

# Chapter 1

## Introduction to Statistics

Likan Zhan

Beijing Language and Culture University

2021-09-04

<https://likan.info>  
[zhanlikan@bncu.edu.cn](mailto:zhanlikan@bncu.edu.cn)

# Table of Contents

1. Statistics, Science, and Observations
2. Data Structures, Research Methods, and Statistics
3. Variables and Measurement
4. Statistical Notation

# Table of Contents

1. Statistics, Science, and Observations
2. Data Structures, Research Methods, and Statistics
3. Variables and Measurement
4. Statistical Notation

# Definitions of Statistics

# Definitions of Statistics

- The term **statistics** (统计学) refers to a set of *mathematical procedures* for *organizing, summarizing, and interpreting* information.

# Definitions of Statistics

# Definitions of Statistics

- Statistics are used to *organize* and *summarize* the information so that the researcher can *see* what happened in the research study and can *communicate* the results to others.

# Definitions of Statistics

- Statistics are used to *organize* and *summarize* the information so that the researcher can *see* what happened in the research study and can *communicate* the results to others.
- Statistics help the researcher to answer the questions that initiated the research by determining exactly what *general conclusions* are justified based on the specific results that were obtained.

# Definitions of Statistics

- Statistics are used to *organize* and *summarize* the information so that the researcher can *see* what happened in the research study and can *communicate* the results to others.
- Statistics help the researcher to answer the questions that initiated the research by determining exactly what *general conclusions* are justified based on the specific results that were obtained.
- Statistical procedures help ensure that the information or observations are presented and interpreted in an accurate and informative way.

# Some Important Concepts

# Some Important Concepts

- Populations and Samples

# Some Important Concepts

- Populations and Samples
- Variables and Data

# Some Important Concepts

- Populations and Samples
- Variables and Data
- Parameters and Statistics

# Some Important Concepts

- Populations and Samples
- Variables and Data
- Parameters and Statistics
- Descriptive and Inferential Statistics

# Populations and Samples

## THE POPULATION

All of the individuals of interest

# Populations and Samples

## THE POPULATION

All of the individuals of interest



The sample  
is selected from  
the population

# Populations and Samples

## THE POPULATION

All of the individuals of interest

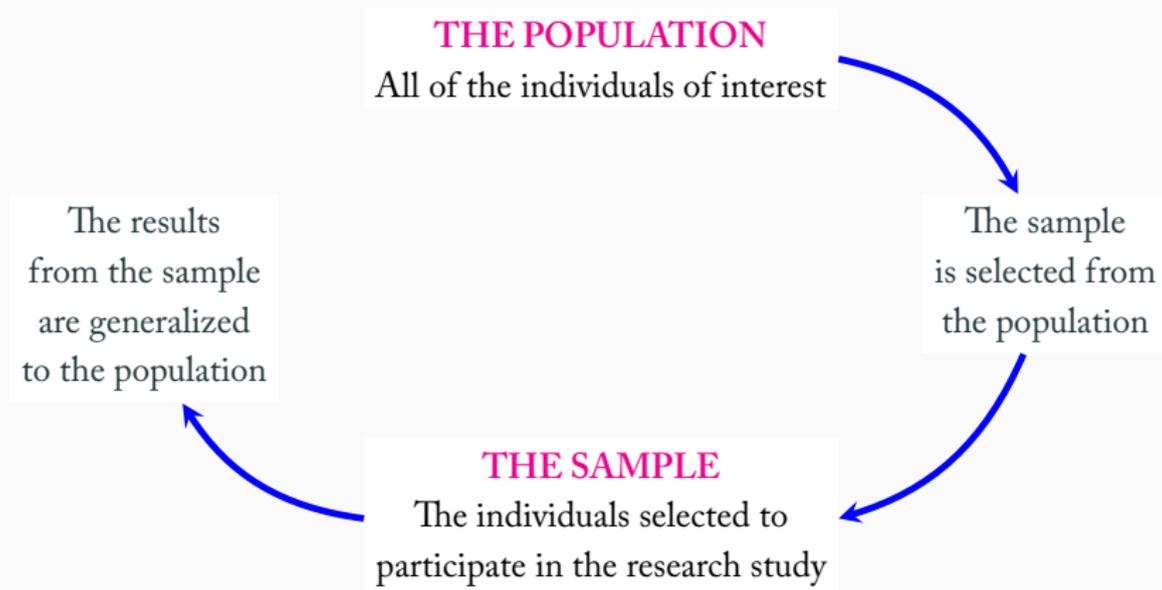
The sample  
is selected from  
the population

```
graph TD; A["THE POPULATION  
All of the individuals of interest"] --> B["The sample  
is selected from  
the population"]; B --> C["THE SAMPLE  
The individuals selected to  
participate in the research study"]
```

