Measurement & Sampling

CRP 501:
Quantitative Methods for Planning Data Analysis

Measurement

- **Measurement is the *assignment* of *numerals or numbers* to objects, events, or variables according to *rules*.**
  - *Numerals or Numbers (represents the attributes)*
  - *Assignment (attach numerals or numbers to subjects)*
  - *Rules (procedure of assignment)*
- **Measurement should be Isomorphic**
  - *Structure of the numeric system*
  - *Structure of the phenomenon*
- **Indicator vs. Direct/Indirect measurement**
- **Estimate, Projection, and Forecast**

Levels of Measurement

- **Nominal - Cases can be placed in categories, but not ranked**
  - *Numerals*
  - *Classify, Mutually-exclusive; Examples?*
- **Ordinal - Cases can placed in categories and ranked from "high" to "low"**
  - *Numerals*
  - *Compare and Sort; Examples?*
- **Interval**
  - *Numbers*
  - *Standard Unit; examples?*
- **Ratio**
  - *Numbers*
  - *TRUE Zero; examples?*
- **Levels of measurement and statistical analysis**
  - *Higher level means more information*
  - *Higher level of measurement requires more resources (e.g., time)*

Measurement Error

- **Definition –**
  - *occurs when a respondent's answer to a given question is inaccurate, imprecise, or cannot be compared in any useful way to other respondents' answer*
- **Y = a + bX + e**
- **Sources**
  - *Influence of associated attributes*
  - *Subject's conditions*
  - *Administration of instrument*
  - *Data Processing*
  - *Operator*
- **Types of Errors**
  - *Systematic errors*
  - *Random Errors*
- **How to reduce measurement error?**

Validity and Reliability

- **Validity**
  - *Validity – correspondence between the measure and the concept it is thought to measure; "Are you measuring what you think you're measuring?"*
  - *Content Validity (face, sampling)*
  - *Empirical Validity (true vs. observed values)*
  - *Construct Validity (relates to theory)*
- **Reliability (pp. 154-157)**
  - *Reliability –*
    - Accuracy of the measuring instrument
    - does a measure yield the same values for a particular case again and again?
    - Measured as the ratio of true-score variance to the total variance
  - *Reliability estimates:*
    - test-retest method
    - parallel-forms technique
    - split-half method

Introduction to Sampling

- **Why Sampling?**
  - *Make inferences about a population*
  - *Why not study the entire population?*
- **Important Concepts**
  - *Population vs. Sample*
  - *Population Parameters (denoted by Greek letters) vs. Sample Statistics (denoted by lower case letters)*
  - *Sampling Unit and Sampling Population*
  - *Sampling Frame (watch out for incomplete frames, clusters of elements, and blank foreign elements).*

Sample Designs

- **Nonprobability Sample Designs**
  - *Convenience Samples*
  - *Purposive Samples*
  - *Quota Samples*
- **Probability Sample Designs**
  - *Simple Random Samples (random digits, p.500)*
  - *Systematic Samples (list)*
  - *Stratified Samples (subgroups)*
  - *Cluster Samples (geographic area)*

Sample Size Determination

- **Standard Error**
  - The difference between a sample statistic and its corresponding population parameter.
  - *Standard Deviation of a Sampling Distribution*
- **Confidence Intervals**
  - *Determined by confidence level (usually 95%)*
  - *1 - confidence level = significance level (e.g., 0.05)*
  - *Characteristics of Normal Distribution*
  - *z-score, p.504*
  - *Normal Distribution and Non-Normal Distributions*
  - *Central Limit Theorem to the rescue*
- **Determining the sample size**

CENTRAL LIMIT THEOREM

- For a population with a mean μ and a variance ^{2}, the sampling distribution of the means of all possible samples of size *n* generated from the population will be approximately normally distributed - with the mean of the sampling distribution equal to μ and the variance equal to σ^2/n - assuming that the sample size is sufficiently large.

95% AND THE 99% CONFIDENCE INTERVALS (CI) FOR m

- **The 95% and the 99% confidence intervals for m are constructed as follows when n ³ 30.**
- **95% CI for the population mean m is given by**

$$\overline{X} \pm 1.96 \frac{s}{\sqrt{n}}$$

- **99% CI for m is given by**

$$\overline{X} \pm 2.58 \frac{s}{\sqrt{n}}$$

CONSTRUCTING A GENERAL CONFIDENCE INTERVAL (CI) FOR m

- **In general, a confidence interval for the mean is computed by:**

$$\overline{X} \pm Z \frac{s}{\sqrt{n}}$$

- **The Z value is obtained from the standard normal table in Appendix E.**