STAT 345 $\diamond$ CONTINUOUS RANDOM VARIABLES

KELLIN RUMSEY $\diamond$ SPRING 2020

Let X be a RV with range $\mathbb{X}$. Recall that X is called *discrete* if the range is finite or countable. We say that X is a *continuous* RV if the range is uncountable, which is to say that $\mathbb{X}$ contains an interval of real numbers. For continuous RVs, the probability mass function

$$p(x) = P(X = x)$$

is no longer a useful/meaningful function. Since the range of X is uncountable, we cannot even list out all of the possible values, so there is no way to assign probabilities so that they sum to 1. Instead, we will begin our discussion with a similar, but different function.

Definition. The Cumulative Distribution Function (CDF) of a RV X is the function $F(x)$ which satisfies.

$$F(x) = P(X \leq x)$$

A CDF is said to be *valid* if it satisfies the following properties.

- $\lim_{x\to-\infty} F(x) = 0$

  This property basically says that $P(X \leq -\infty) = 0$.

- $\lim_{x\to\infty} F(x) = 1$

  This property basically says that $P(X \leq \infty) = 1$.

- Monotonicity: If $x \geq y$, then $F(x) \geq F(y)$.

  This property says that the function $F(x)$ is never "decreasing", as x gets bigger

---

We won't spend too much time on this, but note that these properties are designed to be equivalent to the axioms of probability (i. probabilities are positive, ii. probability of $\mathcal{S}$ is one, iii) probability of the union of disjoint events is the sum of the probabilities). Any function $F(x)$ that satisfies these three properties is a valid CDF.

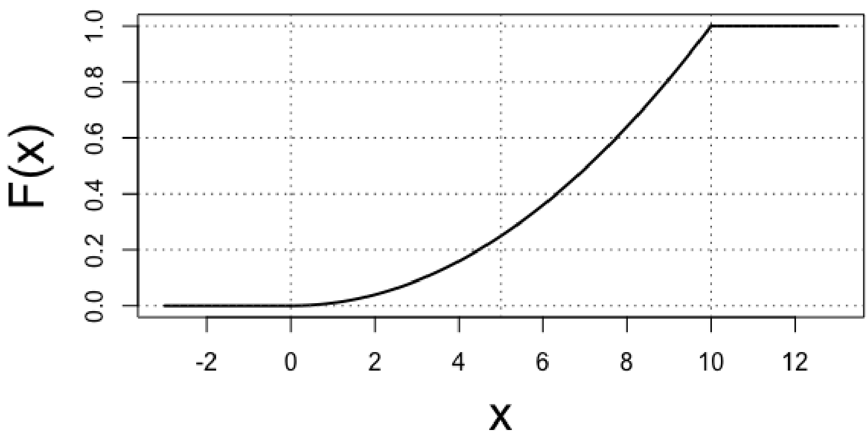

FIGURE 1. CDF for example 1.

Example 1.

1. Let X be a random variable with CDF.

$$F_X(x) = \begin{cases} 0, & x \leq 0 \\ \frac{x^2}{100}, & 0 < x \leq 10 \\ 1, & x > 10 \end{cases}$$

To see that this CDF is valid, we should check the three properties above. Sketching the CDF (as in Figure 1) shows that the three properties are satisfied.

i) What is the probability that X is less than or equal to 7?

$$P(X \leq 7) = F_X(7) = 7^2/100 = 0.49$$

ii) What is the probability that X is greater than 4?

$$P(X > 4) = 1 - P(X \leq 4) = 1 - F_X(4) = 1 - 4^2/100 = 0.84$$

iii) What is the probability that X is less than or equal to 15?

$$P(X \leq 15) = F_X(15) = 1$$

Next, we will look at some important features of continuous variables that result from the definition of the CDF.

A useful formula:

$$P(a < X \leq b) = F(b) - F(a) \tag{1}$$

This is easy to prove using basic probability rules. Let A be the event that $X \leq a$ and let B be the event that $X \leq b$. Then $B \cap A^c$ is the event that $a < X \leq b$, and (by the subtraction rule) we have that

$$P(a < X \leq b) = P(B \cap A^c) = P(B) - P(B \cap A) = P(B) - P(A) = F(b) - F(a).$$

Claim: If X is a continuous RV, then $P(X = x) = 0$ for every x.

