Continuous Random Variables
Class 5, 18.05
Jeremy Orloff and Jonathan Bloom

1 Learning Goals

1. Know the definition of a continuous random variable.
2. Know the definition of the probability density function (pdf) and cumulative distribution function (cdf).
3. Be able to explain why we use probability density for continuous random variables.

2 Introduction

We now turn to continuous random variables. All random variables assign a number to each outcome in a sample space. Whereas discrete random variables take on a discrete set of possible values, continuous random variables have a continuous set of values.

Computationally, to go from discrete to continuous we simply replace sums by integrals. It will help you to keep in mind that (informally) an integral is just a continuous sum.

Example 1. Since time is continuous, the amount of time Jon is early (or late) for class is a continuous random variable. Let's go over this example in some detail.

Suppose you measure how early Jon arrives to class each day (in units of minutes). That is, the outcome of one trial in our experiment is a time in minutes. We'll assume there are random fluctuations in the exact time he shows up. Since in principle Jon could arrive, say, 3.43 minutes early, or 2.7 minutes late (corresponding to the outcome -2.7), or at any other time, the sample space consists of all real numbers. So the random variable which gives the outcome itself has a continuous range of possible values.

It is too cumbersome to keep writing 'the random variable', so in future examples we might write: Let T = "time in minutes that Jon is early for class on any given day."

3 Calculus Warmup

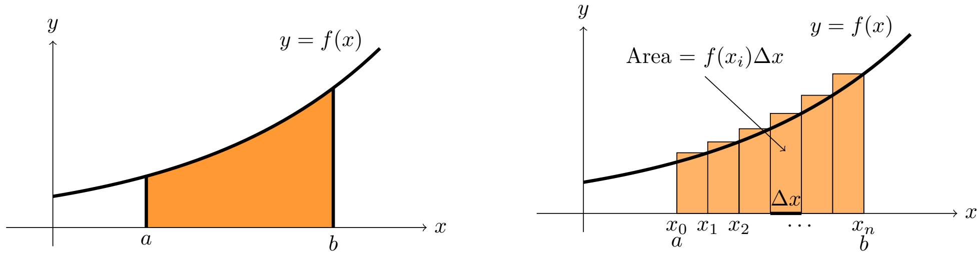
While we will assume you can compute the most familiar forms of derivatives and integrals by hand, we do not expect you to be calculus whizzes. For tricky expressions, we'll let the computer do most of the calculating. Conceptually, you should be comfortable with two views of a definite integral.

1. $\int_a^b f(x)\,dx$ = area under the curve $y = f(x)$.
2. $\int_a^b f(x)\,dx$ = 'sum of $f(x)\,dx$'.

The connection between the two is:

$$\text{area} \approx \text{sum of rectangle areas} \;=\; f(x_1)\Delta x + f(x_2)\Delta x + \ldots + f(x_n)\Delta x = \sum_1^n f(x_i)\Delta x.$$

As the width Δx of the intervals gets smaller the approximation becomes better.



Area is approximately the sum of rectangles

Note: In calculus you learned to compute integrals by finding antiderivatives. This is important for calculations, but don't confuse this method for the reason we use integrals. Our interest in integrals comes primarily from its interpretation as a 'sum' and to a much lesser extent its interpretation as area.

4 Continuous Random Variables and Probability Density Functions

A continuous random variable takes a range of values, which may be finite or infinite in extent. Here are a few examples of ranges: $[0, 1]$, $[0, \infty)$, $(-\infty, \infty)$, $[a, b]$.

Definition: A random variable X is continuous if there is a function $f(x)$ such that for any $c \leq d$ we have

$$P(c \leq X \leq d) \;=\; \int_c^d f(x)\,dx. \tag{1}$$

The function $f(x)$ is called the probability density function (**pdf**).

The pdf always satisfies the following properties:

1. $f(x) \geq 0$ (f is nonnegative).
2. $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$ (This is equivalent to: $P(-\infty < X < \infty) = 1$).

The probability density function $f(x)$ of a continuous random variable is the analogue of the probability mass function $p(x)$ of a discrete random variable. Here are two important differences:

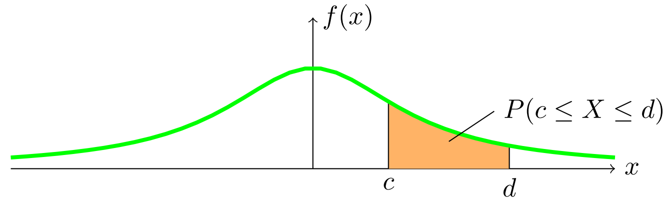
1. Unlike $p(x)$, the pdf $f(x)$ is *not* a probability. You have to integrate it to get probability. (See section 4.2 below.)
2. Since $f(x)$ is not a probability, there is no restriction that $f(x)$ be less than or equal to 1.

