

Psych. 310I

Lecture 3

Samples, Populations, etc.

Reading: Chapter 1: Introduction to Statistics

Populations and Samples

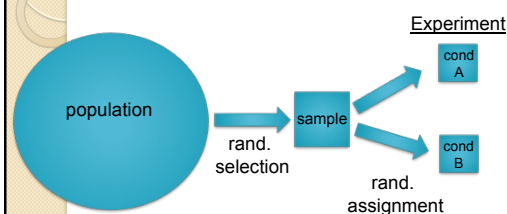
- Population
 - Set of subjects, items, or events we want to learn about
 - Generally very large or infinite
 - All people, all men/women, all pigeons, all people with depression in industrialized countries
- Sample
 - Subset of population assessed in a given study
 - Much smaller
 - Randomly selected (if we want to make a statement about the population)
 - Not perfectly representative of population (sampling variability)

Random Selection

- Every member of population has equal chance of inclusion
- Property of data-gathering process
 - Study design must take random selection into account
- Otherwise sample is biased
 - Only testing students at library
 - Allowing participants to self-select into study

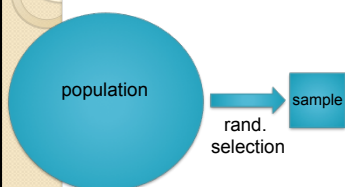
Different than random assignment!

Random Selection vs Random Assignment



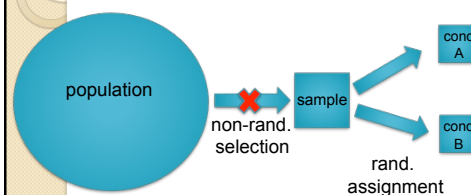
Randomness ensures that sample or subsample won't be too different in composition from what it was drawn from. This certainty increases with n .

Random Selection only



Allows us to generalize to the population, but not make causal inferences

Random Assignment only



Allows us to make causal inference, but not generalize back to the broader population

Parameter

- Characteristic of the population
- Usually theoretically meaningful
- Mean, variance, proportion, rate, correlation
 - What's the average height of college students?
 - How many attempts does it take a normal rat to learn this maze?
 - How many attempts if we cut out its hippocampus?
 - What fraction of words can a subject remember?
 - What's the correlation between height and extraversion?

Statistic

- Mathematical function to be computed from data
- Difference between statistic and its value
 - E.g., *mean* is a statistic (arithmetic average)
 - Value for any dataset will be some number
- Usually serves one of three functions
 - Descriptive statistic: Summarizes some aspect of the sample data
 - Inferential statistic: Aids testing of some hypothesis about the population

Descriptive Statistic

- Summarizes some aspect of the data
 - Mean, median, maximum, quartiles, standard deviation, etc.
- Used only for describing sample data
 - Not for making inferences about population
 - Can be first step of data analysis
 - Also useful if sample is all you're interested in
 - E.g. average age of students in class

Inferential Statistic

- Aids testing of some hypothesis about the population
- Indicates how reliable an effect in the sample is
- Value generally has no physical meaning
 - Not like inches, time, or even psychological variables
- Examples: t , F , χ^2 , β

Basic Logic of Inferential Statistics

- Start with two hypotheses:
 - Null hypothesis – there is no effect of IV
 - Alternative hypothesis – there is some effect of IV
- We pretend that the null is true
- We then ask, “how likely is our observed statistic if the null is true?”
- If our observation is unlikely (given the null), then we reject the null and say there is an effect

Example 1 - tipping

- Patrons in a buffet restaurant tip more when the server is assigned to introduce herself by name to a table compared to when she is assigned to not introduce herself to a table - Garrity & Degelman (1990)
 - null hypothesis?
 - alternative?

Example 2 – social support & depression

- People who have a smaller support network tend to be at higher risk for depression – Cohen & Wills (1985)
 - null hypothesis?
 - alternative hypothesis?