Random Variables, pt. III

ECON 3640–001

Marcio Santetti
Spring 2022

Motivation

Housekeeping

Notes based on `Keller (2009)`, ch. 7

- pp. 222—225.

What else do we need?

By now, we know what a **probability distribution** is.

One of its uses is to give us information about the **probability** that a random variable will be *equal* to some value x; or will lie *between* values a and b, or will be *greater* than c, and so on.

Another use of probability distributions is to express **degrees of uncertainty/belief** in visual and probabilistic terms.

- Later!

But many times, our goal is to simply **summarize** key pieces of information from a distribution.

- Hello again, *summary statistics*!

Expected Value

Expected Value

We already know what a *population* or *sample* means are.

The way we use these concept is through an *arithmetic* mean of from a list of values.

Thus, given a list of values $\{x_1, x_2, x_3, \ldots, x_n\}$, the arithmetic mean is defined by

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

A more **general** definition of a mean is a **weighted average**:

$$\text{weighted-mean}(x) = \sum_{i=1}^{n} x_i p_i$$

where $p_i = p_1, p_2, \ldots, p_n$ are *predetermined* nonnegative numbers (weights) that add up to 1.

- What is p_i in the arithmetic mean formula?

Expected Value

When we are dealing with **random variables**, we use the concept of a weighted average to calculate its **expected value (EV)**:

> The **expected value** (*aka* the *expectation* or *mean*) of a random variable X whose possible values are x_1, x_2,..., x_n is defined by

$$E(X) = \sum_{all\ x} x\ P(X = x) \qquad \text{(Discrete RVs)}$$

$$E(X) = \int_{all\ x} x\ P(X = x)\ dx \qquad \text{(Continuous RVs)}$$

Intuitively, the expected value of X is a weighted average of the possible values that X can take on, **weighted by their probabilities**. $E(X)$ measures the **trend** or **long-run** average of X.

Expected Value

From the *definition* of expected value, we see that it depends only on the **distribution** of X.

Therefore, if we have two different random variables, X and Y, with the **same distribution**, then

$$E(X) = E(Y)$$

However, the **converse** of the above statement is *not true*.

- Why?

Variability

Variability

The expected value informs the **center** of mass of a distribution.

However, it does not tell us how **spread out** the distribution is.

> The *variance* of a random variable X is given by

$$Var(X) = E(X - E(X))^2 \quad \text{or}$$

$$Var(X) = E(X^2) - [E(X)]^2$$

And the **standard deviation** of a random variable is

$$SD(X) = \sqrt{Var(X)}$$

Properties of Expected Value and Variance

Properties of Expected Value and Variance

Many times, we work with random variables that are **functions of** other random variables.

We can **easily** calculate the expected value and variance measures when this is the case.

