Andrew Gard - equitable.equations@gmail.com

Bivariate Distributions of Continuous Random Variables

A distribution of two (or more) continuous random variables is defined by a **joint probability density function** much in the same way that the distribution of a single variable is.

A distribution of two (or more) continuous random variables is defined by a **joint probability density function** much in the same way that the distribution of a single variable is. A joint pdf $f(x.y)$ must satisfy two properties:

- $f(x, y) \geq 0$ for all x and y.
- $\displaystyle\int_{\mathbb{R}^2} f(x, y)\, dx\, dy = 1.$

A distribution of two (or more) continuous random variables is defined by a **joint probability density function** much in the same way that the distribution of a single variable is. A joint pdf $f(x.y)$ must satisfy two properties:

- $f(x, y) \geq 0$ for all x and y.
- $\displaystyle\int_{\mathbb{R}^2} f(x, y)\, dx\, dy = 1$.

Probabilities are calculated by integration. That is, for any region $D \subset \mathbb{R}^2$,

$$P\big((x, y) \in D\big) = \iint_D f(x, y)\, dA$$

A distribution of two (or more) continuous random variables is defined by a **joint probability density function** much in the same way that the distribution of a single variable is. A joint pdf $f(x.y)$ must satisfy two properties:

- $f(x, y) \geq 0$ for all x and y.
- $\displaystyle\int_{\mathbb{R}^2} f(x, y)\, dx\, dy = 1.$

Probabilities are calculated by integration. That is, for any region $D \subset \mathbb{R}^2$,

$$P\big((x, y) \in D\big) = \iint_D f(x, y)\, dA$$

The set of points (x, y) where $f(x, y) > 0$ is called the **support** of the distribution.

We can usually compute double integrals one at a time. This is called **iteration** and is allowed whenever the function $f(x, y)$ is continuous.

We can usually compute double integrals one at a time. This is called **iteration** and is allowed whenever the function $f(x, y)$ is continuous.

Most simply, if the region D is rectangular, given by $a \leq x \leq b$ and $c \leq y \leq d$, then

$$\begin{aligned}
\iint_D f(x,y)\,dA &= \int_{x=a}^{x=b} \int_{y=c}^{y=d} f(x,y)\,dy\,dx \\
&= \int_{y=c}^{y=d} \int_{x=a}^{x=b} f(x,y)\,dx\,dy
\end{aligned}$$

Example. Let X and Y be continuous random variables with probability density function $f(x,y) = \frac{1}{16}(x+3y)$ on $[0,2]\times[0,2]$. Compute $P\big((x,y)\in D\big)$ where $D = [0,1]\times[1,2]$.

Example. Let X and Y be continuous random variables with probability density function $f(x,y) = \frac{1}{16}(x+3y)$ on $[0,2]\times[0,2]$. Compute $P\big((x,y)\in D\big)$ where $D=[0,1]\times[1,2]$.

This amounts to computing

$$\iint_D \frac{1}{16}(x+3y)\,dA$$

For $D = [0,1] \times [1,2]$,

$$\iint_D \frac{1}{16}(x+3y)\,dA \;=\; \int_{x=0}^{x=1} \int_{y=1}^{y=2} \frac{1}{16}(x+3y)\,dy\,dx$$

For $D = [0,1] \times [1,2]$,

$$\begin{aligned}
\iint_D \frac{1}{16}(x+3y)\,dA &= \int_{x=0}^{x=1} \int_{y=1}^{y=2} \frac{1}{16}(x+3y)\,dy\,dx \\
&= \frac{1}{16} \int_{x=0}^{x=1} \left[yx + \frac{3}{2}y^2 \right]_{y=1}^{y=2} dx
\end{aligned}$$

For $D = [0,1] \times [1,2]$,

$$\begin{aligned}
\iint_D \frac{1}{16}(x+3y)\,dA &= \int_{x=0}^{x=1}\int_{y=1}^{y=2} \frac{1}{16}(x+3y)\,dy\,dx \\
&= \frac{1}{16}\int_{x=0}^{x=1}\left[yx + \frac{3}{2}y^2\right]_{y=1}^{y=2} dx \\
&= \frac{1}{16}\int_{x=0}^{x=1}\left(x + \frac{9}{2}\right) dx
\end{aligned}$$