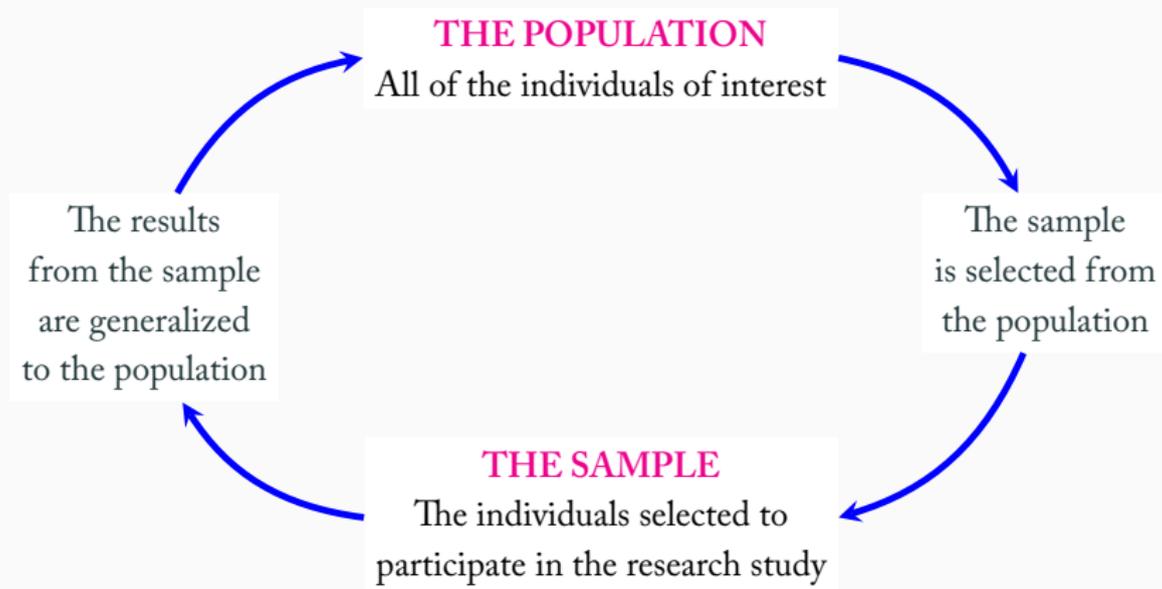
## THE SAMPLE

The individuals selected to  
participate in the research study

# Populations and Samples



# Populations and Samples



# Populations and Samples

# Populations and Samples

- Research in the behavioral sciences typically begins with a general question about a specific group (or groups) of individuals.

# Populations and Samples

- Research in the behavioral sciences typically begins with a general question about a specific group (or groups) of individuals.
- A **population** (整体) is the set of all the individuals of interest in a particular study.

# Populations and Samples

# Populations and Samples

- Because populations tend to be very large, it usually is impossible for a researcher to examine every individual in the population of interest.

# Populations and Samples

- Because populations tend to be very large, it usually is impossible for a researcher to examine every individual in the population of interest.
- Therefore, researchers typically select a smaller, more manageable group from the population and limit their studies to the individuals in the selected group.

# Populations and Samples

- Because populations tend to be very large, it usually is impossible for a researcher to examine every individual in the population of interest.
- Therefore, researchers typically select a smaller, more manageable group from the population and limit their studies to the individuals in the selected group.
- A **sample** (样本) is a set of individuals selected from a population, usually intended to represent the population in a research study.

# Variables and Data

# Variables and Data

- Typically, researchers are interested in specific characteristics of the individuals in the population (or in the sample), or they are interested in outside factors that may influence the individuals.

# Variables and Data

- Typically, researchers are interested in specific characteristics of the individuals in the population (or in the sample), or they are interested in outside factors that may influence the individuals.
- A **variable** (变量) is a characteristic or condition that changes or has different values for different individuals.

# Variables and Data

## Variables and Data

- To demonstrate changes in variables, it is necessary to make measurements of the variables being examined.

# Variables and Data

- To demonstrate changes in variables, it is necessary to make measurements of the variables being examined.
- **Data** (plural) (数据) are measurements or observations.

# Variables and Data

- To demonstrate changes in variables, it is necessary to make measurements of the variables being examined.
- **Data** (plural) (数据) are measurements or observations.
- A **data set** is a collection of measurements or observations.

# Variables and Data

- To demonstrate changes in variables, it is necessary to make measurements of the variables being examined.
- **Data** (plural) (数据) are measurements or observations.
- A **data set** is a collection of measurements or observations.
- A **datum** (singular) is a single measurement or observation and is commonly called a score or raw score.

# Parameters and Statistics

# Parameters and Statistics

- When describing data it is necessary to distinguish whether the data come from a population or a sample.

# Parameters and Statistics

- When describing data it is necessary to distinguish whether the data come from a population or a sample.
- A **parameter** (参数) is a value, usually a numerical value, that describes a population.

# Parameters and Statistics

- When describing data it is necessary to distinguish whether the data come from a population or a sample.
- A **parameter** (参数) is a value, usually a numerical value, that describes a population. A parameter is usually derived from measurements of the individuals in the population.

# Parameters and Statistics

- When describing data it is necessary to distinguish whether the data come from a population or a sample.
- A **parameter** (参数) is a value, usually a numerical value, that describes a population. A parameter is usually derived from measurements of the individuals in the population.
- A **statistic** (统计量) is a value, usually a numerical value, that describes a sample.

# Parameters and Statistics

- When describing data it is necessary to distinguish whether the data come from a population or a sample.
- A **parameter** (参数) is a value, usually a numerical value, that describes a population. A parameter is usually derived from measurements of the individuals in the population.
- A **statistic** (统计量) is a value, usually a numerical value, that describes a sample. A statistic is usually derived from measurements of the individuals in the sample.

