



Module 1: Introduction to Statistics

The Applied Research Center

Module 1 Overview

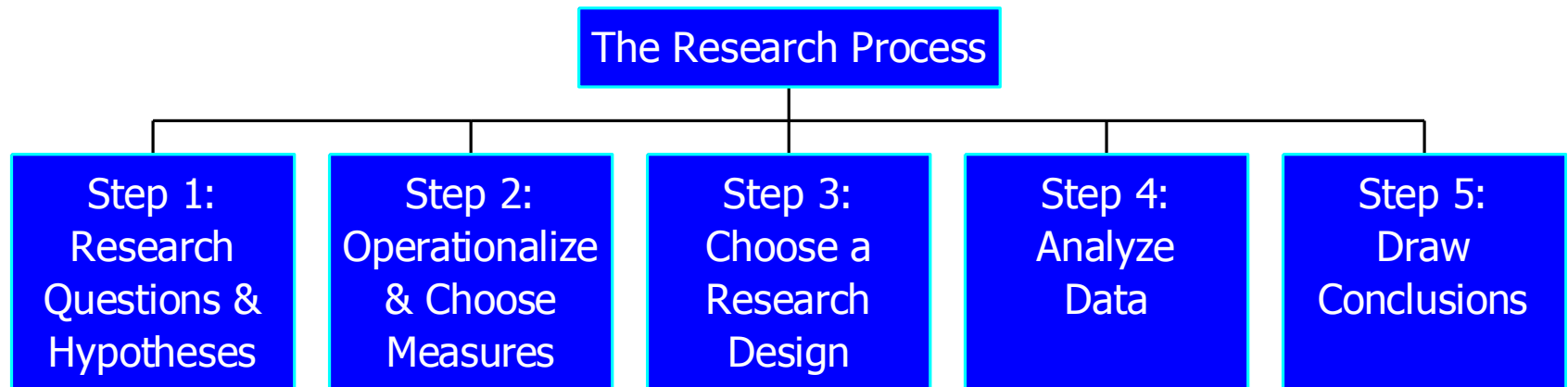
- ▶ The Role of Statistics
- ▶ The Research Process
 - ▶ Threats to Validity
- ▶ Statistical Terminology
- ▶ Scales of Measurement
- ▶ Introduction to Descriptive and Inferential Statistics

The Role of Statistics

- ▶ The goal of virtually all quantitative research studies is to identify and describe relationships among constructs.
- ▶ Data are collected in a very systematic manner and conclusions are drawn based on the data.
- ▶ At a basic level, statistical techniques allow us to aggregate and summarize data in order for researchers to draw conclusions from their study.

The Typical Research Process

- ▶ The typical quantitative study involves a series of steps, one of which is statistical analysis.
- ▶ Note: These are steps in the research process and NOT sections of the dissertation.



Step 1: Research Questions

- ▶ Research questions reflect the problem that the researcher wants to investigate.
- ▶ Research questions can be formulated based on theories, past research, previous experience, or the practical need to make data-driven decisions in a work environment.
- ▶ Research questions are **vitaly** important because they, in large part, dictate what type of statistical analysis is needed as well as what type of research design may be employed.

Examples of Research Questions

- ▶ How is financial need related to retention after the freshman year of college?
- ▶ What types of advertising campaigns produce the highest rates of inquiries among prospective applicants at NSU?
- ▶ How do males and females differ with respect to statistics self-efficacy?
- ▶ How does a body image curriculum improve body image in college females?

Hypotheses

- ▶ While research questions are fairly general, hypotheses are specific predictions about the results, made **prior** to data collection.
 - ▶ As financial need increases, the likelihood of retention decreases.
 - ▶ Personalized letters result in more inquiries than brochures.
 - ▶ Males have higher levels of self-efficacy than females.
 - ▶ Body image will improve as a result of the new curriculum.

Step 2: Operationalize & Choose Measures

- ▶ Many variables of interest in education and psychology are abstract concepts that cannot be directly measured.
- ▶ This doesn't preclude us from studying these things, but requires that we clearly define the specific behaviors that are related to the concept of interest.

Measuring Abstract Concepts

- ▶ How does one measure retention, inquiry rate, statistics self-efficacy, and body image?
- ▶ The process of defining variables and choosing a reliable and accurate measurement tool is called operationalizing your variables.
- ▶ Good measurement is **vital** to the trustworthiness of your results!

Step 3: Choose a Research Design

- ▶ In Step 3, we develop a plan for collecting the data we need (i.e., a “blueprint” for the study)
- ▶ This is called **research design**, and includes things such as:
 - ▶ Who will participate in the study?
 - ▶ Who will receive the intervention?
 - ▶ Will there be a “control group”?
 - ▶ Will data be collected longitudinally?
 - ▶ What instrument will be used to collect data?
 - ▶ What type of data will be collected?

