IV. Functions of Random Variables

Derived distributions

Let X be a continuous random variable with pdf $f_X(x)$, and let $Y = g(X)$ for some known and fixed function g.

What is the pdf of Y?

There is a general two step approach for figuring this out:

1. Calculate the cdf F_Y of Y:

$$F_Y(y) = \mathrm{P}\left(g(X) \leq y\right) = \int_{\{x|g(x)\leq y\}} f_X(x)\ \mathrm{d}x.$$

2. Differentiate to obtain the pdf of Y:

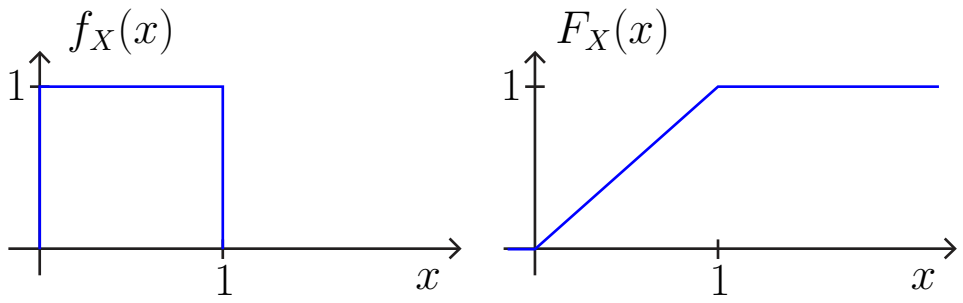
$$f_Y(y) = \frac{\mathrm{d}F_Y}{\mathrm{d}y}(y)$$

Like most things in this course, you will come to appreciate the general approach by looking at a few specific examples.

Example. Take

$$X \sim \text{Uniform}([0,1]),$$
$$Y = \sqrt{X}.$$

Here is what the pdf and cdf of X look like:

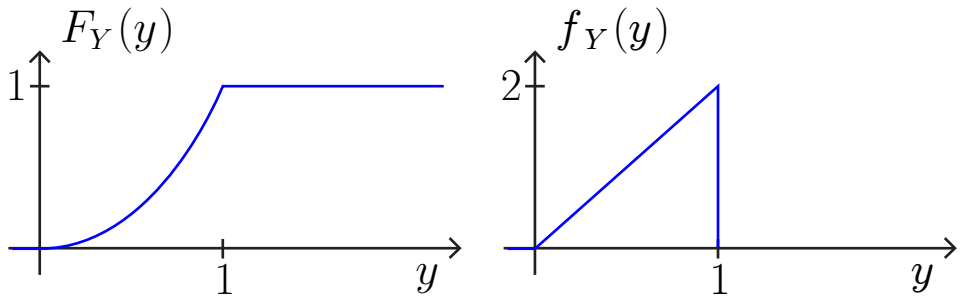


Proceeding through the two steps above, we calculate the cdf for Y as

$$\begin{aligned} F_Y(y) = \mathrm{P}\left(Y \leq y\right) &= \mathrm{P}\left(\sqrt{X} \leq y\right) \\ &= \mathrm{P}\left(X \leq y^2\right) \\ &= \begin{cases} 0, & y < 0 \\ y^2, & 0 \leq y \leq 1 \\ 1, & y > 1. \end{cases} \end{aligned}$$

Differentiating in each of the three regions yields the pdf

$$f_Y(y) = \begin{cases} 2y & 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$



Exercise:
Take

$$X \sim \text{Uniform}([0,1]),$$
$$Y = X^2.$$

Find the pdf for Y.

Exercise:
You are planning to drive from Atlanta to Charlotte, a distance of 240 miles, at a constant speed whose value is uniformly distributed between 50 and 70 miles per hour. What is the pdf of the duration of the trip? Let X be the speed, and set

$$g(X) = \frac{240}{X}.$$

Invertible (monotonic) functions

If $g(x)$ has a well-defined inverse over the range of X, then we can derive a general formula.

To formalize this, suppose that the range of X is contained in an interval I, meaning

$$f_X(x) = 0, \quad \text{for } x \notin I,$$

and suppose that $g(x)$ is **strictly monotonic** in that either

1. $g(x) < g(x')$ for all $x, x' \in I$ with $x < x'$, or
2. $g(x) > g(x')$ for all $x, x' \in I$ with $x < x'$.