# Descriptive and Inferential Statistical Methods

# Descriptive and Inferential Statistical Methods

- Although researchers have developed a variety of different statistical procedures to organize and interpret data, these different procedures can be classified into two general categories:

# Descriptive and Inferential Statistical Methods

- Although researchers have developed a variety of different statistical procedures to organize and interpret data, these different procedures can be classified into two general categories:
- **Descriptive statistics** (描述统计) are statistical procedures used to summarize, organize, and simplify data.

# Descriptive and Inferential Statistical Methods

- Although researchers have developed a variety of different statistical procedures to organize and interpret data, these different procedures can be classified into two general categories:
- **Descriptive statistics** (描述统计) are statistical procedures used to summarize, organize, and simplify data.
- **Inferential statistics** (推论统计) consist of techniques that allow us to study samples and then make generalizations about the populations from which they were selected.

# Sampling error

## Sampling error

- Because populations are typically very large, it usually is not possible to measure everyone in the population. Therefore, a sample is selected to represent the population.

## Sampling error

- Because populations are typically very large, it usually is not possible to measure everyone in the population. Therefore, a sample is selected to represent the population.
- There usually is some discrepancy between a sample statistic and the corresponding population parameter.

## Sampling error

- Because populations are typically very large, it usually is not possible to measure everyone in the population. Therefore, a sample is selected to represent the population.
- There usually is some discrepancy between a sample statistic and the corresponding population parameter.
- A **sampling error** (取样误差) is the naturally occurring discrepancy, or error, that exists between a sample statistic and the corresponding population parameter.

# Sampling error

- Population

```
## [1] 90 102 87 123 104 87 107 111 108 95  
## [1] "Population Mean: 101.4"
```

