



Density functions
Math 217 Probability and Statistics
Prof. D. Joyce, Fall 2014

Today we'll look at the formal definition of a continuous random variable and define its density function.

The cumulative distribution function. Every real random variable X has a cumulative distribution function $F_X : \mathbf{R} \rightarrow [0, 1]$ defined as

$$F_X(b) = P(X \leq b).$$

Recall that if you know the c.d.f, then you can express the probability $P(a \leq X \leq b)$ in terms of it as

$$P(a \leq X \leq b) = F(b) - F(a).$$

In the case of a discrete random variable, F is constant except at countable many points where there are jumps. The jumps occur at numbers b where the probability is positive, and the size of the jumps are the values of the probability mass function $p(b)$, that is, the difference between $F(b)$ and the limit as x approaches b from the left of $F(x)$.

$$\begin{aligned} p(b) &= P(X=b) \\ &= P(x \leq b) - P(x < b) \\ &= F_X(b) - \lim_{x \rightarrow b^+} F_X(x) \end{aligned}$$

In the case of a continuous random variable, F has no jumps; it's continuous.

Definition of continuous random variables and their probability density functions. A continuous random variable is slightly more restrictive than just having a continuous p.d.f. We require

that F be the integral of a function f called the *probability density function*.

$$F(b) = \int_{-\infty}^b f(x) dx.$$

Such functions are called *absolutely continuous* functions. It follows that probabilities for X on intervals are the integrals of the density function:

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

By the Fundamental Theorem of Calculus, wherever F has a derivative, which it will almost everywhere, it will equal f . Typically F will be differentiable everywhere or perhaps everywhere except one or two points.

About the graphs of f and F . The cumulative distribution function F is an increasing function—perhaps it would be better to call it a nondecreasing function—since probability is accumulated as x increases. Indeed, as $x \rightarrow \infty$, F approaches 1. Also, as $x \rightarrow -\infty$, F approaches 0. Thus, the graph $y = F(x)$ is asymptotic on the left to $y = 0$ while it's asymptotic on the right to $y = 1$. In between F is increasing.

The probability density function f is a nonnegative function. Since it's the derivative of F , $f(x)$ is large when F is increasing rapidly, small when F is increasing slowly, and 0 on intervals where F is constant. The total area under the curve $y = f(x)$ is 1. That's equivalent to the statement that $\lim_{x \rightarrow \infty} F(x) = 1$. Both statements are true because the probability of an entire sample space S is 1.

A dart game example. For our first example, consider a dart game in which a dart is thrown at a circular target of radius 1. We assume it will be thrown uniformly randomly so that the probability that any particular region is hit is proportional to its area. We then record the distance X from the dart to the center of the target. We'll determine the cumulative distribution F for X , then differentiate it to get the density function f .

The distance from the thrown dart to the center of the target has to be between 0 and 1, so we know $X \in [0, 1]$. By definition $F(x) = P(X \leq x)$, that is, the dart lands within x units of the center. The region of the unit circle for which $X \leq x$ is a circle of radius x . It has area πx^2 . Since we're assuming uniform continuous probability on the unit circle, therefore, we can find $P(X \leq x)$ by dividing the area πx^2 of that circle by the total area π of the unit circle. Therefore, for $0 \leq x \leq 1$,

$$F(x) = P(X \leq x) = x^2.$$

For $x < 0$, $F(x) = 0$, and for $x > 1$, $F(x) = 1$.

To find the density function f , all we have to do is differentiate F . Then for $0 \leq x \leq 1$, $f(x) = F'(x) = 2x$. Outside the interval $[0, 1]$, f is 0.

Another example, one with unbounded density. When we discussed the axioms of probability, we looked at the probability density function and the c.d.f. for uniform probabilities on intervals.

