

Estimating parameters of a population

This is one of two big ideas in inferential statistics: estimation and hypothesis testing

When trying to make an inference about a population, there are usually 3 quantities at play:

1. a measurement about the population taken from a sample
2. the true value of that quantity for the population
3. an estimate for #2. Sometimes this is the same as #1, other times it may be the result of running #1 through some function (as is the case for standard deviation)

For the case of estimating a population mean:

- \bar{X} = **Sample mean**
 - This is the average of the samples to collect
- μ = **Population mean**
 - This is the *True* mean of the population, or other true quantity about the population. This is almost always impossible to know
- $\hat{\mu}$ = **Estimate of the population mean**
 - This represents our best guess of the population mean based on the data we've collected.

These 3 quantities may not even be numbers. For example, in the case of linear regression, these 3 quantities could be models. There is some true model that best represents the relationship between 2 quantities in a population, we collect a dataset and use that to build a model whose coefficients are estimates of the *true* model's coefficients.

Sources

- Learning statistics with R section 10.4