- One Sample

```
## [1] 108 123 107 90 102  
## [1] "Sample Mean: 106"
```

# Relation among these Concepts

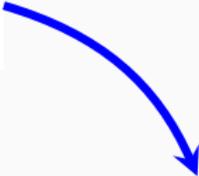
**THE POPULATION**

Parameters

# Relation among these Concepts

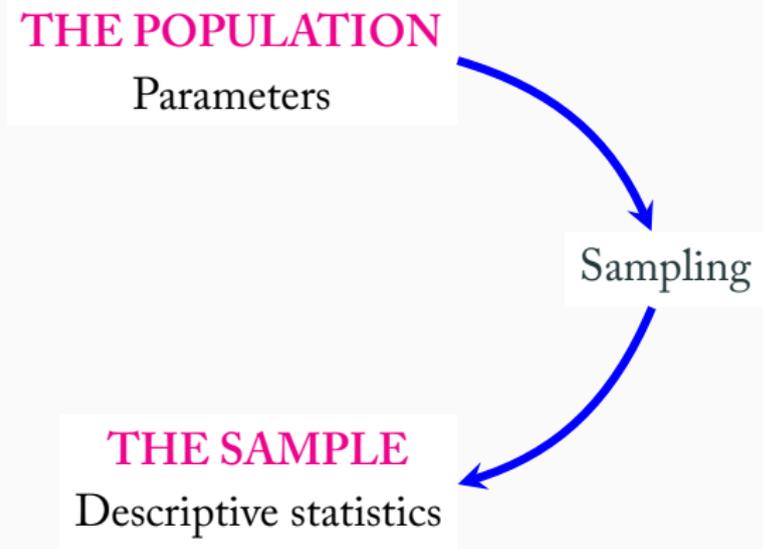
**THE POPULATION**

Parameters

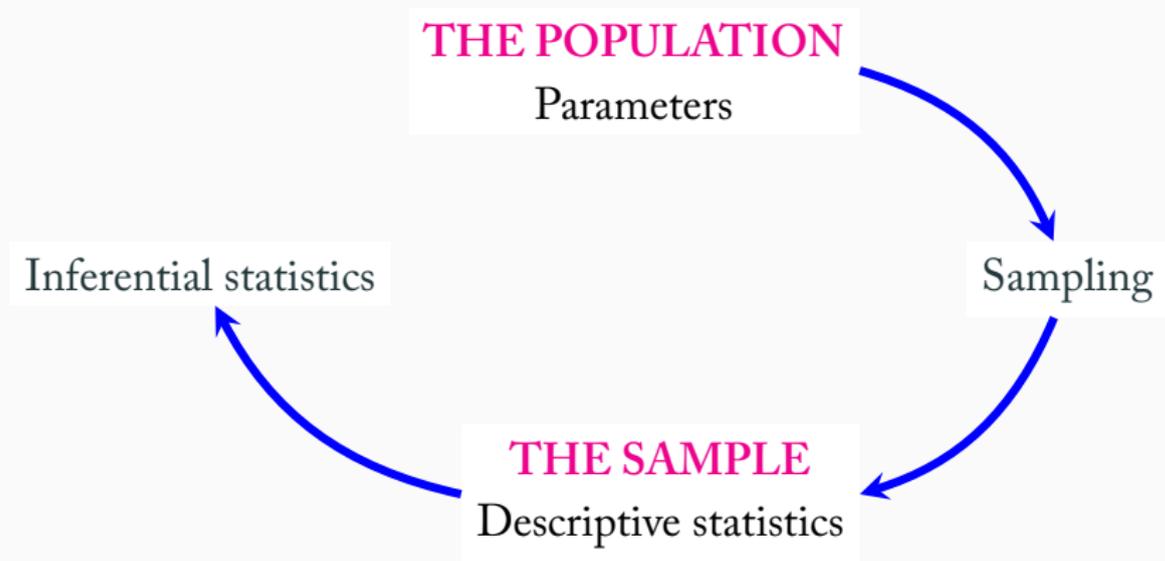


Sampling

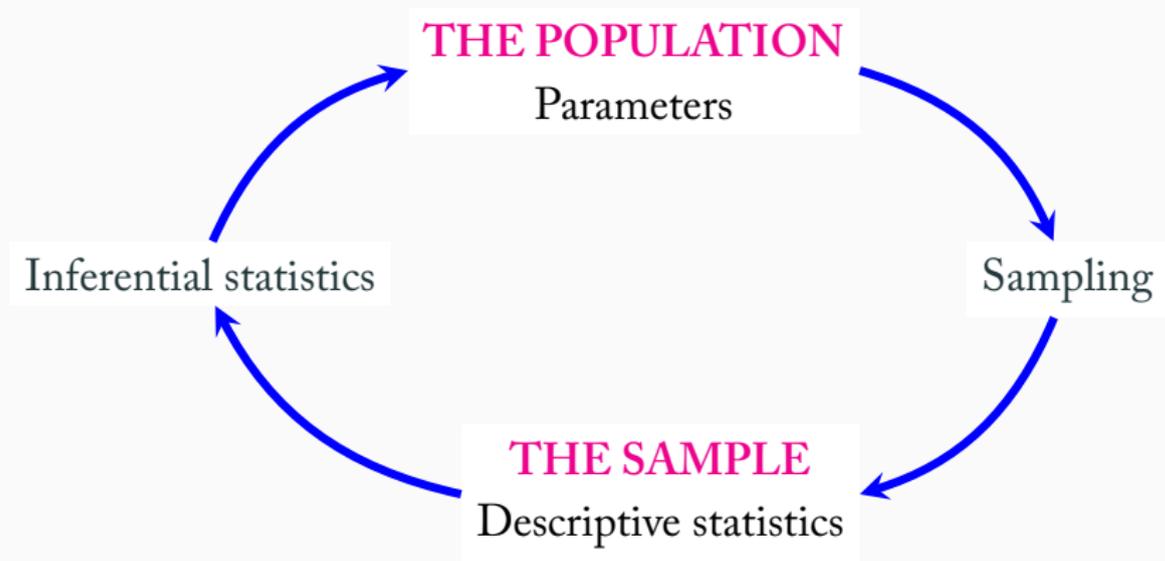
# Relation among these Concepts



# Relation among these Concepts



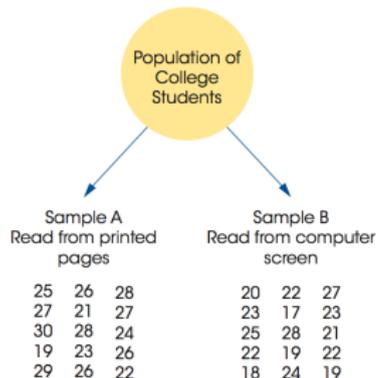
# Relation among these Concepts



# Statistics in the Context of Research

**Step 1**  
*Experiment:*  
Compare two  
studying methods

*Data*  
Test scores for the  
students in each  
sample



**Step 2**  
*Descriptive statistics:*  
Organize and simplify



**Step 3**  
*Inferential statistics:*  
Interpret results

The sample data show a 4-point difference between the two methods of studying. However, there are two ways to interpret the results.

1. There actually is no difference between the two studying methods, and the sample difference is due to chance (sampling error).
2. There really is a difference between the two methods, and the sample data accurately reflect this difference.

The goal of inferential statistics is to help researchers decide between the two interpretations.

# Table of Contents

1. Statistics, Science, and Observations
2. Data Structures, Research Methods, and Statistics
3. Variables and Measurement
4. Statistical Notation

# Research Methods

# Research Methods

- Individual Variables: Descriptive Research

# Research Methods

- Individual Variables: Descriptive Research
- Relationships Between Variables

# Research Methods

- Individual Variables: Descriptive Research
- Relationships Between Variables
  - One Group with Two Variables Measured for Each Individual: The Correlational Method

# Research Methods

- Individual Variables: Descriptive Research
- Relationships Between Variables
  - One Group with Two Variables Measured for Each Individual: The Correlational Method
  - Comparing Two (or More) Groups of Scores: Experimental and Nonexperimental Methods

# Research Methods

# Research Methods

- Descriptive Research: Some research studies are conducted simply to describe individual variables as they exist naturally.

# Research Methods

- **Descriptive Research:** Some research studies are conducted simply to describe individual variables as they exist naturally.
- **Relationships Between Variables:** Most research, however, is intended to examine relationships between two or more variables.

# Relationships Between Variables

## Relationships Between Variables

- To establish the existence of a relationship, researchers must make observations - that is, measurements of the two variables.

## Relationships Between Variables

- To establish the existence of a relationship, researchers must make observations - that is, measurements of the two variables.
- The resulting measurements can be classified into two distinct data structures that also help to classify different research methods and different statistical techniques.