This is why the PMF is not meaningful for continuous RVs. The basic issue is, there are too many possible values to assign probabilities to every single x in a way that they still sum up to 1. Here's a (hopefully) useful way to think about this. Professor Halfbrain tells me that he is 65 inches tall. If we are rounding to the nearest inch (as in the discrete case), then I believe him. But on a continuous scale, to know his exact height, we would need to know infinitely many decimal places (i.e. $65.149275089382738\cdots$ inches). There are just too many possible numbers to assign a probability to every single one. The proof is simple.

$$P(x - h < X \leq x) = F(x) - F(x - h)$$

For each side of this equation, we take the limit as $h \to 0$ (from the right) and we get

$$P(X = x) = F(x) - F(x) = 0.$$

Another claim:

$$P(X < x) = P(X \leq x) \qquad \text{and} \qquad P(X > x) = P(X \geq x)$$

This is a direct consequence of the fact that $P(X = x) = 0$.

$$\begin{aligned} P(X \leq x) &= P(X < x \ \cup \ X = x) \\ &= P(X < x) + P(X = x) && \text{(disjoint events)} \\ &= P(X < x) && \text{(since } P(X = x) = 0) \end{aligned}$$

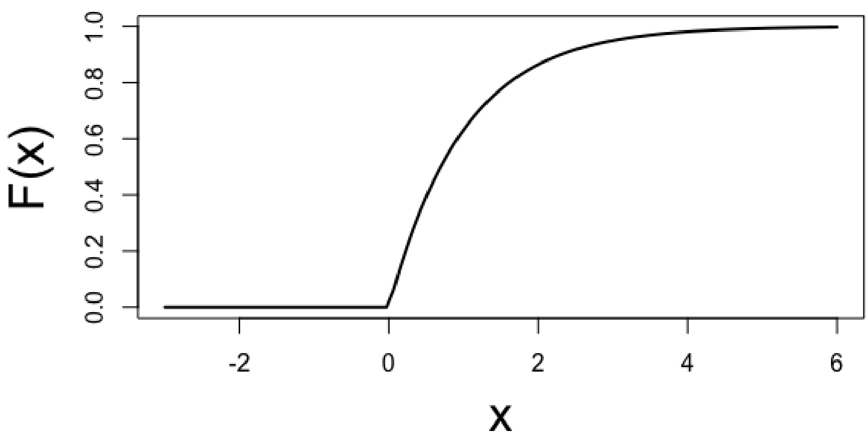

FIGURE 2. CDF for example 2.

Example 2. Timothy has just placed an order at McDonalds. Let X be the time (in minutes) it takes for him to get his food. Suppose that the CDF of X is

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x}, & x \geq 0 \end{cases}$$

i) What is the probability that Timothy waits at least 2 minutes to get his food?

$$P(X > 2) = 1 - F(2) = 1 - (1 - e^{-2}) = 0.135$$

ii) What is the probability that Timothy waits between 1.5 and 4 minutes for his food?

$$P(1.5 < X \leq 4) = F(4) - F(1.5) = (1 - e^{-4}) - (1 - e^{-1.5}) = 0.205$$

iii) What is the probability that Timothy waits less than 1 minute OR more than 4 minutes?

$$P(X < 1 \cup X > 4) = P(X < 1) + P(X > 4) = F(1) + 1 - F(4) = 1 - e^{-1} + e^{-4} = 0.65$$

iv) What is the probability that Timothy waits more than 2 minutes (total) given that he has already waited 1 minute?

$$P(X > 2|X > 1) = \frac{P(X > 2 \cap X > 1)}{P(X > 1)} = \frac{P(X > 2)}{P(X > 1)} = \frac{e^{-2}}{e^{-1}} = 0368$$

Definition. Let X be a continuous RV with CDF $F(x)$. The *probability density function* (PDF) of X is the function $f(x)$ such that

$$f(x) = \frac{dF(x)}{dx}$$

A PDF is *valid* if it satisfies the following properties.

- $f(x) \geq 0$ for all x
- $\int_{-\infty}^{\infty} f(x)dx = 1$

Note: If $F(x)$ is a valid CDF, then $f(x)$ is guaranteed to be a valid PDF.

Note the similarity between the PMF and PDF. The properties required for a PDF to be valid are almost the same as the PDF (for discrete RVs) with an integral swapped in for the sum.