Then g has a well-defined inverse over I. That is, there exists a function $h(y)$ such that for all $x \in I$,

$$y = g(x) \quad \text{if and only if} \quad x = h(y).$$

Two easy examples (assume $a \neq 0$ in both cases below):

$$g(x) = ax + b, \qquad h(y) = \frac{y - b}{a},$$

$$g(x) = \mathrm{e}^{ax}, \quad x \geq 0, \qquad h(y) = \frac{\ln y}{a}.$$

For g monotonically increasing (first case above), if $Y = g(X)$ then

$$\begin{aligned} F_Y(y) &= \mathrm{P}\left(g(X) \leq y\right) \\ &= \mathrm{P}\left(X \leq h(y)\right) \\ &= F_X(h(y)), \end{aligned}$$

and then using the chain rule for derivatives

$$\begin{aligned} f_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} &= \frac{\mathrm{d}F_X(h(y))}{\mathrm{d}y} \\ &= f_X(h(y)) \frac{\mathrm{d}h(y)}{\mathrm{d}y}. \end{aligned}$$

Similarly, if g is monotonically decreasing, then

$$\begin{aligned} F_Y(y) = \mathrm{P}\left(g(X) \leq y\right) &= 1 - \mathrm{P}\left(g(X) \geq y\right) \\ &= 1 - \mathrm{P}\left(X \leq h(y)\right) \\ &= 1 - F_X(h(y)), \end{aligned}$$

and

$$f_Y(y) = -f_X(h(y)) \frac{\mathrm{d}h(y)}{\mathrm{d}y}.$$

We combine the two statements above to get the general result for invertible g:

$$f_Y(y) = f_X(h(y)) \left| \frac{\mathrm{d}h(y)}{\mathrm{d}y} \right|$$

where h is the inverse of g.

Example. Let X be a continuous random variable with pdf

$$f_X(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Set $Y = \ln(X)$. Find the pdf for Y and sketch it.

First, note that for $0 \leq x \leq 1$, $-\infty < \ln(x) \leq 0$. Thus $f_Y(y)$ will be supported only on the negative half of the axis.

For $0 \leq x \leq 1$, $y = \ln(x) \Leftrightarrow x = \mathrm{e}^y$, so we can take $h(y) = \mathrm{e}^y$, and

$$\begin{aligned} f_Y(y) &= f_X(\mathrm{e}^y) \left| \frac{\mathrm{d}\mathrm{e}^y}{\mathrm{d}y} \right| \\ &= f_X(\mathrm{e}^y)\mathrm{e}^y \\ &= \mathrm{e}^y, \quad y \leq 0. \end{aligned}$$

Sketch:

Linear functions

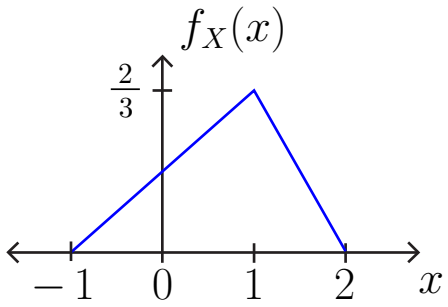
It is worth looking at the special case when $g(x)$ is linear:

$$Y = aX + b, \quad a, b \in \mathbb{R}, \quad a \neq 0.$$

Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).$$

Exercise:
Suppose that $f_X(x)$ is given by



Plot the pdf for Y in each of the following cases.

1. $Y = 2X$

2. $Y = -2X$

3. $Y = 3X + 2$

4. $Y = 3X - 2$

5. $Y = -3X + 2$

6. $Y = -3X - 2$

Things can get a little trickier when $g(x)$ is **not invertible** over the range of X.

For example, if X is a random variable on all of $\mathbb{R}$, then $g(x) = x^2$ is not invertible (since $a^2 = (-a)^2$).

But by *carefully* following the steps for the derived distribution, we can still get a pdf:

$$\begin{aligned} F_Y(y) &= \mathrm{P}\left(Y \leq y\right) \\ &= \mathrm{P}\left(X^2 \leq y\right) \\ &= \mathrm{P}\left(-\sqrt{y} \leq X \leq \sqrt{y}\right) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}), \end{aligned}$$

and then differentiating with the chain rule:

$$f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}), \quad y \geq 0.$$

Exercise:
Let $X \sim \text{Normal}(0, 1)$, and set $Y = X^2$. Calculate the pdf $f_Y(y)$ and sketch it.