# Relationships Between Variables

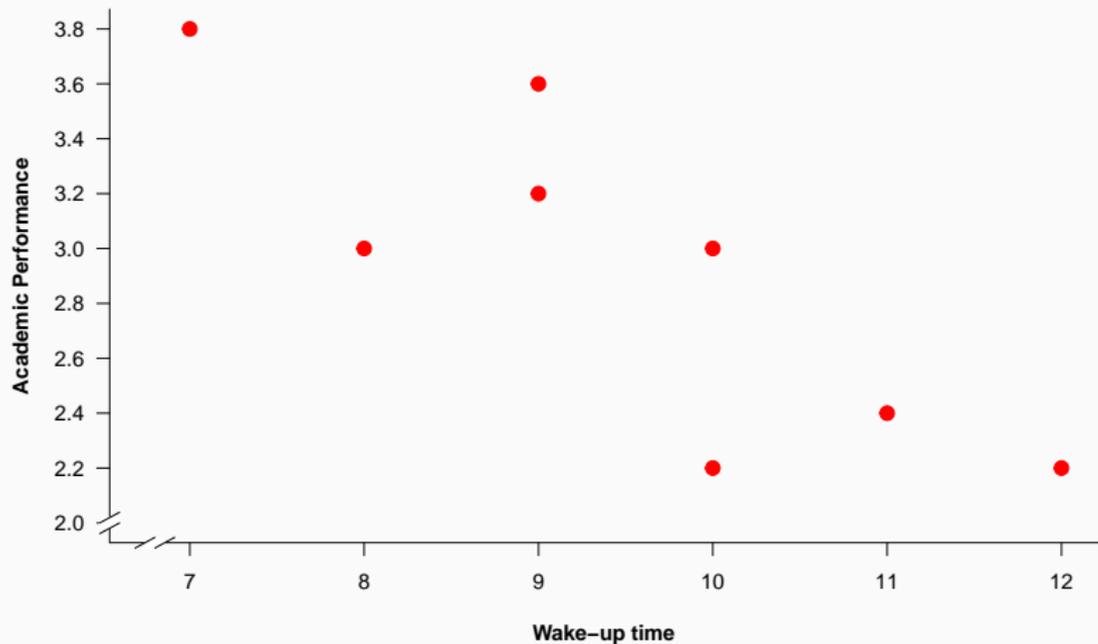
## Relationships Between Variables

- In the **Correlational Method** (相关方法), two different variables are observed to determine whether there is a relationship between them.

## Statistics for the Correlational Method

Student	Wake_up_time	Academic_Performance
A	11	2.4
B	9	3.6
C	9	3.2
D	12	2.2
E	7	3.8
F	10	2.2
G	10	3.0
H	8	3.0

# Statistics for the Correlational Method



# Statistics for the Correlational Method

- Occasionally, the measurement process used for a correlational study simply classifies individuals into categories that do not correspond to numerical values.
- For example,

**Table 1: Cell Phone Preference**

	Text	Talk	
Males	30	20	50
Females	25	25	50

# Relationships Between Variables

## Relationships Between Variables

- The results from a correlational study can demonstrate the existence of a relationship between two variables, but they do not provide an explanation for the relationship.

## Relationships Between Variables

- The results from a correlational study can demonstrate the existence of a relationship between two variables, but they do not provide an explanation for the relationship.
- In particular, a correlational study cannot demonstrate a cause-and-effect relationship.

# Experimental and Nonexperimental Methods

# Experimental and Nonexperimental Methods

- The second method for examining the relationship between two variables involves the comparison of two or more groups of scores.

## Experimental and Nonexperimental Methods

- The second method for examining the relationship between two variables involves the comparison of two or more groups of scores.
- In this situation, the relationship between variables is examined by using one of the variables to define the groups, and then measuring the second variable to obtain scores for each group.

# Experimental and Nonexperimental Methods

# Experimental and Nonexperimental Methods

- There are two distinct research methods that both produce groups of scores to be compared: the experimental and the nonexperimental strategies.

# Experimental and Nonexperimental Methods

- There are two distinct research methods that both produce groups of scores to be compared: the experimental and the nonexperimental strategies.
- These two research methods use exactly the same statistics and they both demonstrate a relationship between two variables.

# Experimental and Nonexperimental Methods

- There are two distinct research methods that both produce groups of scores to be compared: the experimental and the nonexperimental strategies.
- These two research methods use exactly the same statistics and they both demonstrate a relationship between two variables.
- The distinction between the two research strategies is how the relationship is interpreted.

# Experimental and Nonexperimental Methods

- There are two distinct research methods that both produce groups of scores to be compared: the experimental and the nonexperimental strategies.
- These two research methods use exactly the same statistics and they both demonstrate a relationship between two variables.
- The distinction between the two research strategies is how the relationship is interpreted.
- The results from an experiment allow a cause-and-effect explanation.

## Experimental and Nonexperimental Methods

- There are two distinct research methods that both produce groups of scores to be compared: the experimental and the nonexperimental strategies.
- These two research methods use exactly the same statistics and they both demonstrate a relationship between two variables.
- The distinction between the two research strategies is how the relationship is interpreted.
- The results from an experiment allow a cause-and-effect explanation.
- A nonexperimental study does not permit a cause-and-effect explanation.

# The Experimental Method

ID	Violent	Nonviolent
1	7	8
2	8	4
3	10	8
4	7	3
5	9	6
6	8	5
7	6	3
8	10	4
9	9	4
10	6	5

# The Experimental Method

ID	Type_of_Video_Game	Aggressive_Behavior
1	Violent	7
2	Violent	8
3	Violent	10
4	Violent	7
5	Violent	9
6	Violent	8
7	Violent	6
8	Violent	10
9	Violent	9
10	Violent	6
1	Nonviolent	8
2	Nonviolent	4
3	Nonviolent	8
4	Nonviolent	3
5	Nonviolent	6
6	Nonviolent	5
7	Nonviolent	3
8	Nonviolent	4
9	Nonviolent	4
10	Nonviolent	5

# The Experimental Method

# The Experimental Method

- The goal of an experimental study is to demonstrate a cause-and-effect relationship between two variables.