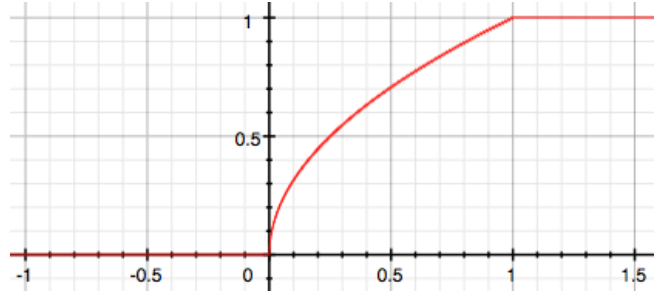
Now let Y be the square of a uniform distribution on $[0, 1]$. In other words, let X have a uniform distribution on $[0, 1]$, so that $f_X(x) = 1$ and $F_X(x) = x$ for $x \in [0, 1]$. Then let $Y = X^2$. Now to figure out what $F_Y(y)$ and $f_Y(y)$ are. We'll start with F_Y .

By the definition of F_Y , we have $F_Y(y) = P(Y \leq y) = P(X^2 \leq y)$. If y is negative of 0, then $P(X^2 \leq y) = 0$, so $F_Y(y) = 0$.

Now suppose $0 < y < 1$. Then $P(X^2 \leq y) = P(0 \leq X \leq \sqrt{y})$, which, since X is uniform, is equal to the length of the interval, \sqrt{y} .

Finally, if $1 \leq y$, then $F_Y(y) = 1$. Thus we have this formula for F_Y

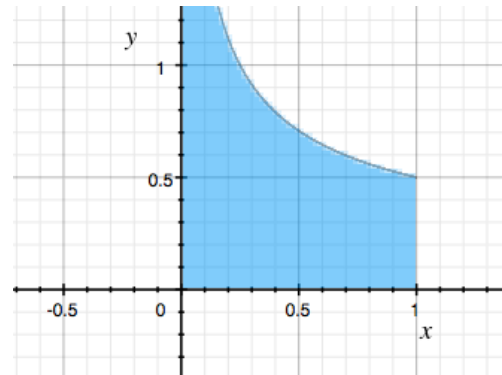
$$F_Y(y) = \begin{cases} 0 & \text{if } y \leq 0 \\ \sqrt{y} & \text{if } 0 \leq y \leq 1 \\ 1 & \text{if } 1 \leq y \end{cases}$$



We can differentiate F_Y to get the density function f_Y . Since F_Y is constant on the intervals $(-\infty, 0)$ and $(1, \infty)$, therefore f_Y will be 0 on those intervals. On the interval $[0, 1]$ the derivative of $F_Y(y)$ is $1/\sqrt{y}$. Therefore, f_y is given by the formula

$$f_Y(y) = \frac{1}{\sqrt{y}} \text{ on } [0, 1].$$

Usually the intervals where a probability density function is 0 aren't mentioned.



When I graph a density function, I usually shade in the region below the curve as a reminder that it's the area below the curve that's important. Notice that this density function approaches infinity at 0. In other words, its integral is improper, but the area under the curve is still 1.

For our next example, we'll let Z be the sum of two independent uniform distributions on $[0, 1]$. That means X and Y are both uniform on $[0, 1]$ so that (X, Y) is uniform on the unit square $[0, 1] \times [0, 1]$. In other words, the point (X, Y) is a uniformly randomly chosen point in the unit square. Then let Z be the sum of these two numbers, $Z = X + Y$. Of course, Z is some number between 0 and 2. We'll figure out what $F_Z(z)$ and $f_Z(z)$ are.

Functions of random variables. Frequently, one random variable Y is a function ϕ of another random variable X , that is, $Y = \phi(X)$. Given the cumulative distribution function $F_X(x)$ and the probability density function $f_X(x)$ for X , and we want to determine the cumulative distribution function $F_Y(y)$ and the probability density function $f_Y(y)$ for Y . How do we find them?

Let's illustrate this with an example. Suppose X is a uniform distribution on $[0, 10]$ and $Y = \phi(X) = X^2$. The probability density function for X is $f_X(x) = \frac{1}{10}$ for $x \in [0, 10]$, and the cumulative distribution function for X is

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0, \\ x/10 & \text{for } 0 \leq x \leq 10, \\ 1 & \text{for } x > 10 \end{cases}$$

Can we find $F_Y(y)$? Well,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(X^2 \leq y) \\ &= P(X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}). \end{aligned}$$

Thus,

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 0, \\ \sqrt{y}/10 & \text{for } 0 \leq y \leq 100, \\ 1 & \text{for } y > 100 \end{cases}$$

In general, if ϕ is an increasing function on the domain of F_X , then we can do the same thing. It will be the case that

$$F_Y(y) = F_X(\phi^{-1}(y))$$

where ϕ^{-1} is the function inverse to ϕ . So, it's pretty easy to find the cumulative distribution function for Y from that for X .