Threats to Validity

- ▶ **Internal Validity**

- ▶ Problems associated with the experimental procedures or experiences of participants

- ▶ **External Validity**

- ▶ Problems that affect the generalizability of the results

- ▶ **The choice of design impacts the validity of your final results**

Step 4: Analyze The Data

- ▶ Once the data have been collected, the results must be organized and summarized so that we can answer the research questions.
- ▶ This is the purpose of statistics
- ▶ The choice of analysis at this stage depends **entirely** on two prior steps:
 - ▶ The research questions
 - ▶ How the variable is measured

Step 5: Draw Conclusions

- ▶ After analyzing the data, we can make judgments about our initial research questions and hypotheses.
- ▶ Are these results consistent with previous studies?
- ▶ The conclusions drawn from a study may provide a starting point for new research.

The Role of Statistics

- ▶ Despite the anxiety usually associated with statistics, data analysis is a relatively small piece of the larger research process.
- ▶ There is a misconception that the trustworthiness of statistics is independent of the research process itself.
- ▶ This is **absolutely** incorrect!
- ▶ A statistical analysis can in **no way** compensate for a poorly designed study!!!!



Statistical Terminology

Population

- ▶ A population is the **entire** set of individuals that we are interested in studying.
- ▶ This is the group that we want to generalize our results to.
- ▶ Although populations can vary in size, they are usually quite large.
- ▶ Thus, it is usually not feasible to collect data from the entire population.

Sample

- ▶ A sample is a **subset** of individuals selected from the population.
- ▶ In the best case, the sample will be representative of the population.
- ▶ That is, the characteristics of the individuals in the sample will mirror those in the population.

Parameters vs. Statistics

- ▶ In most studies, we wish to quantify some characteristic of the population.
- ▶ Example:
 - ▶ The retention rate, inquiry rate, average level of self-efficacy, average level of body image
- ▶ This is the population parameter
- ▶ Parameters are generally unknown, and must be estimated from a sample
- ▶ The sample estimate is called a statistic

Variables

- ▶ A characteristic that takes on different values for different individuals in a sample is called a variable.
- ▶ **Examples:**
 - ▶ Retention (yes/no)
 - ▶ Inquiry about NSU (yes/no)
 - ▶ Self-efficacy (score on self-efficacy questionnaire)
 - ▶ Body image (score on body image questionnaire)

Independent Variables (IV)

- ▶ The “explanatory” variable
- ▶ The variable that attempts to explain or is purported to cause differences in a second variable.
- ▶ In experimental designs, the intervention is the IV.
- ▶ Example:
 - ▶ Does a new curriculum improve body image?
 - ▶ The curriculum is the IV

Dependent Variables (DV)

- ▶ The “outcome” variable
- ▶ The variable that is **thought to be influenced** by the independent variable
- ▶ Example:
 - ▶ Does a new curriculum improve body image?
 - ▶ Body image is the DV

Examples

- ▶ Do students prefer learning statistics online or face to face?
 - ▶ What is the IV? DV?

- ▶ How do students who have never had statistics compare to students who have previously had statistics in terms of their anxiety levels?
 - ▶ What is the IV? DV?

Confounding Variables

- ▶ Researchers are usually only interested in the relationship between the IV and DV.
- ▶ Confounding variables represent unwanted sources of influence on the DV, and are sometimes referred to as “nuisance” variables.
- ▶ Example:
 - ▶ Does a new curriculum improve body image?
 - ▶ Such things as heredity, family background, previous counseling experiences, etc. can also impact the DV.

Controlling Confounding Variables

- ▶ Typically, researchers are interested in excluding, or controlling for, the effects of confounding variables.
- ▶ This is generally **not** a statistical issue, but is accomplished by the research design.
- ▶ Certain types of designs (e.g., experiments) better control the effects of confounding variables.
- ▶ If an experiment or an equivalent control group is not possible → ANCOVA



Scales of Measurement

Variable Measurement Scales

- ▶ For any given variable that we are interested in, there may be a variety of measurement scales that can be used:
 - ▶ What is your annual income? _____
 - ▶ What is your annual income?
 - a. 10,000-20,000
 - b. 20,000-30,000
 - c. 30,000-40,000
 - d. 40,000-50,000
 - e. 50,000 or above
- ▶ Variable measurement is the second factor that influences the choice of statistical procedure.

