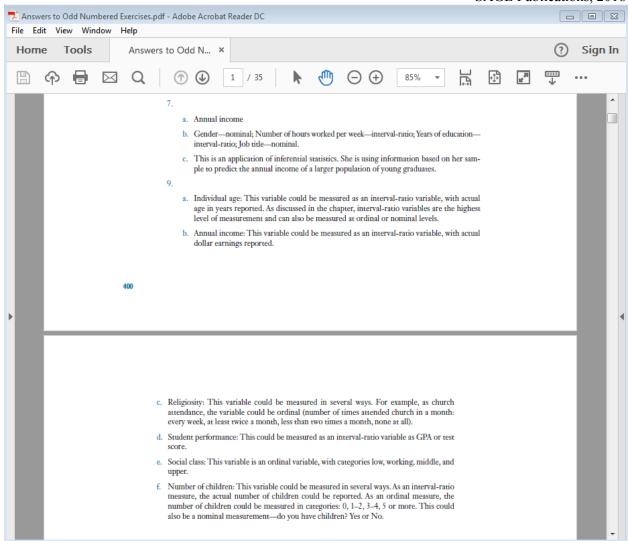


## Instructor Resource Frankfort-Nachmias, Social Statistics for a Diverse Society, 8e SAGE Publications, 2018



# Lecture Notes

# **Chapter 1: The What and the Why of Statistics**

# **Summary**

- Statistics are procedures used by social scientists to organize, summarize, and communicate numerical information. Only information represented by numbers can be the subject of statistical analysis.
- The research process is a set of activities in which social scientists engage to answer questions, examine ideas, or test theories. It consists of the following stages: asking the research question, formulating the hypotheses, collecting data, analyzing data, and evaluating the hypotheses.
- A theory is a set of assumptions and propositions used by social scientists to explain,
   predict, and understand social phenomena.
- Theories offer specific concrete predictions about the way observable attributes of people or groups would be interrelated in real life. These predictions, called hypotheses, are tentative answers to research problems.
- A variable is a property of people or objects that takes on two or more values. The
  variable that the researcher wants to explain (the "effect") is called the dependent
  variable. The variable that is expected to "cause" or account for the dependent variable is
  called the independent variable.
- Three conditions are required to establish causal relations: (1) The cause has to precede the effect in time; (2) there has to be an empirical relationship between the cause and the effect; and (3) this relationship cannot be explained by other factors.
- At the nominal level of measurement, numbers or other symbols are assigned to a set of categories to name, label, or classify the observations. At the ordinal level of measurement, categories can be rank ordered from low to high (or vice versa). At the

SAGE Publications, 2018

interval-ratio level of measurement, measurements for all cases are expressed in the same unit.

- A population is the total set of individuals, objects, groups, or events in which the researcher is interested. A sample is a relatively small subset selected from a population.
- Descriptive statistics includes procedures that help us organize and describe data
  collected from either a sample or a population. Inferential statistics is concerned with
  making predictions or inferences about a population from observations and analyses of a
  sample.

#### **Outline**

- The Research Process
  - A set of activities in which social scientists engage to answer questions, examine ideas, or test theories
- Asking Research Questions
  - The starting point for most research
  - o Questions can be answered by conducting empirical research
  - o The first step in selecting a research question involves considering what is interesting
- The Role of Theory
  - The relationship between attributes or characteristics of individuals and groups lies at the heart of social scientific inquiry
  - A theory is a set of assumptions and propositions used for explanation, prediction, and understanding of social phenomena.
  - The theory attempts to establish a link between what we observe (the data) and our conceptual understanding of why certain phenomena are related to each other in a particular way
- Formulating the Hypothesis
  - Hypothesis is a statement predicting the relationship between two or more observable attributes
  - They can be generated in many ways

- O Hypotheses often make statements about two variables
- o Variables must include categories that are exhaustive and mutually exclusive
- o Units of analysis are the level of social life on which social scientists focus
- The variable is a property of the unit of analysis
- o The dependent variable is the effect, or variable to be explained
- o The independent variable is expected to be the cause of the dependent variable
- o To establish that two variables are causally related, three conditions must be met:
  - The cause has to precede the effect in time
  - There has to be an empirical relationship between the cause and the effect
  - This relationship cannot be explained by other factors
- There are also guidelines for establishing the dependent and independent variables:
  - The dependent variable is always the property that you are trying to explain; it is always the object of the research
  - The independent variable usually occurs earlier in time than the dependent variable
  - The independent variable is often seen as influencing, directly or indirectly, the dependent variable

## • Collecting Data

 The choice of a particular data collection method or instrument to measure our variables depends on the study objective

### • Levels of Measurement

- At the nominal level of measurement, numbers or symbols are assigned a set of categories for the purpose of naming, labeling, or classifying the observations.
   Nominal categories cannot be rank-ordered.
- The ordinal level of measurement involves assigning numbers to rank-ordered categories ranging from low to high
- An interval-ratio level is achieved when the categories of a variable can be rankordered and the measurements for all cases are expressed in the same units

- Variables that can be measured at the interval-ratio level of measurement can also be measured at the ordinal and nominal levels
- As a rule, properties that can be measured at a higher level (interval-ratio is the highest) can also be measured at lower levels, but not vice versa
- Dichotomous variables are variables that have only two values and can be measured at the ordinal or interval-ratio level

#### Discrete and Continuous Variables

- Discrete variables have a minimum-sized unit of measurement which cannot be subdivided
- Continuous variables do not have a minimum-sized unit of measurement; their range of values can be subdivided into increasingly smaller fractional values
- This attribute of variables affects subsequent research operations, particularly
   measurement procedures, data analysis, and methods of inference and generalization

#### Measurement Error

- o Reliability and validity are important aspects of the research process
- Analyzing Data and Evaluating the Hypotheses
  - After collecting data, researchers must find a systematic way to organize these data,
     analyze them, and use some set of procedures to decide what they mean
  - A population is the total set of individuals, objects, groups, or events in which the researcher is interested
  - o A subset selected from a population is called a sample
  - Researchers usually collect their data from a sample and then generalize their observations to the larger population
  - Descriptive statistics are procedures that help us organize and describe data collected from either a sample or a population
  - Inferential statistics involves the logic and procedures concerned with making predictions or inferences about a population from observations and analyses of a sample

## • Evaluating the Hypotheses

Instructor Resource Frankfort-Nachmias, *Social Statistics for a Diverse Society, 8e* SAGE Publications, 2018

- o Assess and evaluate our hypothesis in light of analyzed data
- o Examine how data relate to the theoretical framework that guided our research
- o Statistics provides an important link between theory and research
- Looking at Social Differences
  - o How culture may influence how we count
  - o Are you anxious about statistics?