Note: In Property 2, we integrated over $(-\infty, \infty)$ since we did not know the range of values taken by X. Formally, this makes sense because we just define $f(x)$ to be 0 outside of the range of X. In practice, we would integrate between bounds given by the range of X.

4.1 Graphical View of Probability

If you graph the probability density function of a continuous random variable X then

$$\boxed{P(c \le X \le d) = \text{area under the graph between } c \text{ and } d.}$$



Think: What is the total area under the pdf $f(x)$?

4.2 The terms 'probability mass' and 'probability density'

Why do we use the terms mass and density to describe the pmf and pdf? What is the difference between the two? The simple answer is that these terms are completely analogous to the mass and density you saw in physics and calculus. We'll review this first for the probability mass function and then discuss the probability density function.

Mass as a sum:

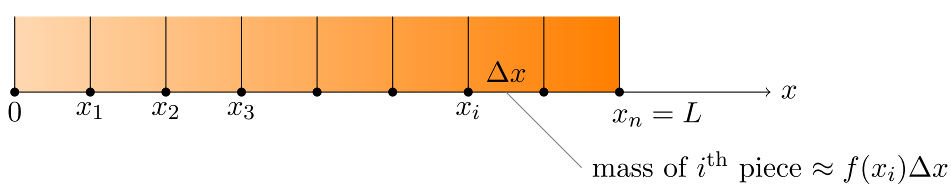
If masses m_1, m_2, m_3, and m_4 are set in a row at positions x_1, x_2, x_3, and x_4, then the total mass is $m_1 + m_2 + m_3 + m_4$.

We can define a 'mass function' $p(x)$ with $p(x_j) = m_j$ for $j = 1, 2, 3, 4$, and $p(x) = 0$ otherwise. In this notation the total mass is $p(x_1) + p(x_2) + p(x_3) + p(x_4)$.

The probability mass function behaves in exactly the same way, except it has the dimension of probability instead of mass.

Mass as an integral of density:

Suppose you have a rod of length L meters with varying density $f(x)$ kg/m. (Note the units are mass/length.)



If the density varies continuously, we must find the total mass of the rod by integration:

$$\textbf{total mass} \; = \int_0^L f(x)\, dx.$$

This formula comes from dividing the rod into small pieces and 'summing' up the mass of each piece. That is:

$$\text{total mass} \; \approx \sum_{i=1}^{n} f(x_i)\, \Delta x$$

In the limit as Δx goes to zero the sum becomes the integral.

The probability density function behaves exactly the same way, except it has units of probability/(unit x) instead of kg/m. Indeed, equation (1) is exactly analogous to the above integral for total mass.

While we're on a physics kick, note that for both discrete and continuous random variables, the expected value is simply the center of mass or balance point.

Example 2. Suppose X has pdf $f(x) = 3$ on $[0, 1/3]$ (this means $f(x) = 0$ outside of $[0, 1/3]$). Graph the pdf and compute $P(.1 \le X \le .2)$ and $P(.1 \le X \le 1)$.

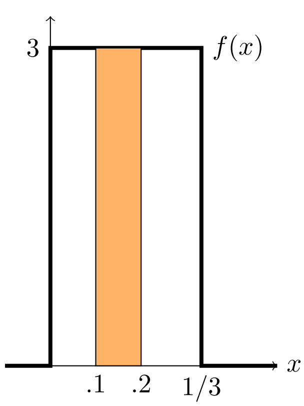
answer: $P(.1 \le X \le .2)$ is shown below at left. We can compute the integral:

$$P(.1 \le X \le .2) = \int_{.1}^{.2} f(x)\, dx = \int_{.1}^{.2} 3\, dx = .3.$$

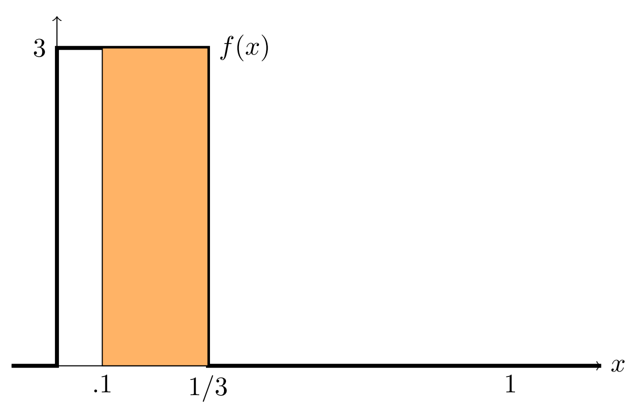
Or we can find the area geometrically:

$$\text{area of rectangle} \; = 3 \cdot .1 = .3.$$

$P(.1 \le X \le 1)$ is shown below at right. Since there is only area under $f(x)$ up to 1/3, we have $P(.1 \le X \le 1) = 3 \cdot (1/3 - .1) = .7$.



$P(.1 \le X \le .2)$



$P(.1 \le X \le 1)$

Think: In the previous example $f(x)$ takes values greater than 1. Why does this not violate the rule that probabilities are always between 0 and 1?

Note on notation. We can define a random variable by giving its range and probability density function. For example we might say, let X be a random variable with range [0,1] and pdf $f(x) = x/2$. Implicitly, this means that X has no probability density outside of the given range. If we wanted to be absolutely rigorous, we would say explicitly that $f(x) = 0$ outside of [0,1], but in practice this won't be necessary.

Example 3. Let X be a random variable with range [0,1] and pdf $f(x) = Cx^2$. What is the value of C?

answer: Since the total probability must be 1, we have

$$\int_0^1 f(x)\,dx = 1 \quad \Leftrightarrow \quad \int_0^1 Cx^2\,dx = 1.$$

By evaluating the integral, the equation at right becomes

$$C/3 = 1 \;\Rightarrow\; \boxed{C = 3}.$$

Note: We say the constant C above is needed to normalize the density so that the total probability is 1.

Example 4. Let X be the random variable in the Example 3. Find $P(X \leq 1/2)$.

answer: $P(X \leq 1/2) = \displaystyle\int_0^{1/2} 3x^2\,dx = x^3\Big|_0^{1/2} = \boxed{\frac{1}{8}.}$

Think: For this X (or any continuous random variable):

- What is $P(a \leq X \leq a)$?
- What is $P(X = 0)$?
- Does $P(X = a) = 0$ mean that X can never equal a?

In words the above questions get at the fact that the probability that a random person's height is exactly 5'9" (to infinite precision, i.e. no rounding!) is 0. Yet it is still possible that someone's height is exactly 5'9". So the answers to the thinking questions are 0, 0, and No.

4.3 Cumulative Distribution Function

The cumulative distribution function (**cdf**) of a continuous random variable X is defined in exactly the same way as the cdf of a discrete random variable.

$$F(b) \,=\, P(X \leq b).$$

Note well that the definition is about probability. When using the cdf you should first think of it as a probability. Then when you go to calculate it you can use

$$F(b) \,=\, P(X \leq b) \,=\, \int_{-\infty}^{b} f(x)\,dx, \quad \text{where } f(x) \text{ is the pdf of } X.$$

Notes:
1. For discrete random variables, we defined the cumulative distribution function but did

not have much occasion to use it. The cdf plays a far more prominent role for continuous random variables.

2. As before, we started the integral at $-\infty$ because we did not know the precise range of X. Formally, this still makes sense since $f(x) = 0$ outside the range of X. In practice, we'll know the range and start the integral at the start of the range.

3. In practice we often say 'X has distribution $F(x)$' rather than 'X has cumulative distribution function $F(x)$.'

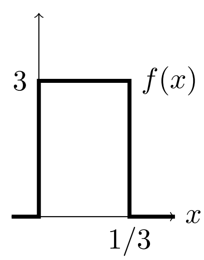
Example 5. Find the cumulative distribution function for the density in Example 2.

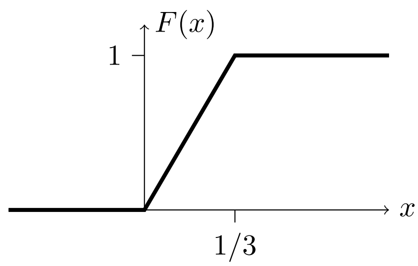
answer: For a in [0,1/3] we have $F(a) = \int_0^a f(x)\,dx = \int_0^a 3\,dx = 3a.$

Since $f(x)$ is 0 outside of [0,1/3] we know $F(a) = P(X \le a) = 0$ for $a < 0$ and $F(a) = 1$ for $a > 1/3$. Putting this all together we have

$$F(a) = \begin{cases} 0 & \text{if } a < 0 \\ 3a & \text{if } 0 \le a \le 1/3 \\ 1 & \text{if } 1/3 < a. \end{cases}$$

Here are the graphs of $f(x)$ and $F(x)$.