$E(c) = c$

$E(X + c) = E(X) + c$

$E(cX) = cE(X)$

$Var(c) = 0$

$Var(X + c) = Var(X)$

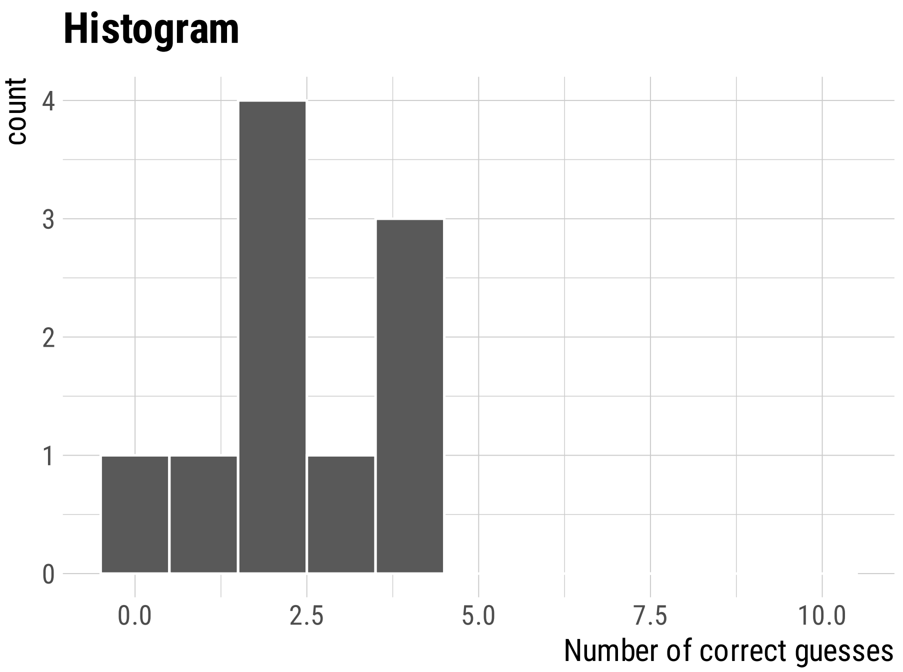
$Var(cX) = c^2\, Var(X)$

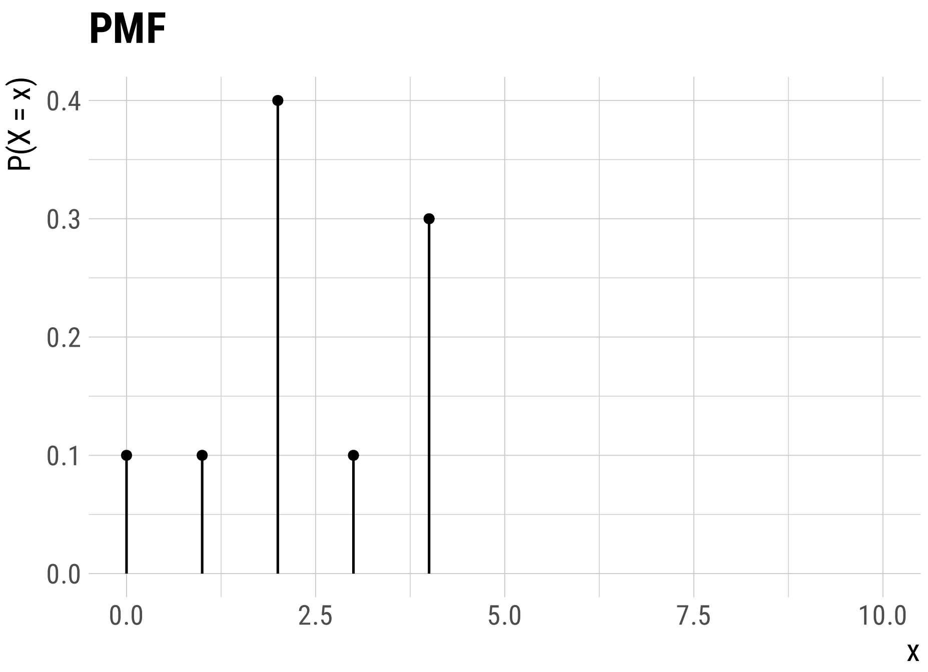
where c is a constant.

Pictures

Pictures

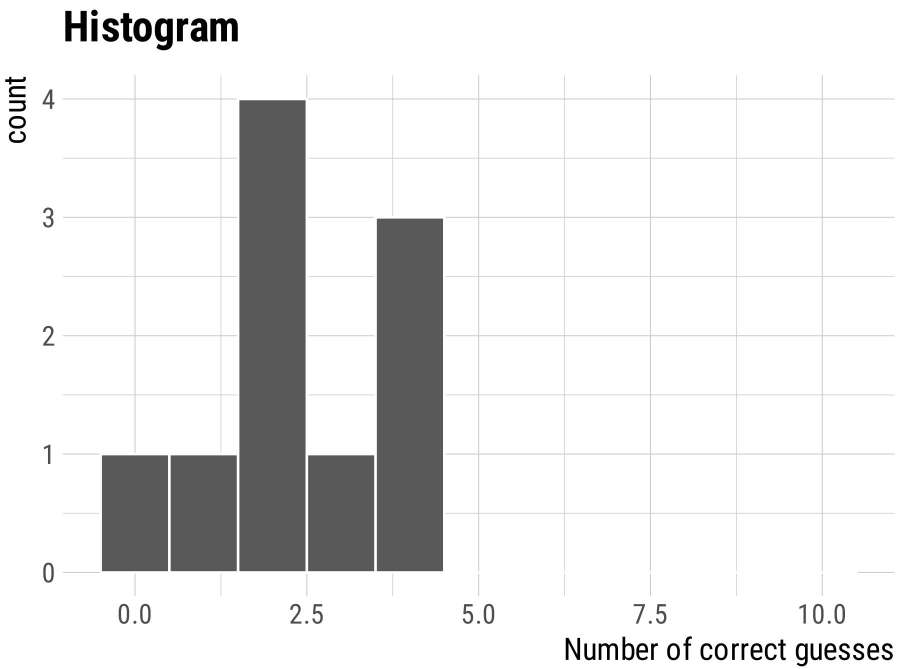
Recall the example seen in class of a **binomial random variable** (students guessing answers in a multiple-choice quiz).