Of course, we can find the probability density function $f_Y(y)$ by differentiating $F_Y(y)$. In our example we find

$$f_Y(y) = \frac{1}{20\sqrt{y}} \quad \text{for } y \in [0, 100].$$

Can we find f_Y directly from f_X instead of going through the cumulative distribution function? Yes, at least when ϕ is an increasing function. We've seen that $F_Y(y) = F_X(\phi^{-1}(y))$. Differentiate both sides of that equation with respect to y , using the chain rule on the way. You get

$$F'_Y(y) = F'_X(\phi^{-1}(y)) \frac{d}{dy} \phi^{-1}(y).$$

Therefore,

$$f_Y(y) = f_X(\phi^{-1}(y)) \frac{d}{dy} \phi^{-1}(y).$$

How does that work in our example where $y = \phi(x) = x^2$? The inverse function is $\phi^{-1}(y) = \sqrt{y}$, so $\frac{d}{dy} \phi^{-1}(y) = \frac{1}{2\sqrt{y}}$. Therefore,

$$\begin{aligned} f_Y(y) &= f_X(\phi^{-1}(y)) \frac{d}{dy} \phi^{-1}(y) \\ &= \frac{1}{10} \frac{1}{2\sqrt{y}} = \frac{1}{20\sqrt{y}} \end{aligned}$$

for $0 \leq y \leq 100$.

The Cauchy distribution. This distribution isn't as important as some of the others we've looked at, but it does allow us to use the results we just found. Also, it is a very unusual distribution that we'll use later for theoretical purposes.

Imagine a spinner located at the origin of \mathbf{R}^2 . Draw the vertical line $x = 1$ and look at where on that line the spinner is pointing. Ignore the half of the time the spinner points to the left, and just use the times it points to the right. In other words, the spinner uniformly selects an angle X between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. Let Y be the coordinate on the line $x = 1$ where the spinner points. Then $Y = \tan(X)$. The distribution for Y is called the *Cauchy distribution*.

Let's determine the density for Y . First, note that the density for X is $f_X(x) = \frac{1}{\pi}$ for $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. The density for Y is $f_Y(y) = f_X(\phi^{-1}(y)) \frac{d}{dy} \phi^{-1}(y)$, where $\phi(x) = \tan x$. So $\phi^{-1}(y) = \arctan y$, and that has the derivative

$$\frac{d}{dy} \arctan y = \frac{1}{1 + y^2}.$$

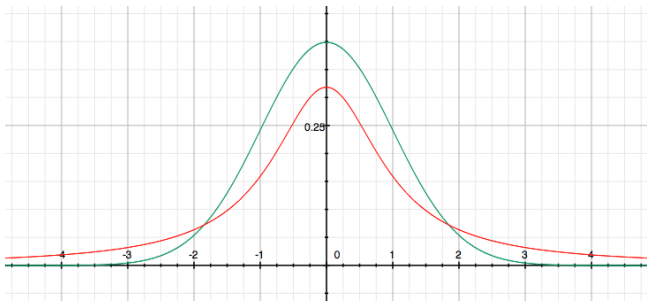
Therefore,

$$f_Y(y) = \frac{1}{\pi} \frac{1}{1+y^2}.$$

The graph of f_Y is a very wide bell-shaped curve. The cumulative density function for the Cauchy distribution is

$$F_Y(y) = \frac{1}{\pi} \arctan y.$$

In the figure below, the graph of the standard normal density function is green and that of the Cauchy density function is red.



What makes a Cauchy random variable so interesting is that the values that it takes on swing so wildly. It takes on values very far from 0 often enough so that the Cauchy distribution has no mean, no variance or standard deviation, doesn't satisfy the law of large numbers, and doesn't satisfy the conclusion of the central limit theorem. When we look at those concepts, we'll show how they don't work for the Cauchy distribution.

Math 217 Home Page

at <http://math.clarku.edu/~djoyce/ma217/>