Mapping Uniform Random Variables

Now that we have an understanding of derived distributions, we will look at an important special case.

If $X \sim \text{Uniform}([0,1])$, then there is a principled way to choose g so that $Z = g(X)$ has any distribution you like, continuous or discrete. This is especially handy when you want to **generate** random variables with a specified distribution — a good random number generator for the uniform distribution gives you a good random number generator for *any* distribution.

Suppose that we would like for Z to be a continuous random variable with pdf $f_{\text{target}}(z)$ and corresponding cdf $F_{\text{target}}(z)$. How should we choose g?

Answer: Take $g(x) = F^{-1}_{\text{target}}(x)$.

To see why this works, first suppose that $f_{\text{target}}(z) > 0$ for all $z \in \mathbb{R}$ (example: we would like Z to be normally distributed). Then $F_{\text{target}}(z)$ is *monotonically increasing* over the entire real line, and so F^{-1}_{target} is well-defined.

Now calculate the cdf for $Z = g(X) = F^{-1}_{\text{target}}(X)$:

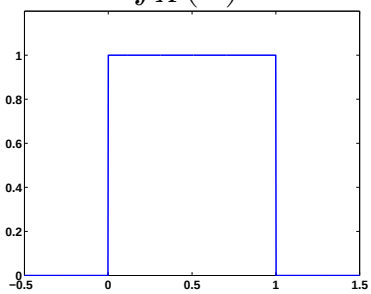
$$\begin{aligned} F_Z(z) = \mathrm{P}\left(Z \leq z\right) &= \mathrm{P}\left(F^{-1}_{\text{target}}(X) \leq z\right) \\ &= \mathrm{P}\left(X \leq F_{\text{target}}(z)\right) \\ &= F_{\text{target}}(z) \quad \text{for } \ 0 \leq F_{\text{target}}(z) \leq 1 \\ &= F_{\text{target}}(z) \quad \text{for all } z \in \mathbb{R}, \end{aligned}$$

where the last step simply follows from that fact that since it is a cdf, $F_{\text{target}}(z)$ is *always* between 0 and 1.

Example. Suppose we want $Z = g(X)$ to be Normal$(0, 1)$. To accomplish this, we choose $g(x) = \Phi^{-1}(x)$, where Φ is the cdf of a standard normal random variable.

To summarize: A random variable with this pdf (uniform):

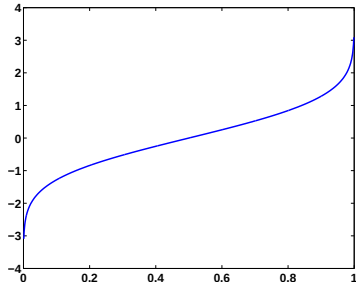
$f_X(x)$



$x \longrightarrow$

mapped to $Z = g(X)$ through this g

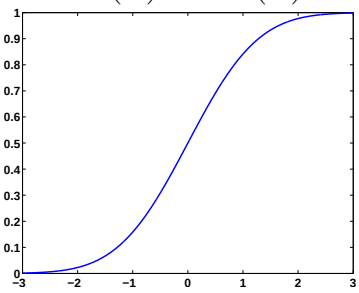
$g(x) = \Phi^{-1}(x)$



$x \longrightarrow$

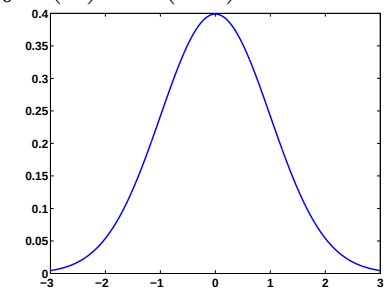
yields a random variable Z with this cdf and pdf:

$F_Z(z) = \Phi(z)$



$z \longrightarrow$

$f_Z(z) = (2\pi)^{-1/2} \mathrm{e}^{-z^2/2}$

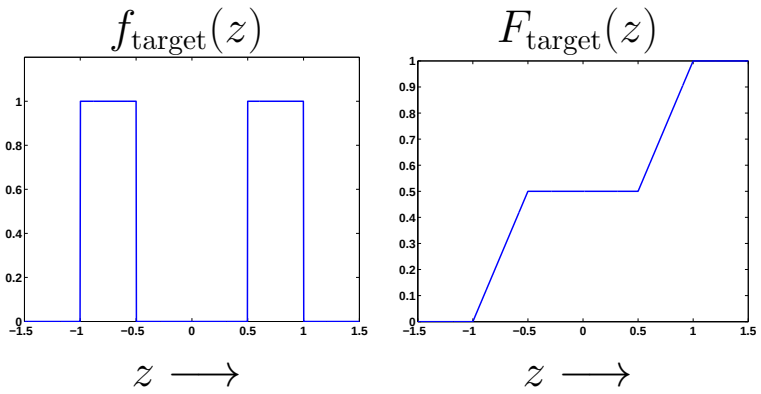


$z \longrightarrow$

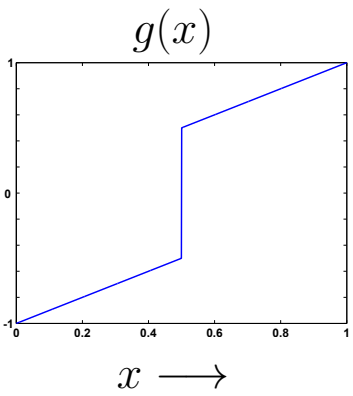
When there are intervals on the real line where $f_{\text{target}}(z) = 0$, then $F_{\text{target}}(z)$ will have plateaus (i.e., it will not be monotonically increasing), and so it is not invertible in the formal sense. These cases are still easily treated, we just need to make our definition of "inverse" more precise. Specifically, we will define

$$F^{-1}_{\text{target}}(z) = \min\{\ u\ : F_{\text{target}}(u) \geq z\}.$$

For example, suppose we want the target pdf and cdf shown below:



Then $g(x) = F^{-1}_{\text{target}}(z)$ will be discontinuous:



where at the discontinuity

$$g(0.5) = \min\{\ u\ : F_{\text{target}}(u) \geq 0.5\} = -0.5.$$

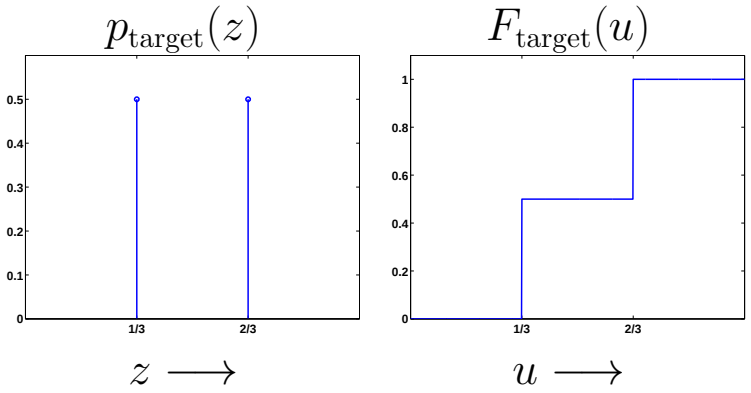
Exercise:
Suppose we would like $Z = g(X)$, where $X \sim \text{Uniform}([0,1])$, to be exponential:

$$f_{\text{target}}(z) = \begin{cases} \lambda \, \mathrm{e}^{-\lambda z}, & z \geq 0 \\ 0, & z < 0. \end{cases}$$

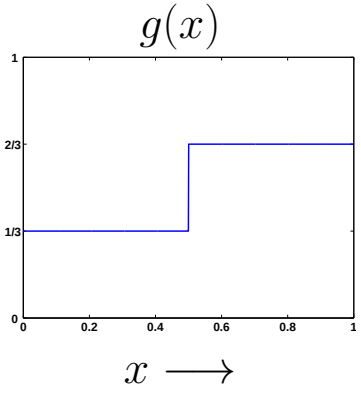
What should $g(x)$ be?

We can also map $X \sim \text{Uniform}([0,1])$ into a discrete random variable with a prescribed pmf. In this case, the target cdf is piecewise constant (i.e., it is all plateaus).