# The Experimental Method

- The goal of an experimental study is to demonstrate a cause-and-effect relationship between two variables.
- Specifically, an experiment attempts to show that changing the value of one variable causes changes to occur in the second variable.

# The Experimental Method

- The goal of an experimental study is to demonstrate a cause-and-effect relationship between two variables.
- Specifically, an experiment attempts to show that changing the value of one variable causes changes to occur in the second variable.
- To accomplish this goal, the experimental method has two characteristics that differentiate experiments from other types of research studies:

# The Experimental Method

# The Experimental Method

- **Manipulation** (操纵): The researcher manipulates one variable by changing its value from one level to another.

# The Experimental Method

- **Manipulation** (操纵): The researcher manipulates one variable by changing its value from one level to another. A second variable is observed (measured) to determine whether the manipulation causes changes to occur.

# The Experimental Method

- **Manipulation** (操纵): The researcher manipulates one variable by changing its value from one level to another. A second variable is observed (measured) to determine whether the manipulation causes changes to occur.
- **Control** (控制): The researcher must exercise control over the research situation to ensure that other, extraneous variables do not influence the relationship being examined.

# The Experimental Method

There are two general categories of variables that researchers must consider:

# The Experimental Method

There are two general categories of variables that researchers must consider:

- **Participant Variables** (被试变量): These are characteristics such as age, gender, and intelligence that vary from one individual to another.

# The Experimental Method

There are two general categories of variables that researchers must consider:

- **Participant Variables** (被试变量): These are characteristics such as age, gender, and intelligence that vary from one individual to another.
- **Environmental Variables** (环境变量): These are characteristics of the environment such as lighting, time of day, and weather conditions.

# The Experimental Method

Researchers typically use three basic techniques to control other variables:

# The Experimental Method

Researchers typically use three basic techniques to control other variables:

- **Random assignment** (随机指派): Each participant has an equal chance of being assigned to each of the treatment conditions.

# The Experimental Method

Researchers typically use three basic techniques to control other variables:

- **Random assignment** (随机指派): Each participant has an equal chance of being assigned to each of the treatment conditions.
- A second technique for controlling variables is to use **Matching** (匹配) to ensure equivalent groups or equivalent environments.

# The Experimental Method

Researchers typically use three basic techniques to control other variables:

- **Random assignment** (随机指派): Each participant has an equal chance of being assigned to each of the treatment conditions.
- A second technique for controlling variables is to use **Matching** (匹配) to ensure equivalent groups or equivalent environments.
- Finally, the researcher can control variables by holding the control variables **constant** (保持恒定).

# The Experimental Method: Summary

# The Experimental Method: Summary

- In the **experimental method**, one variable is manipulated while another variable is observed and measured.

## The Experimental Method: Summary

- In the **experimental method**, one variable is manipulated while another variable is observed and measured.
- To establish a cause-and-effect relationship between the two variables, an experiment attempts to control all other variables to prevent them from influencing the results.

# Terminology in the Experimental Method

## Terminology in the Experimental Method

- An experimental study evaluates the relationship between two variables by manipulating one variable (the independent variable) and measuring one variable (the dependent variable).

## Terminology in the Experimental Method

- An experimental study evaluates the relationship between two variables by manipulating one variable (the independent variable) and measuring one variable (the dependent variable).
- The **independent variable** (自变量) is the variable that is manipulated by the researcher.

## Terminology in the Experimental Method

- An experimental study evaluates the relationship between two variables by manipulating one variable (the independent variable) and measuring one variable (the dependent variable).
- The **independent variable** (自变量) is the variable that is manipulated by the researcher. In behavioral research, the independent variable usually consists of the two (or more) treatment conditions to which subjects are exposed.

## Terminology in the Experimental Method

- An experimental study evaluates the relationship between two variables by manipulating one variable (the independent variable) and measuring one variable (the dependent variable).
- The **independent variable** (自变量) is the variable that is manipulated by the researcher. In behavioral research, the independent variable usually consists of the two (or more) treatment conditions to which subjects are exposed. The independent variable consists of the antecedent conditions that were manipulated prior to observing the dependent variable.

## Terminology in the Experimental Method

- An experimental study evaluates the relationship between two variables by manipulating one variable (the independent variable) and measuring one variable (the dependent variable).
- The **independent variable** (自变量) is the variable that is manipulated by the researcher. In behavioral research, the independent variable usually consists of the two (or more) treatment conditions to which subjects are exposed. The independent variable consists of the antecedent conditions that were manipulated prior to observing the dependent variable.
- The **dependent variable** (因变量) is the one that is observed to assess the effect of the treatment.

# Terminology in the Experimental Method

# Terminology in the Experimental Method

- Individuals in a The **control condition** (控制条件) do not receive the experimental treatment.

## Terminology in the Experimental Method

- Individuals in a The **control condition** (控制条件) do not receive the experimental treatment. Instead, they either receive no treatment or they receive a neutral, placebo treatment.

## Terminology in the Experimental Method

- Individuals in a The **control condition** (控制条件) do not receive the experimental treatment. Instead, they either receive no treatment or they receive a neutral, placebo treatment. The purpose of a control condition is to provide a baseline for comparison with the experimental condition.