For $D = [0,1] \times [1,2]$,

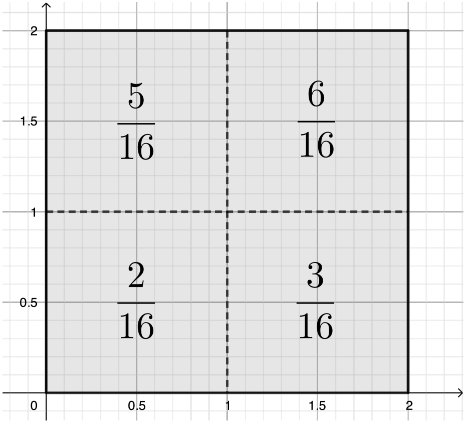
$$\begin{aligned}
\iint_D \frac{1}{16}(x+3y)\,dA &= \int_{x=0}^{x=1}\int_{y=1}^{y=2} \frac{1}{16}(x+3y)\,dy\,dx \\
&= \frac{1}{16}\int_{x=0}^{x=1}\left[yx + \frac{3}{2}y^2\right]_{y=1}^{y=2} dx \\
&= \frac{1}{16}\int_{x=0}^{x=1}\left(x + \frac{9}{2}\right) dx \\
&= \frac{1}{16}\left[\frac{1}{2}x^2 + \frac{9}{2}x\right]_0^1
\end{aligned}$$

For $D = [0,1] \times [1,2]$,

$$\begin{aligned}
\iint_D \frac{1}{16}(x+3y)\,dA &= \int_{x=0}^{x=1}\int_{y=1}^{y=2} \frac{1}{16}(x+3y)\,dy\,dx \\
&= \frac{1}{16}\int_{x=0}^{x=1}\left[yx + \frac{3}{2}y^2\right]_{y=1}^{y=2} dx \\
&= \frac{1}{16}\int_{x=0}^{x=1}\left(x + \frac{9}{2}\right) dx \\
&= \frac{1}{16}\left[\frac{1}{2}x^2 + \frac{9}{2}x\right]_0^1 \\
P\big((x,y)\in D\big) &= \frac{5}{16}
\end{aligned}$$

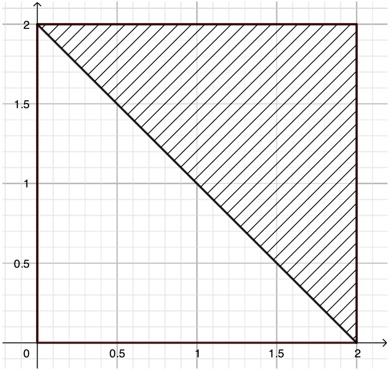
Notice that the region D is 1/4 the area of the region where $f(x, y)$ is nonzero, but includes 5/16 of the probability. The joint pdf $f(x, y) = \frac{1}{16}(x + 3y)$ is larger for larger values of x and y, so the probability is more concentrated in the upper-right of the support.

Notice that the region D is 1/4 the area of the region where $f(x, y)$ is nonzero, but includes 5/16 of the probability. The joint pdf $f(x, y) = \frac{1}{16}(x + 3y)$ is larger for larger values of x and y, so the probability is more concentrated in the upper-right of the support. The probabilities for each quarter are:

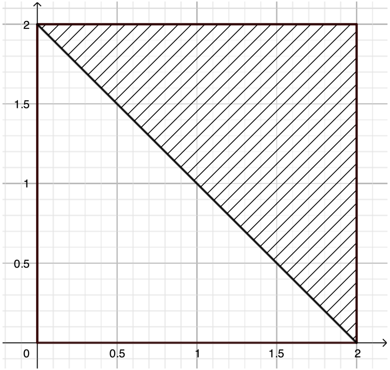


Things get a little more interesting when the region D isn't a rectangle.

Things get a little more interesting when the region D isn't a rectangle. For instance, we might want to compute the probability that $(x, y) \in D$ where D is the shaded triangle below.

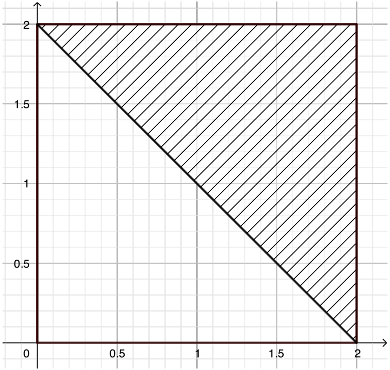


Things get a little more interesting when the region D isn't a rectangle. For instance, we might want to compute the probability that $(x, y) \in D$ where D is the shaded triangle below.

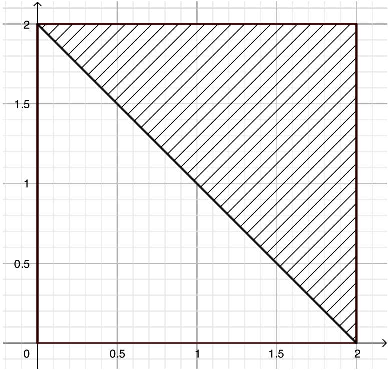


In order to write an iterated integral, we need to come up with inequalities describing the region.