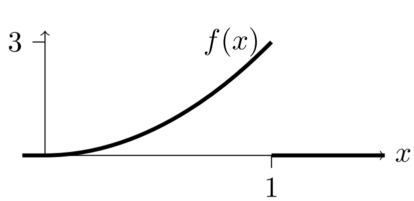
Note the different scales on the vertical axes. Remember that the vertical axis for the pdf represents probability density and that of the cdf represents probability.

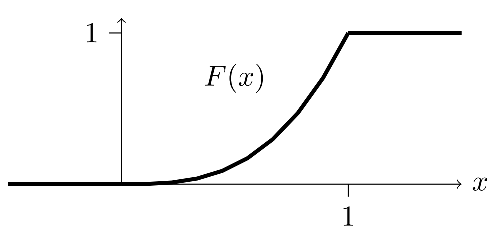
Example 6. Find the cdf for the pdf in Example 3, $f(x) = 3x^2$ on $[0, 1]$. Suppose X is a random variable with this distribution. Find $P(X < 1/2)$.

answer: $f(x) = 3x^2$ on [0,1] $\Rightarrow$ $F(a) = \int_0^a 3x^2\,dx = a^3$ on [0,1]. Therefore,

$$F(a) = \begin{cases} 0 & \text{if } a < 0 \\ a^3 & \text{if } 0 \le a \le 1 \\ 1 & \text{if } 1 < a \end{cases}$$

Thus, $P(X < 1/2) = F(1/2) = 1/8$. Here are the graphs of $f(x)$ and $F(x)$:





4.4 Properties of cumulative distribution functions

Here is a summarry of the most important properties of cumulative distribution functions (cdf)

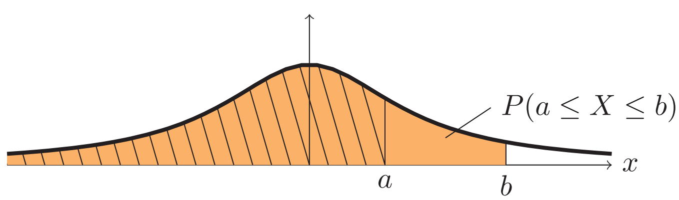
1. (Definition) $F(x) = P(X \le x)$
2. $0 \le F(x) \le 1$
3. $F(x)$ is non-decreasing, i.e. if $a \le b$ then $F(a) \le F(b)$.
4. $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$
5. $P(a \le X \le b) = F(b) - F(a)$
6. $F'(x) = f(x)$.

Properties 2, 3, 4 are identical to those for discrete distributions. The graphs in the previous examples illustrate them.

Property 5 can be seen algebraically:

$$\begin{aligned} \int_{-\infty}^{b} f(x)\,dx &= \int_{-\infty}^{a} f(x)\,dx + \int_{a}^{b} f(x)\,dx \\ \Leftrightarrow \int_{a}^{b} f(x)\,dx &= \int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx \\ \Leftrightarrow P(a \le X \le b) &= F(b) - F(a). \end{aligned}$$

Property 5 can also be seen geometrically. The orange region below represents $F(b)$ and the striped region represents $F(a)$. Their difference is $P(a \le X \le b)$.



Property 6 is the fundamental theorem of calculus.

4.5 Probability density as a dartboard

We find it helpful to think of sampling values from a continuous random variable as throwing darts at a funny dartboard. Consider the region underneath the graph of a pdf as a dartboard. Divide the board into small equal size squares and suppose that when you throw a dart you are equally likely to land in any of the squares. The probability the dart lands in a given region is the fraction of the total area under the curve taken up by the region. Since the total area equals 1, this fraction is just the area of the region. If X represents the x-coordinate of the dart, then the probability that the dart lands with x-coordinate between a and b is just

$$P(a \le X \le b) \quad = \quad \text{area under } f(x) \text{ between } a \text{ and } b \quad = \quad \int_{a}^{b} f(x)\,dx.$$