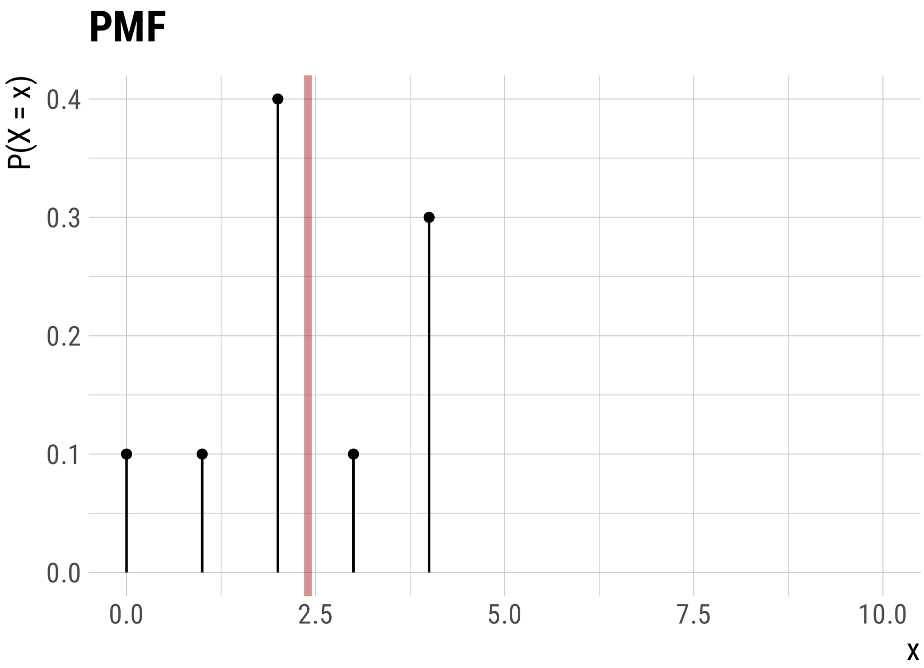




Pictures

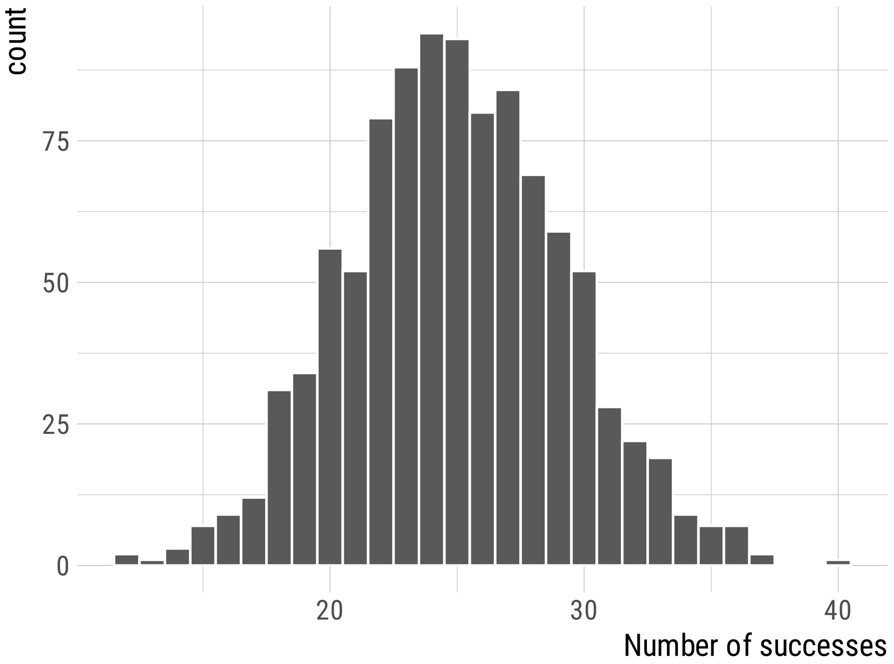
Now, highlighting the expected value (**red** vertical line):

