Sampling Distributions and Confidence Intervals

M. Drew LaMar
September 02, 2016

> "…a hypothesis test tells us whether the observed data are consistent with the null hypothesis, and a confidence interval tells us which hypotheses are consistent with the data."
>
> - William C. Blackwelder

Introduction to Quantitative Biology, Fall 2016

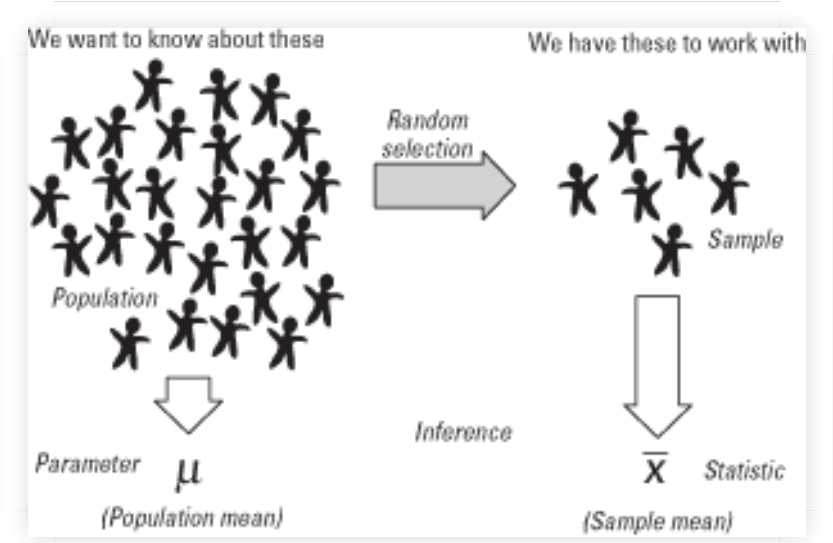
Class announcements

- Update on CoursePack: Still in processing…
- Blogs will start NEXT week - stay tuned for blogging prompt
  - Homework Assignment #1 UPDATE (due Monday, September 5):
  - OpenStats, Chapter 4: 4.6.1 Variability in estimates (p. 203) - #4.3, 4.5
  - OpenStats, Chapter 4: 4.6.2 Confidence intervals (p. 206) - #4.8, 4.10, 4.14, 4.16
  - Depending on class today, exercises on Hypotheses Testing
- Reading assignment for Monday (posted on BB): **OpenIntro Stats, Chapter 7: Introduction to Linear Regression**

Populations vs Samples

Definition: A ***parameter*** is a quantity describing a population, whereas an ***estimate*** or ***statistic*** is a related quantity calculated from a sample.

Parameter examples: Averages, proportions, measures of variation, and measures of relationship



Sampling Distributions

Definition: The ***sampling distribution*** represents the distribution of the point estimates ***based on samples of a fixed size*** from a certain population. It is useful to think of a particular point estimate as being drawn from such a distribution. Understanding the concept of a sampling distribution is central to understanding statistical inference.

Definition: The standard deviation associated with an estimate is called the ***standard error***. It describes the typical error or uncertainty associated with the estimate.

The standard error is also the standard deviation of the sampling distribution.

http://www.zoology.ubc.ca/~whitlock/kingfisher/SamplingNormal.htm

Confidence Intervals

Definition: The standard error represents the standard deviation associated with the estimate, and roughly 95% of the time the estimate will be within 2 standard errors of the parameter.

An *approximate* 95% confidence interval for a point estimate is given by

$$\text{point estimate} \pm 1.96 \times SE$$

Note: For a yuge number of computed 95% confidence intervals, the population parameter will be contained in 95% of the confidence intervals.

http://www.zoology.ubc.ca/~whitlock/kingfisher/CIMean.htm