Scales of Measurement

- ▶ Nominal
- ▶ Ordinal
- ▶ Interval
- ▶ Ratio

Nominal Scale

- ▶ Observations fall into different **categories** or groups.
- ▶ Differences among categories are **qualitative**, not quantitative.
- ▶ Examples:
 - ▶ Gender
 - ▶ Ethnicity
 - ▶ Counseling method (cognitive vs. humanistic)
 - ▶ Retention (retained vs. not retained)

Ordinal Scale

- ▶ Categories can be rank ordered in terms of amount or magnitude.
- ▶ Categories possess an inherent order, but the amount of difference between categories is unknown.
- ▶ Examples:
 - ▶ Class standing
 - ▶ Letter grades (A,B,C,D,F)
 - ▶ Likert-scale survey responses (SD, D, N, A, SA)

Interval Scale

- ▶ Categories are ordered, but now the intervals for each category are exactly the same size.
- ▶ That is, the distance between measurement points represent equal magnitudes (e.g., the distance between point A and B is the same as the distance between B and C).
- ▶ Examples:
 - ▶ Fahrenheit scale of measuring temperature
 - ▶ Chronological scale of dates (1997 A.D.)
 - ▶ Standard scores (z-scores)

Ratio Scale

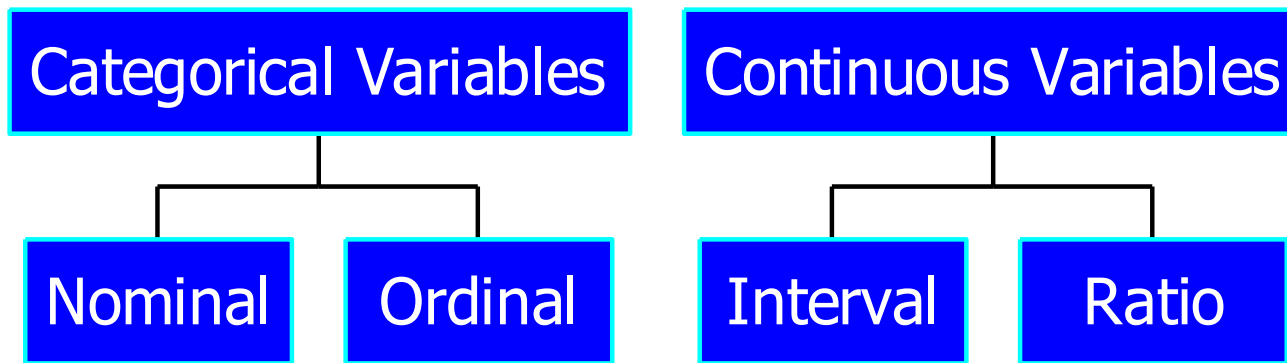
- ▶ Same properties as the interval scale, but with an additional feature
- ▶ Ratio scale has an absolute 0 point.
- ▶ Absolute 0 point permits the use of ratios (e.g., A is “twice as large” as B).
- ▶ Examples:
 - ▶ Number of children
 - ▶ Weight
 - ▶ Annual income

Categorical vs. Continuous Variables

- ▶ In practice, it is not usually necessary to make such fine distinctions between measurement scales.
- ▶ Two distinctions, categorical and continuous are usually sufficient.
- ▶ Categorical variables consist of separate, indivisible categories (i.e., men/women).
- ▶ Continuous variables yield values that fall on a numeric continuum, and can (theoretically) take on an infinite number of values.

Level of Measurement Summary

- ▶ In practice, the four levels of measurement can usually be classified as follows:



- ▶ Continuous variables are generally preferable because a wider range of statistical procedures can be applied

Examples

- ▶ What is the level of measurement of
 - ▶ Temperature $^{\circ}\text{C}$?
 - ▶ Color?
 - ▶ Income of professional baseball players?
 - ▶ Degree of agree (1 = Strongly Disagree, 5 = Strongly Agree)?

Descriptive Statistics

- ▶ Procedures used to **summarize, organize, and simplify** data (data being a collection of measurements or observations) taken from a sample (i.e., mean, median, mode).
- ▶ Examples:
 - ▶ The average score on the Rosenberg Self-Esteem Scale was 7.5
 - ▶ 63% of the sample described themselves as Caucasian

Inferential Statistics

- ▶ Techniques that allow us to make inferences about a population based on data that we gather from a sample.
- ▶ Study results will vary from sample to sample strictly due to random chance (i.e., sampling error).
- ▶ Inferential statistics allow us to determine how likely it is to obtain a set of results from a single sample.
- ▶ This is also known as testing for “statistical significance.”

Module 1 Summary

- ▶ The Role of Statistics
- ▶ Statistical Terminology
- ▶ Scales of Measurement
- ▶ Introduction to Descriptive and Inferential Statistics

Review Activity and Quiz

- ▶ Please complete the Module 1 Review Activity: Statistical Terminology located in Module 1.
- ▶ Upon completion of the Review Activity, please complete the Module 1 Quiz.
- ▶ Please note that all modules in this course build on one another; as a result, completion of the Module 1 Review Activity and Module 1 Quiz are required before moving on to Module 2.
- ▶ You can complete the review activities and quizzes as many times as you like.

Upcoming Modules

- ▶ Module 1: Introduction to Statistics
- ▶ Module 2: Introduction to SPSS
- ▶ Module 3: Descriptive Statistics
- ▶ Module 4: Inferential Statistics
- ▶ Module 5: Correlation
- ▶ Module 6: *t*-Tests
- ▶ Module 7: ANOVAs
- ▶ Module 8: Linear Regression
- ▶ Module 9: Nonparametric Procedures