## Terminology in the Experimental Method

- Individuals in a The **control condition** (控制条件) do not receive the experimental treatment. Instead, they either receive no treatment or they receive a neutral, placebo treatment. The purpose of a control condition is to provide a baseline for comparison with the experimental condition.
- Individuals in the **experimental condition** (实验条件) do receive the experimental treatment.

# Nonexperimental Methods

# Nonexperimental Methods

- A real experiment must include manipulation of an independent variable and rigorous control of other, extraneous variables.

# Nonexperimental Methods

- A real experiment must include manipulation of an independent variable and rigorous control of other, extraneous variables.
- As a result, there are a number of other research designs that are not true experiments but still examine the relationship between variables by comparing groups of scores.

# Nonexperimental Methods

- A real experiment must include manipulation of an independent variable and rigorous control of other, extraneous variables.
- As a result, there are a number of other research designs that are not true experiments but still examine the relationship between variables by comparing groups of scores.
- Two examples of nonexperimental studies that involve comparing two groups of scores:

## Preexisting or time groups

ID	Boys (Before therapy)	Girls (After therapy)
1	17	12
2	19	10
3	16	14
4	12	15
5	17	13
6	18	12
7	15	11
8	16	13

---

# Nonexperimental Methods

## Nonexperimental Methods

- In a nonexperimental study, the “independent variable” that is used to create the different groups of scores is often called the **quasi-independent variable** (伪自变量).

# Table of Contents

1. Statistics, Science, and Observations
2. Data Structures, Research Methods, and Statistics
3. Variables and Measurement
4. Statistical Notation

# Constructs and Operational Definitions

# Constructs and Operational Definitions

- **Constructs** (建构) are internal attributes or characteristics that cannot be directly observed but are useful for describing and explaining behavior.

# Constructs and Operational Definitions

# Constructs and Operational Definitions

- An **operational definition** (操作定义) identifies a measurement procedure (a set of operations) for measuring an external behavior and uses the resulting measurements as a definition and a measurement of a hypothetical construct.

# Constructs and Operational Definitions

- An **operational definition** (操作定义) identifies a measurement procedure (a set of operations) for measuring an external behavior and uses the resulting measurements as a definition and a measurement of a hypothetical construct.
- Note that an operational definition has two components.

## Constructs and Operational Definitions

- An **operational definition** (操作定义) identifies a measurement procedure (a set of operations) for measuring an external behavior and uses the resulting measurements as a definition and a measurement of a hypothetical construct.
- Note that an operational definition has two components. First, it describes a set of operations for measuring a construct.

## Constructs and Operational Definitions

- An **operational definition** (操作定义) identifies a measurement procedure (a set of operations) for measuring an external behavior and uses the resulting measurements as a definition and a measurement of a hypothetical construct.
- Note that an operational definition has two components. First, it describes a set of operations for measuring a construct. Second, it defines the construct in terms of the resulting measurements.

# Discrete and Continuous Variables

# Discrete and Continuous Variables

- The variables in a study can be characterized by the type of values that can be assigned to them.

## Discrete and Continuous Variables

- The variables in a study can be characterized by the type of values that can be assigned to them.
- A **discrete variable** (离散变量) consists of separate, indivisible categories. No values can exist between two neighboring categories.

# Discrete and Continuous Variables

# Discrete and Continuous Variables

- For a **continuous variable** (连续变量), there are an infinite number of possible values that fall between any two observed values.

## Discrete and Continuous Variables

- For a **continuous variable** (连续变量), there are an infinite number of possible values that fall between any two observed values. A continuous variable is divisible into an infinite number of fractional parts.

## Discrete and Continuous Variables

- For a **continuous variable** (连续变量), there are an infinite number of possible values that fall between any two observed values. A continuous variable is divisible into an infinite number of fractional parts.
- When measuring a continuous variable, it should be very rare to obtain identical measurements for two different individuals.

## Discrete and Continuous Variables

- For a **continuous variable** (连续变量), there are an infinite number of possible values that fall between any two observed values. A continuous variable is divisible into an infinite number of fractional parts.
- When measuring a continuous variable, it should be very rare to obtain identical measurements for two different individuals.
- When measuring a continuous variable, each measurement category is actually an interval that must be defined by boundaries.

# Continuous variables and real limits

## Continuous variables and real limits

- **Real limits** (精确阈限) are the boundaries of intervals for scores that are represented on a continuous number line. The real limit separating two adjacent scores is located exactly halfway between the scores.

## Continuous variables and real limits

- **Real limits** (精确阈限) are the boundaries of intervals for scores that are represented on a continuous number line. The real limit separating two adjacent scores is located exactly halfway between the scores.
- Each score has two real limits. The **upper real limit** (精确上限) is at the top of the interval, and the **lower real limit** (精确下限) is at the bottom.

# Continuous variables and real limits



# Scales of Measurement

## Scales of Measurement

- The terms continuous and discrete apply to the variables that are being measured and not to the scores that are obtained from the measurement.

## Scales of Measurement

- The terms continuous and discrete apply to the variables that are being measured and not to the scores that are obtained from the measurement.
- Data collection requires that we make measurements of our observations. **Measurement** (測量) involves assigning individuals or events to categories.