Example 1 revisited. If X is a RV with the CDF given in example 1, then the probability density function (PDF) of X is

$$f(x) = \frac{dF(x)}{dx} = \begin{cases} \frac{x}{50}, & 0 < x \leq 10 \\ 0, & \text{otherwise} \end{cases}$$

The "integrates to 1" property of a PDF means that the area underneath the density curve is equal to 1. For this example, we can easily check that this is true with geometry (although for more complicated PDFs this may not be the case).

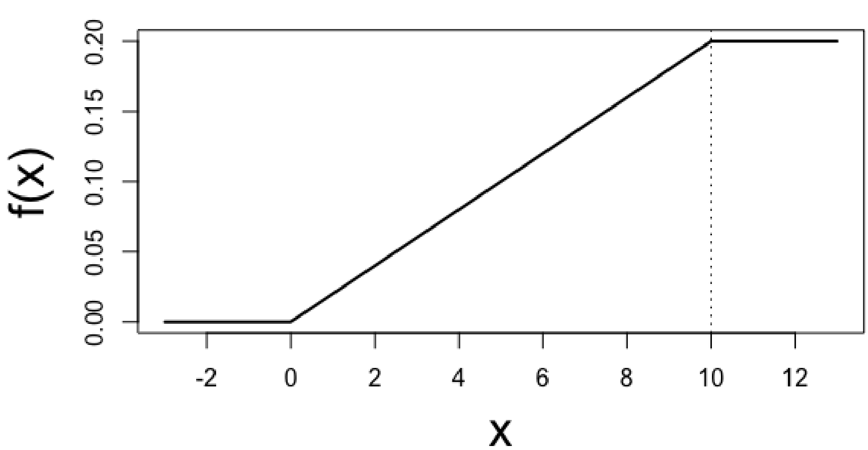

Figure 3. PDF for example 1.

Example 2 revisited. If X is a RV with the CDF given in example 2, then the PDF of X is

$$f(x) = e^{-x},\ x > 0.$$

To check that this density function "integrates" to 1, we can't use geometry, but the integral is fairly straightforward.

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{\infty} e^{-x}dx = 0 - e^{-x}|_0^{\infty} = 0 - (0 - 1) = 1$$

Interpretation of PDF. Probability density functions don't have the same interpretation as a probability mass function, but they are used similarly. There are many ways to think about a PDF, but I want to focus on an interpretation which connects them with *data*. Consider a collection of n data points, all from the same distribution. If we make a histogram of the data, the bars will tend to form a shape. As n increases (to infintiy), the outline of the histogram bars represents the probability density curve. Figure 5 illustrates this concept using the PDF in example 2.

Getting from PDF to CDF. In this lecture, we began with the CDF and presented the PDF as the derivative of the CDF. Naturally, we can reverse this process through integration. If X is a random variable

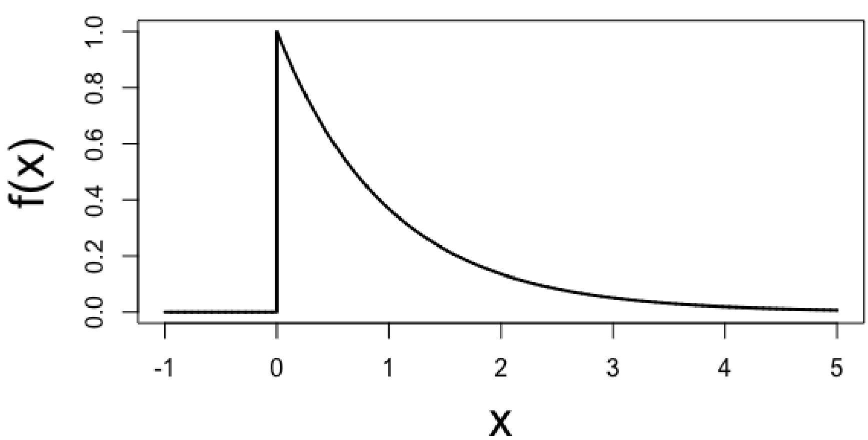

Figure 4. PDF for example 2.

with PDF $f(x)$, then the CDF of X is given by

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

PDFs and Continuous RVs. The probability density function can be used for a continuous RV to find probabilities. Let X be a continuous RV with PDF $f(x)$ and CDF $F(x)$.

$$P(a < X \leq b) = \int_{a}^{b} f(t)dt = F(b) - F(a)$$
$$P(X < a) = \int_{-\infty}^{a} f(t)dt = F(a)$$
$$P(X > b) = \int_{b}^{\infty} f(t)dt = 1 - F(b)$$