In order to describe this region, we need two inequalities, one for x and one for y.

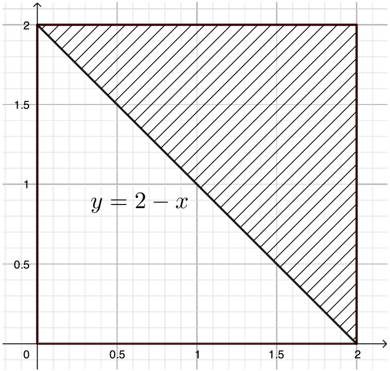


In order to describe this region, we need two inequalities, one for x and one for y.



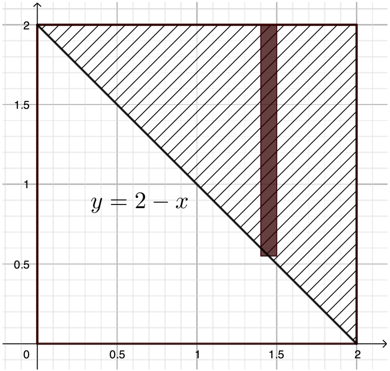
Whichever we write first should just capture the range of possible values of that variable. For instance, $0 \leq x \leq 2$.

The second inequality bounds the second variable between functions of the first.



When $0 \leq x \leq 2$, the shaded region is defined by $2 - x \leq y \leq 2$.

When writing the second inequality, it can be helpful to draw a rectangle as if setting up an integral for the area between curves.



This makes it more clear that the y-values are bounded by $y = 2 - x$ and $y = 2$.

We've managed to parameterize the region D.

$$\begin{aligned} 0 &\leq x \leq 2 \\ 2 - x &\leq y \leq 2 \end{aligned}$$

We're ready to set up an integral.

We've managed to parameterize the region D.

$$\begin{array}{rcccl} 0 & \leq & x & \leq & 2 \\ 2-x & \leq & y & \leq & 2 \end{array}$$

We're ready to set up an integral.

$$\int_{x=0}^{x=2} \int_{y=2-x}^{y=2} \frac{1}{16}(x+3y)\, dy\, dx$$

Notice that the variable with the constant limits goes on the outside and so will be evaluated second.

Here's the calculation:

$$\begin{aligned}
\int_{x=0}^{x=2}\int_{y=2-x}^{y=2}\frac{1}{16}(x+3y)\,dy\,dx &= \frac{1}{16}\int_{x=0}^{x=2}\left(xy+\frac{3}{2}y^2\right)_{y=2-x}^{y=2}dx \\
&= \frac{1}{16}\int_{x=0}^{x=2}\left(6x-\frac{1}{2}x^2\right)dx \\
&= \frac{1}{16}\left(3x^2-\frac{1}{6}x^3\right)_{x=0}^{x=2} \\
&= \frac{2}{3}
\end{aligned}$$

In general, this technique can be used whenever D can be represented as the region between two curves $y = g_1(x)$ and $y = g_2(x)$.

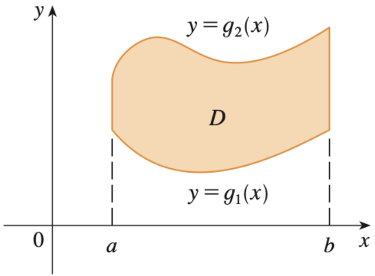
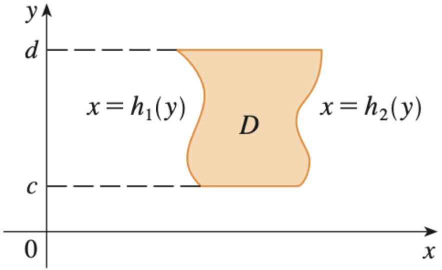


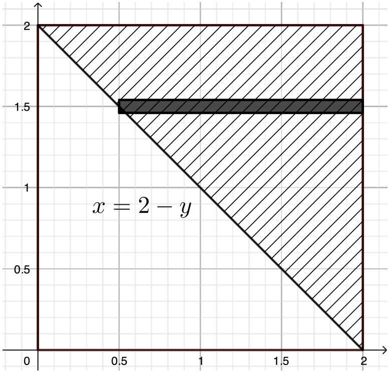
Image source: Stewart's *Calculus* 10e

The same technique works if D can be described as the region between two function of y.



In such a case, we parameterize the region as $c \leq y \leq d$, $h_1(y) \leq x \leq h_2(y)$ and integrate with respect to y first.

The region in the last example can also be described by bounding x between functions of y. You should try this for practice!



You answer should be the same as before: $P\big((x, y) \in A\big) = \frac{2}{3}$.