Statistics One

Lecture 5

Measurement

Two segments

- Reliability & Validity
- Sampling

Lecture 5
Segment 1
Reliability & Validity

Reliability & Validity

- Important concepts/topics
  - Classical test theory
  - Reliability estimates
  - Construct validity

Reliability

- Classical test theory
  - Raw scores (X) are not perfect
  - They are influenced by bias and chance error
  - In a perfect world, we would obtain a "true score"
    - X = true score + bias + error
    - Also known as "true score theory"

Reliability

- A measure (X) is considered to be reliable as it approaches the true score
  - The problem is we don't know the true score
  - So, we estimate reliability

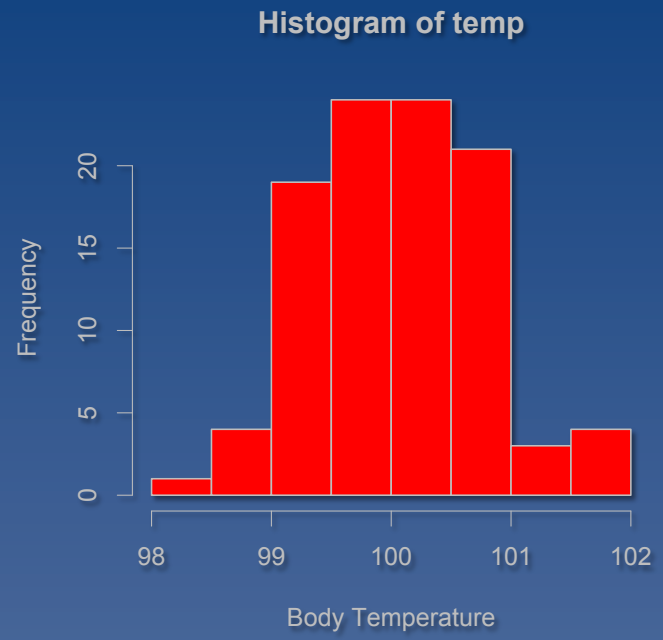
Reliability

- Methods to estimate reliability
  - Test / re-test
  - Parallel tests
  - Inter-item estimates

Reliability

- Example: Body temperature
  - Infrared meter: “The wand”

Body temperature



Reliability

- Test / re-test
  - Measure everyone twice
    - X1
    - X2

Reliability

- Test / re-test
  - The correlation between X1 and X2 is an estimate of reliability
    - However, if the bias is uniform then we won't detect it with the test / re-test method

Reliability

- Parallel tests
  - Measure body temperature with the wand (X1) and with an oral thermometer (X2)
  - The correlation between X1 and X2 is an estimate of reliability
    - AND, now the bias of the wand will be revealed

Reliability

- Inter-item
  - Inter-item is the most commonly used method in the social sciences
    - Test / re-test and parallel tests are time consuming
    - Inter-item is therefore more cost efficient

Reliability

- Inter-item
  - For example, suppose a 20-item survey is designed to measure extraversion
    - Randomly select 10 items to get sub-set A (X1)
    - The other 10 items become sub-set B (X2)
    - The correlation between X1 and X2 is an estimate of reliability

Validity

- What is a construct?
  - How do we operationalize a construct?
  - Construct validity
    - Content validity
    - Convergent validity
    - Divergent validity
    - Nomological validity

Validity

- What is a construct?
  - An ideal "object" that is not directly observable
    - As opposed to "real" observable objects
    - For example, "intelligence" is a construct

Validity

- How do we operationalize a construct?
  - The process of defining a construct to make it observable and quantifiable
    - For example, intelligence tests

Validity

- Construct validity
  - Content validity
  - Convergent validity
  - Divergent validity
  - Nomological validity

Validity

- An example:
  - Construct: Verbal ability in children

Validity

- How to operationalize?
  - A vocabulary test

Validity

- Construct validity
  - Content validity
    - Does the test consist of words that children should know?

Validity

- Construct validity
  - Convergent validity
    - Does the test correlate with other, established measures of verbal ability?
      - For example, reading comprehension

Validity

- Construct validity
  - Divergent validity
    - Does the test correlate less well with measures designed to test a different type of ability?
      - For example, spatial reasoning

Validity

- Construct validity
  - Nomological validity
    - Are scores on the test consistent with more general theories, for example, of child development and neuroscience
      - For example, a child with neural damage or disease to brain regions associated with language development should score lower on the test

Reliability & Validity: Review

- Important concepts/topics
  - Classical test theory
  - Reliability estimates
  - Construct validity