otherwize} \end{cases}
$$

What is the probability that both Alice and Bob arrive in the next minute?

As with all pdfs, the joint pdf must be **normalized** in that it integrates to 1:

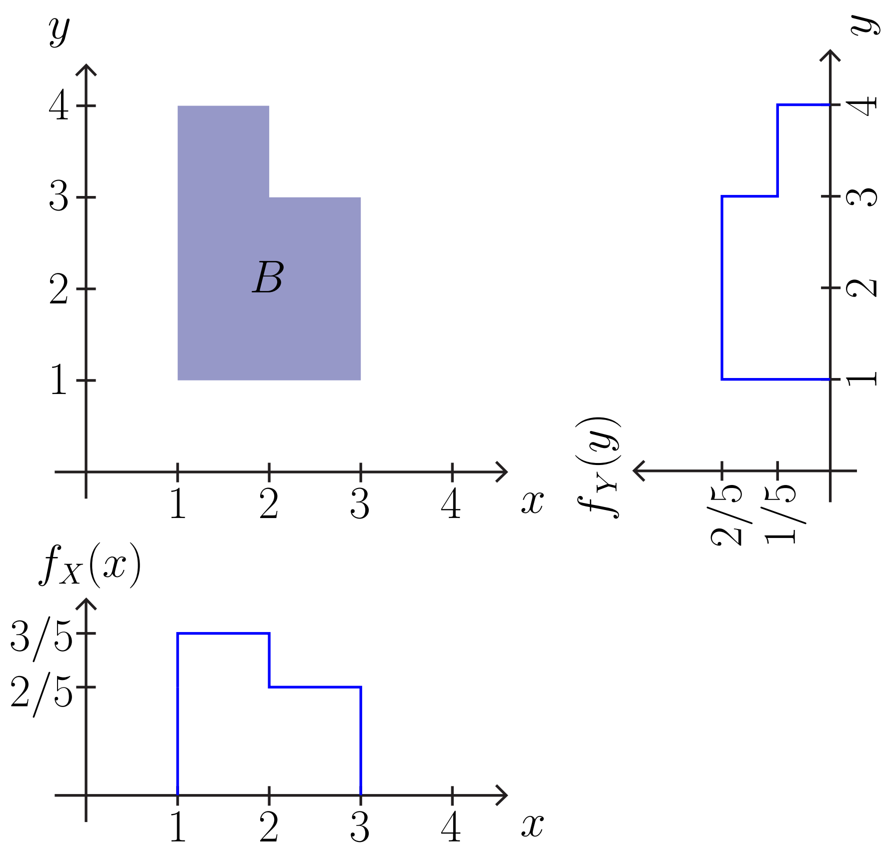
$$\mathrm{P}\left(-\infty \leq X \leq \infty, -\infty \leq Y \leq \infty\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1.$$

Also, as before, given the joint pdf $f_{X,Y}(x,y)$, we can compute the **marginal** pdfs for X and Y using

$$f_X(x) = \int_{y=-\infty}^{\infty} f_{X,Y}(x,y)\ \mathrm{d}y, \quad f_Y(y) = \int_{x=-\infty}^{\infty} f_{X,Y}(x,y)\ \mathrm{d}x.$$

Example. Consider (X,Y) with the joint pdf shown below, where

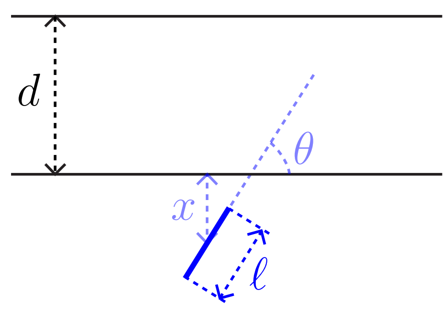
$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{5} & (x,y) \in B \\ 0 & (x,y) \notin B \end{cases}$$



Exercise (Buffon's Needle):
This is a classic problem, the answer to which was used by several people to estimate π.

A surface is ruled with parallel lines at a distance d from each other. Suppose we throw a needle of length ℓ on the surface at random. What is the probability the needle will intersect one of the lines? (Assume that $\ell < d$, so it never intersects more than one line.)

Hint: Let x be the distance from the middle of the needle to the closest line and let θ be the acute angle that the needle makes with the lines, and note that the needle crosses a line if $x \leq \frac{\ell}{2} \sin\theta$.



Joint CDFs

If X and Y are random variables associated with the same experiment, then their joint cdf is

$$\begin{aligned} F_{X,Y}(x,y) &= \mathrm{P}\left(X \leq x,\ Y \leq y\right) \\ &= \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\ \mathrm{d}s\ \mathrm{d}t \end{aligned}$$

Given the joint cdf, we can recover the joint pdf by differentiating:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

Expectation

The expectation of a function of two jointly continuous random variables generalizes as well:

$$\mathrm{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{X,Y}(x,y)\ \mathrm{d}x\ \mathrm{d}y$$

and as always

$$\mathrm{E}[aX + bY + c] = a\,\mathrm{E}[X] + b\,\mathrm{E}[Y] + c$$

Example. Suppose X and Y have joint pdf

$$f_{X,Y}(x,y) = \begin{cases} 1 & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\begin{aligned} \mathrm{E}[XY] = \int_0^1 \int_0^1 xy\, 1\, \mathrm{d}x\ \mathrm{d}y &= \int_0^1 y \left[\frac{x^2}{2}\right]_0^1 \mathrm{d}y \\ &= \frac{1}{2}\int_0^1 y\ \mathrm{d}y \\ &= \frac{1}{4} \end{aligned}$$

Exercise:
For X, Y distributed as above, what is $\mathrm{E}\left[\frac{X}{Y+1}\right]$?

More than two random variables

Three random variables X, Y, Z all associated with the same experiment can be characterized by a joint pdf with three arguments which satisfies

$$\mathrm{P}(B) = \iiint_{(x,y,z)\in B} f_{X,Y,Z}(x,y,z)\ \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z$$

for any event $B \subset \mathbb{R}^3$. For any function $g(x,y,z)$,

$$\mathrm{E}[g(X,Y,Z)] = \int\int\int g(x,y,z)\, f_{X,Y,Z}(x,y,z)\ \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z.$$

The extension to more that three random variables is conceptually the same: $X_1, X_2, \ldots, X_n$ are completely characterized by a positive function $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ of n variables such that

$$\mathrm{P}(B) = \int \underset{B}{\cdots} \int\ f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\ \mathrm{d}x_1\ \cdots \mathrm{d}x_n$$

etc.