Terminology: estima**nd**, estima**tor**, estima**te**

Suppose we are interested in a particular feature of an unknown distribution. It could be a common summary statistic — such as the mean or the variance of the distribution.

Perhaps we are interested in the probability of a particular event.

We may be interested in the optimal action under some problem-specific utility function (such as our “milk demand” example).

In each case, any quantity of interest of an unknown probability distribution is referred to as an **estimand**.

Estima**nd**, estima**tor**, estima**te**

Next, we proceed to use data to figure out what the *estimand* is.

To do this, we come up with a recipe for taking observed data and spitting out our best guess as to what the *estimand* is.

This recipe itself is called an **estimator**.

Note that this recipe defines a random variable, induced by the data generating process: estimators are random variables!

Estima**nd**, estima**tor**, estima**te**

Finally, after we observe actual data, we can apply the estimator recipe to it, to obtain an actual number.

This number — our post-data guess as to the value of the unknown estimand — is called our **estimate**.

Estimates are fixed numbers based on the observed data.

Estima**nd**, estima**tor**, estima**te**

An **estimand** is a fixed but unknown property of a probability distribution. It is the thing we want to *estimate*.

An **estimator** is a recipe for taking observed data and formulating a guess about the unknown value of the *estimand*.

An **estimate** is a specific value obtained when the *estimator* is applied to specific observed data.

Example: population mean: $\mathrm{E}(X)$; sample mean: $\bar{X}$; sample mean: $\bar{x}$

Suppose we are interested in $\mathrm{E}(X)$, the mean of a random variable X. This is sometimes referred to as the **population mean**, in reference to polling problems.

The standard estimator of the population mean is the mean of the observed data, or the **sample mean**. Before any data is observed, this defines our estimator, commonly denoted $\bar{X} \equiv \frac{1}{n}\sum_{i=1}^{n} X_i$.

Once the data has been observed, we have an estimate in hand, denoted $\bar{x} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i$. (Note the lower-case x's here.) This is the **observed sample mean**.