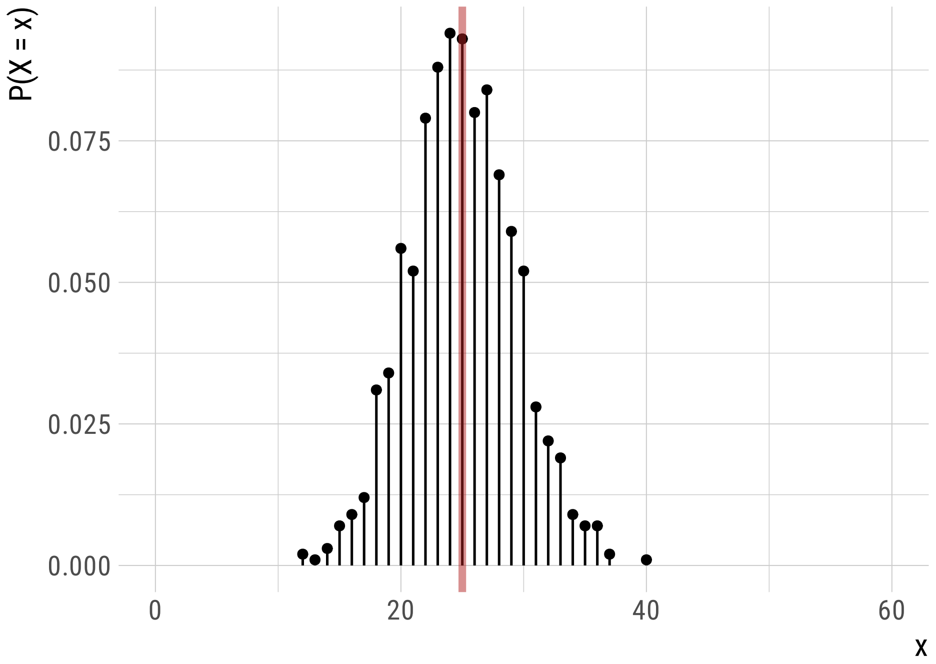




Pictures

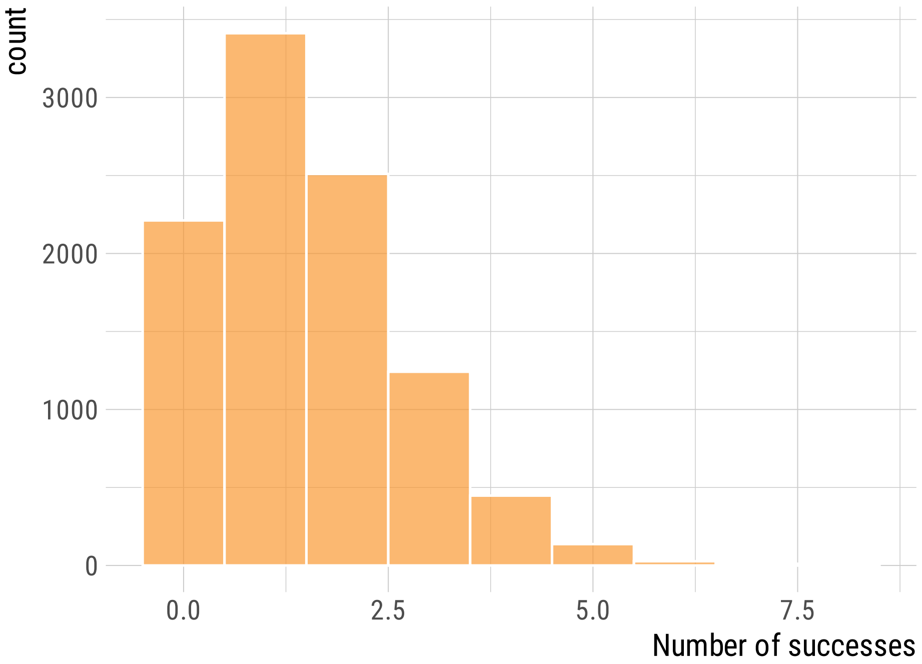
Now, suppose $X \sim \text{Binom}(100, 0.25)$:

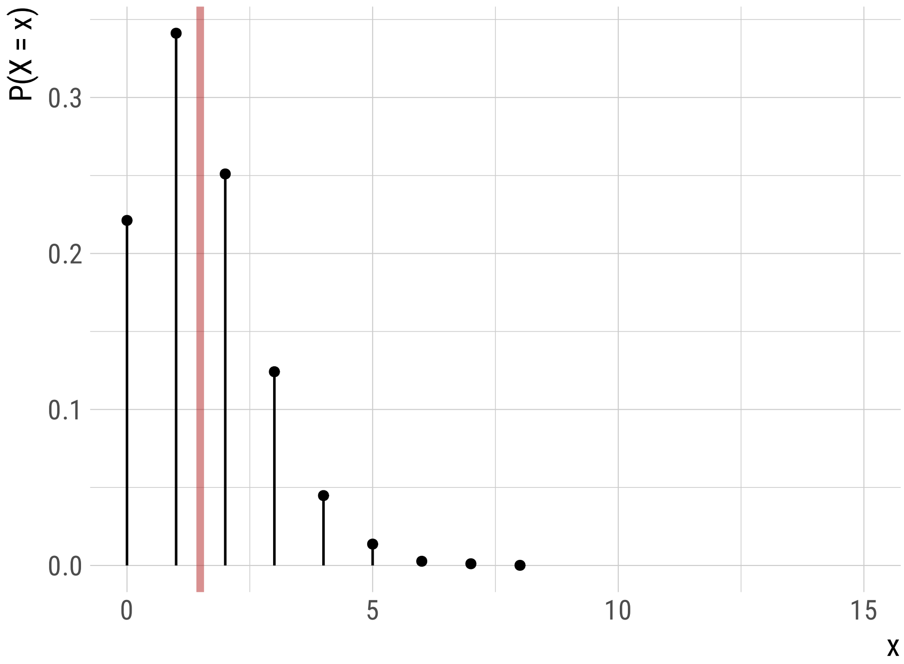




Pictures

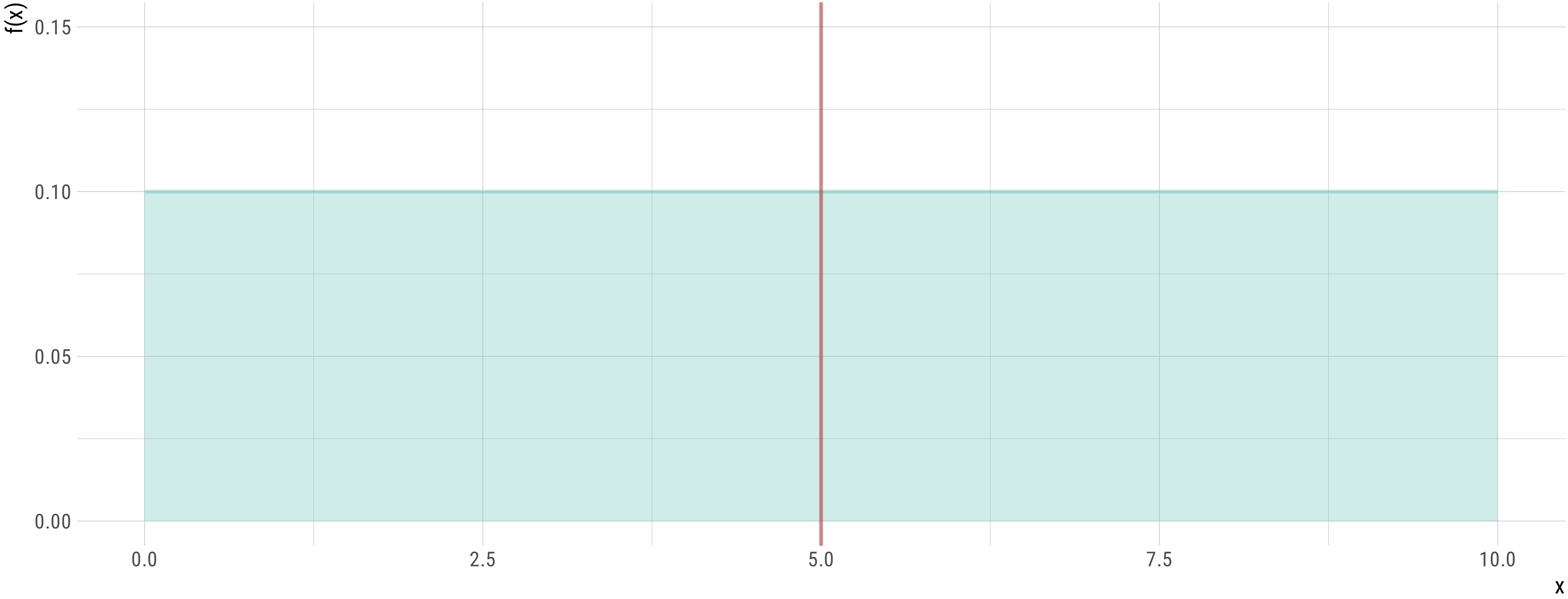
Assume a Poisson-distributed RV, $X \sim \mathrm{Pois}(1.5)$:





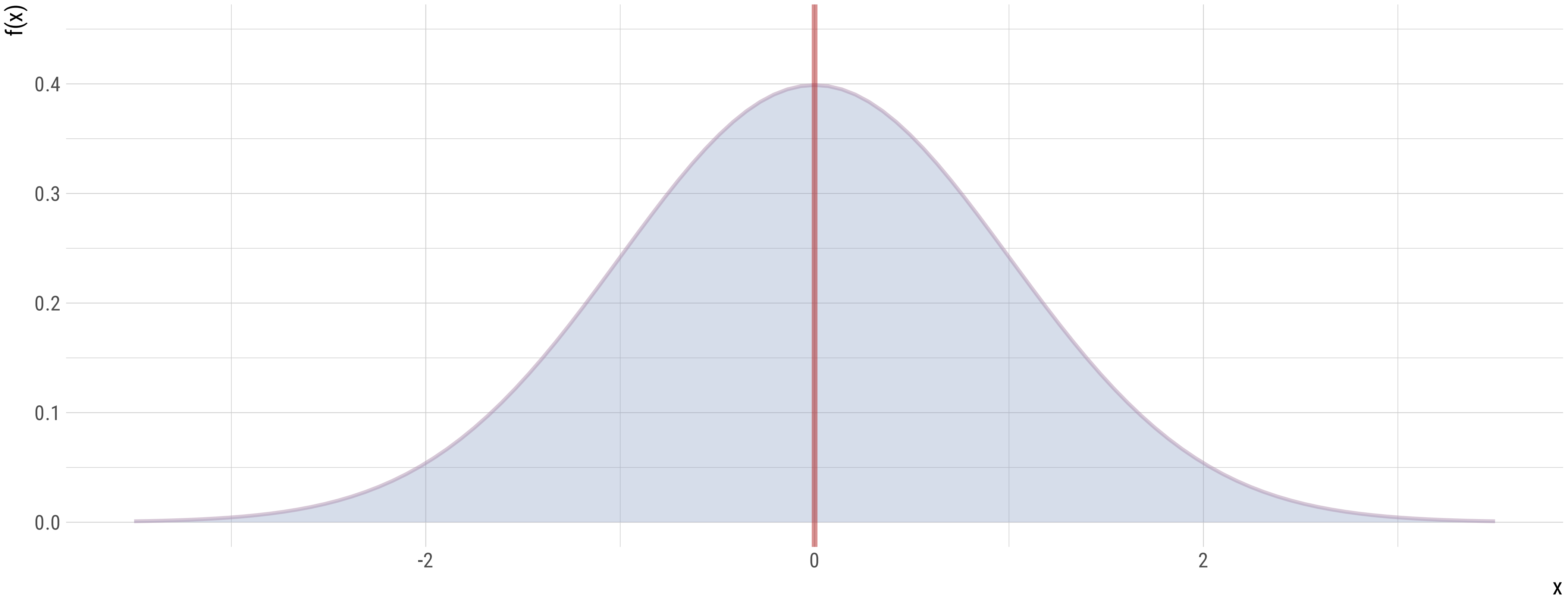
Pictures

Now, assume a continuous random variable, uniformly distributed between 0 and 10: $X \sim \mathbf{Unif}(0, 10)$.



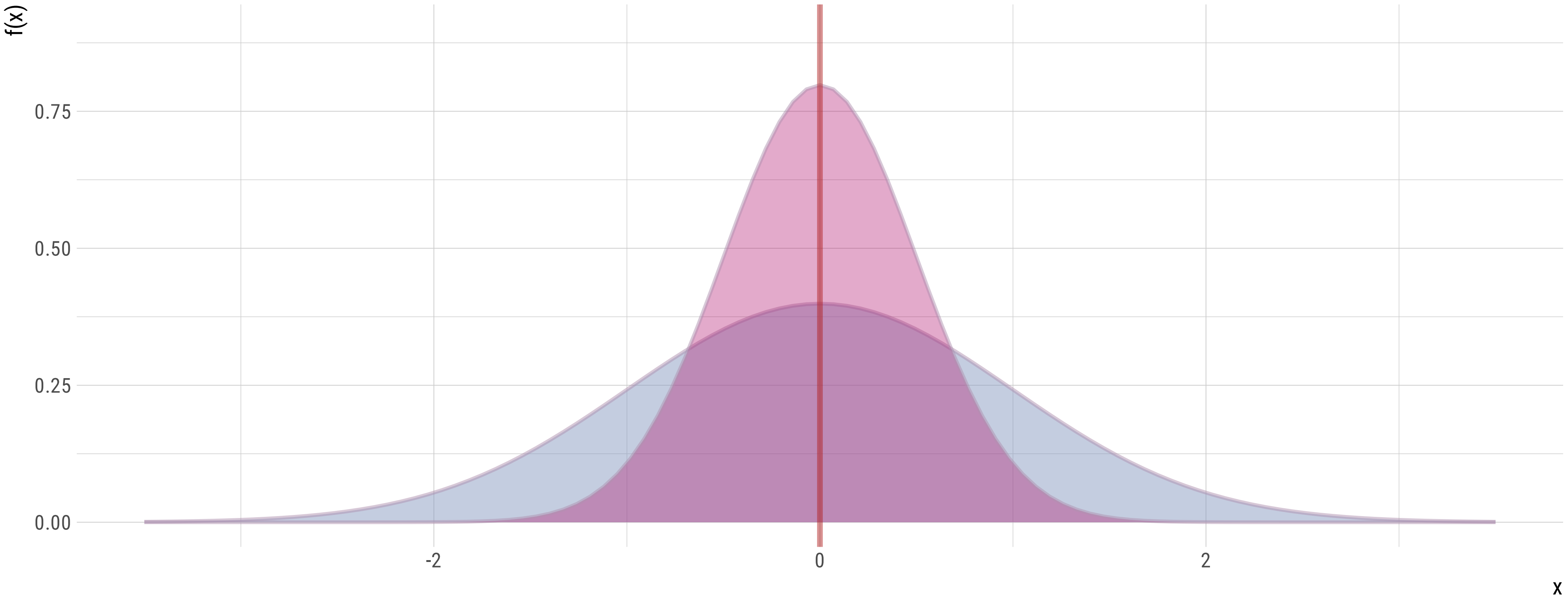
Pictures

Now, consider a **standard normally** distributed random variable: $X \sim \mathcal{N}(0, 1)$



Pictures

What is the **difference** between these two?



Next time: Back to R!