For example, suppose we want Z with the following pmf and cdf:



In this case, we take $g(x)$ to be the following discontinuous function:



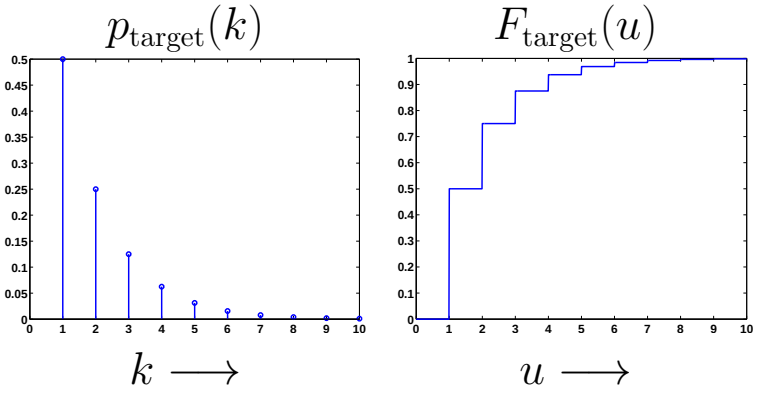
Which corresponds to taking

$$Z = \begin{cases} 1/3 & \text{if } 0 \leq X < 1/2 \\ 2/3 & \text{if } 1/2 \leq X < 1. \end{cases}$$

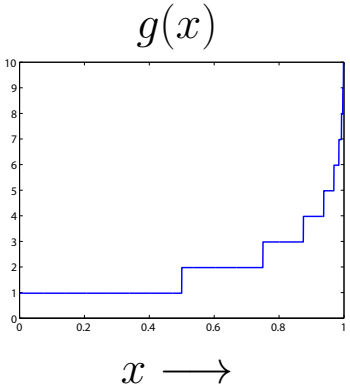
Example. Suppose we would like $Z = g(X)$, where $X \sim \text{Uniform}([0,1])$, to be a **geometric** discrete random variable with $p = 1/2$:

$$p_{\text{target}}(k) = \left(\frac{1}{2}\right)^k, \quad k \geq 1.$$

The target pmf and cdf look like:



The cdf above has an infinite number of plateaus, jumping up to close the distance towards 1 by a factor of two at every integer. In this case, $g(x)$ looks like



which we can write in functional form as

$$g(x) = k, \quad \text{for} \quad 1 - 2^{-k+1} \leq x \leq 1 - 2^{-k},$$

or

$$g(x) = \text{floor}\left(-\log_2(1-x)\right).$$

This is the same as taking a binary expansion of X, then taking as $g(X)$ the location of the first "1."

Functions of two independent random variables

Another common setting is where we wish to calculate the distribution of a function of multiple independent random variables. We can handle this case in much the same way as before by extending the trick of "computing the cdf then differentiating." This is best illustrated through examples.

Example. Suppose that X and Y are independent and uniformly distributed on $[0, 1]$. Let

$$Z = \max(X, Y).$$

Suppose that we would like to find the pdf of Z. We begin by computing the cdf of Z:

$$\begin{aligned} F_Z(z) &= \mathrm{P}\left(\max(X, Y) \leq z\right) \\ &= \mathrm{P}\left(\{X \leq z\} \cap \{Y \leq z\}\right) \\ &= \mathrm{P}\left(X \leq z\right) \cdot \mathrm{P}\left(Y \leq z\right) \qquad \text{(since } X, Y \text{ independent)} \\ &= \begin{cases} 0, & z < 0 \\ z^2, & 0 \leq z \leq 1 \\ 1, & z > 1. \end{cases} \end{aligned}$$

And so

$$f_Z(z) = \begin{cases} 2z, & 0 \leq z \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

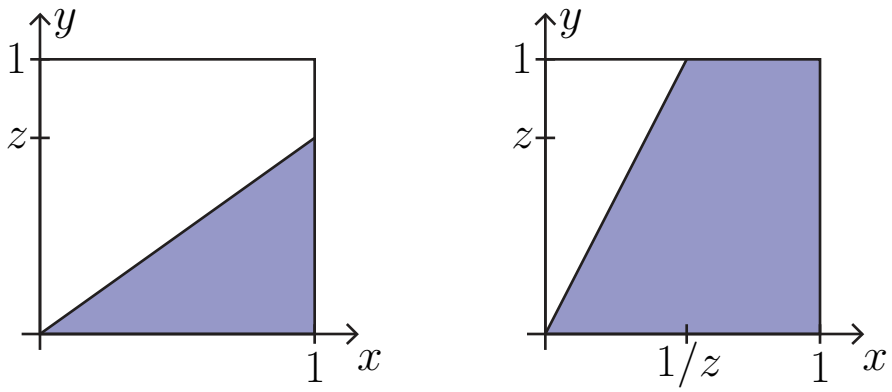
Example. Let X and Y be independent and uniformly distributed on $[0, 1]$. Consider

