Need for Formal Statistical Inference

In a randomized, double-blind experiment:

17 of 20 subjects in the control group died.

11 of 20 subjects in the treatment group died.

How confident should we be that the treatment is beneficial?

In a simple random sample of eligible voters:

598 voters in the sample of 1000 say they will vote for the Liberals.

How confident should we be that the Liberals will receive a majority of votes in the election?

Using an apparatus to measure a property of protons:

We take 20 measurements, and find that their mean is 0.512, and their standard deviation is 0.014.

We have a theory that says the true value is exactly 1/2. Should we abandon this theory?

What Formal Statistical Inference Can and Cannot Do

The mathematical theory of statistics *cannot*:

- Tell you which population you should be interested in.
- Ensure that you sampled properly from the population.
- Determine whether measurements made by your apparatus are systematically wrong.

Mathematical statistics *can*:

- Give you a *quantitative* indication of how much random variation may have affected your results.

But this indication alone *cannot*:

- Tell you what decision to make. That should depend also on other information you have, and on possible consequences.

Sources of Randomness

Where does the "random variation" that formal statistics can tell us about come from?

Sometimes, we deliberately introduce randomness:

- We randomly assign subjects to control and treatment groups.
- We randomly select a sample from a population.

Other times, we use randomness in a model of reality:

- We may model the errors our measuring apparatus makes as being random, with a normal distribution.
- We may model the relationship of crop yield to amount of fertilizer applied as linear, with random residuals.

Parameters, Statistics, and Estimators

In many statistical problems, we want to in infer some characteristic of a *population*, based on a *sample* from that population.

Terminology

Parameter: A number describing the population.

Statistic: A number we can compute from the sample.

Estimator: A statistic that we hope will be close to a parameter we are interested in.

Example:

The fraction of Canadians who own cars is a parameter.

The fraction who own cars in a sample of Canadians is a statistic.

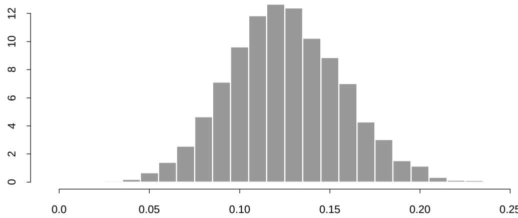
We use the fraction in the sample to estimate the fraction in the population.

Sampling Distribution of a Statistic

A statistic is not a definite number, but rather a rule for computing a number from a sample.

The *value* of the statistic will vary from one sample to another. Many samples of a given size will reveal the statistic's *distribution*:

Population of size 30 million, Simple random sample of size 100
Parameter, π, is the fraction of population supporting the Liberals
Statistic, p, is the fraction supporting the Liberals in the sample



Histogram of p obtained from 10,000 samples, when $\pi = 0.125$

If a statistic is used as an estimator, we call its value for a particular sample the *estimate* of the parameter derived from that sample.

The Bias of an Estimator

If the mean of an estimator's distribution is equal to the population parameter, we say the estimator is *unbiased*.

The proportion, p, from a simple random sample is an unbiased estimator of the proportion in the population. (We will see this mathematically later.)

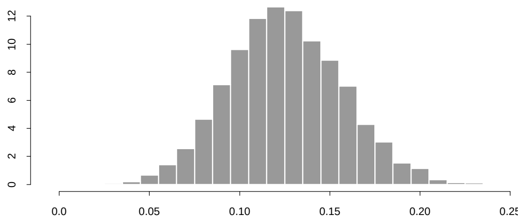
An unbiased estimator is "fair", in a sense.

But will the value of an unbiased estimator necessarily be close to the true parameter value?

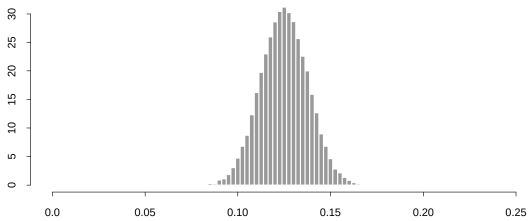
The Variability of an Estimator

To tell how accurate an estimator is likely to be, we need to know how *variable* it is — the spread of its distribution.

Getting more data reduces variability:



Distribution of p: population of 10000, sample of 100, $\pi = 0.125$



Distribution of p: population of 10000, sample of 400, $\pi = 0.125$

What Makes a Good Estimator?

An estimator is likely to be close to the true parameter value if it has:

- **Low bias**: Its distribution is centred near the parameter.
- **Low variabililty**: It does not vary much from this central value.