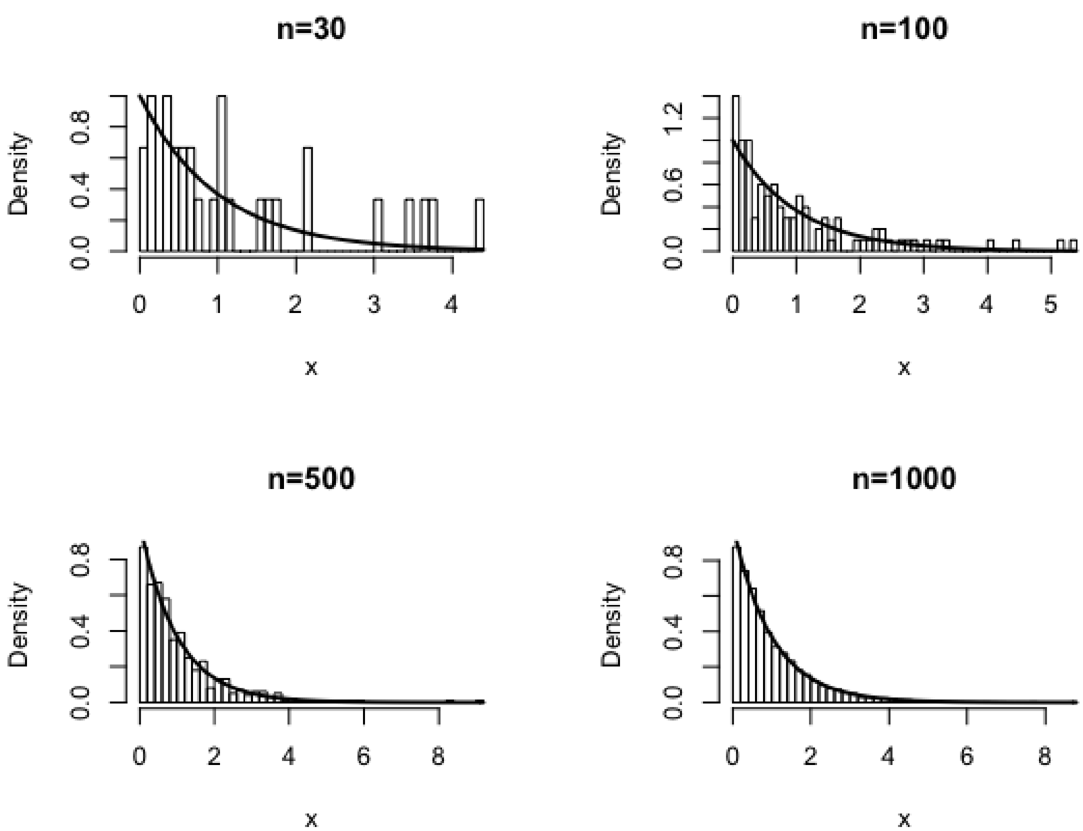

Figure 5. Density curve (example 2) and data.

Calculation of expected values and variance is similar to the discrete case, using the PDF (rather than PMF) and using an integral (rather than a sum).

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$
$$E(h(X)) = \int_{-\infty}^{\infty} h(x) f(x) dx$$
$$Var(X) = E\left((X - E(X))^2\right) = E(X^2) - E(X)^2 \qquad \text{(same as before)}$$

Example 1 revisited. For instance, let X be a RV with the CDF and PDF of example 1. Then the expected value is

$$E(X) = \int_0^{10} x \frac{x}{50} dx = \frac{1}{150} x^3 |_0^{10} = 0.66\cdots$$

To find the variance, we first compute $E(X^2)$ as

$$E(X^2) = \int_0^{10} x^2 \frac{x}{50} dx = \frac{1}{200} x^4 |_0^{10} = 50$$

so that the variance is

$$Var(X) = E(X^2) - E(X)^2 = 50 - (6.6667)^2 = 5.55\cdots$$

and the standard deviation is

$$SD(X) = \sqrt{Var(X)} = \sqrt{5.55} = 2.36.$$

For a more complicated example, we could try finding the expected value of $E(2^{3X})$, which requires solving the integral

$$E(2^{3X}) = \int_0^{10} 2^{3x} \frac{x}{50} dx.$$

To solve this integral, we need to do integration by parts. We will do many integrals of this form next week. Try it out on your own and see if you get the right answer (6361985).