

- Sampling terminology
  - Parameter
    - Fixed, unknown number that describes the population
  - Statistic
    - Known value calculated from a sample
    - A statistic is used to estimate a parameter
  - Bias
    - In repeated samples, the sample statistic consistently misses the population parameter in the same direction
- Sampling Strategy
  - To reduce bias, use random sampling
  - Increasing sample size does not help with Bias
  - To reduce variability, use larger samples
- Proportions (p-hat)
  - P = Proportion
  - ^ = Comes from sample
- Margin of error
  - The amount by which the answer will differ from p-hat
- Key Concepts
  - Parameters vs statistics
  - Bias variability
  - Margin of error
  - Confidence statements
- Errors in Sample Surveys
  - Sampling Errors
  - Random Sampling Errors
    - Measured by margin of error
  - Non-sampling errors
- Concerns when Asking Survey Questions
  - Deliberate Bias
  - Unintentional Bias
  - Desire to Please
  - Asking the Uninformed
  - Ordering of questions
  - Unnecessary Complexity
  - Confidentiality and Anonymity
- Probability Sampling Plans
  - Simple Random Sampling
  - Stratified random sampling
  - Cluster sampling
  - Systematic sampling
  - Random digit dialing
  - Multistage sampling

- Experiments: Difficulties and Disasters
  - Extraneous variables
    - Confounding variables
    - Interacting variables
  - Hawthorne, placebo and experimenter effects
  - Refusals, no adherers, dropouts
- Interacting Variables
  - The problem:
    - Effect of explanatory variable on response variable may vary over levels of other variables
  - The solution:
    - Measure and study potential interacting variables.
- Hawthorne, placebo and experimenter effects
  - The problem:
    - People may respond differently when they know they are part of an experiment
  - The solution:
    - Use placebos, control groups, and double-blind studies when possible
- Extending the results
  - Taking what you found in your sample and applying it to the whole population
  - The problem:
    - Unrealistic treatment
    - Unnatural setting
    - Sample that is not representative of the population
- Key concepts
  - Double blind
  - Difficulties and disasters
  - Experimental designs
    - Completely randomized design
    - Matched pairs design
    - Blocked design
- Data ethics
  - Experiments: ethical requirements
    - Institutional review board
      - To protect subjects from possible harm
    - Informed consent
    - Confidentiality
  - Confidentiality and Anonymity
    - Confidential Answer
      - Respondent is known, but the information is a secret
    - Anonymous answer

- The respondent is not known, or cannot be linked to his/her response
  - o Clinical trials
    - Experiments that study the effectiveness of medical treatments on actual patients
- Measuring
  - o Measurement Validity
    - A measurement is valid if it is an appropriate representation of the property of interest
    - A measurement has predictive validity if it can be used to adequately predict some outcome related to the property of interest
  - o Measurement Reliability
    - Errors in measurement
      - Measurement value = true value + bias + random error
      - Bias means that each measurement has the same systematic error in one direction
      - Random Error means that repeated measurements on the same individual give different results
    - A measurement is reliable if the random error is small (i.e., we get “repeated results”)
      - Reliability can be improved by taking multiple measurements and taking the average of those results