$$Z = \frac{Y}{X}.$$

We can find the cdf of Z with a geometrical argument. We will divide

$$F_Z(z) = \mathrm{P}\left(\frac{Y}{X} \leq z\right)$$

into two cases: $0 \leq z \leq 1$ and $z > 1$. From this figure:



we see that area in the shaded region corresponds to the area where $y/x \leq z$ and is equal to $z/2$ when $z \leq 1$ and $1 - 2/z$ when $z > 1$:

$$F_Z(z) = \mathrm{P}\left(\frac{Y}{X} \leq z\right) = \begin{cases} 0, & z < 0 \\ z/2, & 0 \leq z \leq 1 \\ 1 - \frac{1}{2z}, & z > 1. \end{cases}$$

Thus

$$f_Z(z) = \begin{cases} 0, & z < 0 \\ \frac{1}{2}, & 0 \leq z \leq 1 \\ \frac{1}{2z^2}, & z > 1. \end{cases}$$

Sums of independent random variables

Suppose that X and Y are independent random variables with pdfs $f_X(x)$ and $f_Y(y)$ and that we wish to find the pdf of

$$Z = X + Y.$$

To calculate the pdf of Z, let's begin by computing the *conditional* cdf of Z:

$$\begin{aligned} F_{Z|X}(z|x) &= \mathrm{P}\left(Z \leq z \;\middle|\; X = x\right) \\ &= \mathrm{P}\left(X + Y \leq z \;\middle|\; X = x\right) \\ &= \mathrm{P}\left(x + Y \leq z\right) \\ &= \mathrm{P}\left(Y \leq z - x\right) \\ &= F_Y(z - x), \end{aligned}$$

and so

$$f_{Z|X}(z|x) = \frac{\mathrm{d}F_{Z|X}}{\mathrm{d}z} = \frac{\mathrm{d}F_Y(z-x)}{\mathrm{d}z} = f_Y(z - x).$$

Now we can use the conditional pdf to find $f_Z(z)$:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{X,Z}(x, z)\, \mathrm{d}x \\ &= \int_{-\infty}^{\infty} f_{Z|X}(z|x)\, f_X(x)\, \mathrm{d}x \\ &= \int_{-\infty}^{\infty} f_Y(z - x)\, f_X(x)\, \mathrm{d}x \\ &= \text{CONVOLUTION!!!} \end{aligned}$$

Moral: Given independent random variables X, Y we can find the pdf for their sum $Z = X + Y$ by convolving their pdfs:

$$f_Z = f_X * f_Y.$$

Exercise:
Suppose that X and Y are independent with

$$X \sim \text{Uniform}([0,1]),$$
$$Y \sim \text{Uniform}([0,1]).$$

Find the pdf for $Z = X + Y$.

Exercise:
Suppose that X and Y are independent with

$$X \sim \text{Exp}(\lambda),$$
$$Y \sim \text{Exp}(\lambda).$$

Find the pdf for $Z = X + Y$.

Example. Suppose that X and Y are independent with

$$X \sim \text{Normal}(\mu_X, \sigma_X^2),$$
$$Y \sim \text{Normal}(\mu_Y, \sigma_Y^2).$$

If $Z = X + Y$ then the pdf of Z is given by

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_X} e^{-(x-\mu_X)^2/2\sigma_X^2} \cdot \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-((z-x)-\mu_Y)^2/2\sigma_Y^2} \, \mathrm{d}x.$$

Calculating this integral is a bit involved, but if you go through the details you will find that the integral reduces to

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(\sigma_X^2 + \sigma_Y^2)}} e^{-(z-(\mu_X+\mu_Y))^2/2(\sigma_X^2+\sigma_Y^2)},$$

thus

$$Z \sim \text{Normal}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2).$$

Observe that when we added two uniform random variables, or two exponential random variables, the distribution of the sum was something entirely different — but the sum of two normal random variables is itself another normal random variable. It turns out that the normal distribution is actually the only distribution that has this property. This hints at why the normal distribution is so central to the study of probability.