## Scales of Measurement

- The terms continuous and discrete apply to the variables that are being measured and not to the scores that are obtained from the measurement.
- Data collection requires that we make measurements of our observations. **Measurement** (测量) involves assigning individuals or events to categories.
- The categories used to measure a variable make up a **scale of measurement** (测量的水平), and the relationships between the categories determine different types of scales.

# Scales of Measurement

## Scales of Measurement

- A **nominal scale** (称名尺度) consists of a set of categories that have different names.

## Scales of Measurement

- A **nominal scale** (称名尺度) consists of a set of categories that have different names.
- Measurements on a nominal scale label and categorize observations, but do not make any quantitative distinctions between observations.

# Scales of Measurement

- A **nominal scale** (称名尺度) consists of a set of categories that have different names.
- Measurements on a nominal scale label and categorize observations, but do not make any quantitative distinctions between observations.
- An **ordinal scale** (顺序尺度) consists of a set of categories that are organized in an ordered sequence.

# Scales of Measurement

- A **nominal scale** (称名尺度) consists of a set of categories that have different names.
- Measurements on a nominal scale label and categorize observations, but do not make any quantitative distinctions between observations.
- An **ordinal scale** (顺序尺度) consists of a set of categories that are organized in an ordered sequence.
- Measurements on an ordinal scale rank observations in terms of size or magnitude.

# Scales of Measurement

## Scales of Measurement

- An **interval scale** (等距尺度) consists of ordered categories that are all intervals of exactly the same size.

## Scales of Measurement

- An **interval scale** (等距尺度) consists of ordered categories that are all intervals of exactly the same size.
- Equal differences between numbers on scale reflect equal differences in magnitude. However, the zero point on an interval scale is arbitrary and does not indicate a zero amount of the variable being measured.

## Scales of Measurement

- An **interval scale** (等距尺度) consists of ordered categories that are all intervals of exactly the same size.
- Equal differences between numbers on scale reflect equal differences in magnitude. However, the zero point on an interval scale is arbitrary and does not indicate a zero amount of the variable being measured.
- A **ratio scale** (等比尺度) is an interval scale with the additional feature of an absolute zero point.

## Scales of Measurement

- An **interval scale** (等距尺度) consists of ordered categories that are all intervals of exactly the same size.
- Equal differences between numbers on scale reflect equal differences in magnitude. However, the zero point on an interval scale is arbitrary and does not indicate a zero amount of the variable being measured.
- A **ratio scale** (等比尺度) is an interval scale with the additional feature of an absolute zero point.
- With a ratio scale, ratios of numbers do reflect ratios of magnitude.

# Statistics and Scales of Measurement

# Statistics and Scales of Measurement

- Scales of measurement are important because they help determine the statistics that are used to evaluate the data.

# Statistics and Scales of Measurement

- Scales of measurement are important because they help determine the statistics that are used to evaluate the data.
- Specifically, there are certain statistical procedures that are used with numerical scores from interval or ratio scales and other statistical procedures that are used with non-numerical scores from nominal or ordinal scales.

# Table of Contents

1. Statistics, Science, and Observations
2. Data Structures, Research Methods, and Statistics
3. Variables and Measurement
4. Statistical Notation

# Scores

# Scores

- Measuring a variable in a research study yields a value or a score for each individual.

# Scores

- Measuring a variable in a research study yields a value or a score for each individual.
- **Raw scores** (原始分数) are the original, unchanged scores obtained in the study.

# Scores

- Measuring a variable in a research study yields a value or a score for each individual.
- **Raw scores** (原始分数) are the original, unchanged scores obtained in the study.
- Scores for a particular variable are typically represented by the letter **X**. A set of scores can be presented in a column that is headed by **X**.

# Quiz Scores

<u>X</u>
37
35
35
30
25
17
<u>16</u>

# Scores

# Scores

- When observations are made for two variables, there will be two scores for each individual. The data can be presented as two lists labeled **X** and **Y** for the two variables.

# Scores

- When observations are made for two variables, there will be two scores for each individual. The data can be presented as two lists labeled **X** and **Y** for the two variables.
- The letter **N** is used to specify how many scores are in a set. An uppercase letter **N** identifies the number of scores in a population and a lowercase letter **n** identifies the number of scores in a sample.

# Height and Weight

X	Y
72	165
68	151
67	160
67	160
68	146
70	160
66	133

# Summation Notation

## Summation Notation

- The **summation sign**  $\Sigma$  (求和符号) can be read as “the sum of”.

# Summation Notation

- The **summation sign**  $\Sigma$  (求和符号) can be read as “the sum of”.
- The summation sign is always followed by a symbol or mathematical expression.

# Summation Notation

- The **summation sign**  $\Sigma$  (求和符号) can be read as “the sum of”.
- The summation sign is always followed by a symbol or mathematical expression.
- The summation process is often included with several other mathematical operations, such as multiplication or squaring.

# Order of Mathematical Operations

# Order of Mathematical Operations

- Any calculation contained within **parentheses** is done first.

# Order of Mathematical Operations

- Any calculation contained within **parentheses** is done first.
- **Squaring** (or raising to other exponents) is done second.

# Order of Mathematical Operations

- Any calculation contained within **parentheses** is done first.
- **Squaring** (or raising to other exponents) is done second.
- **Multiplying** and/or **dividing** is done third. A series of multiplication and/or division operations should be done in order from left to right.

# Order of Mathematical Operations

- Any calculation contained within **parentheses** is done first.
- **Squaring** (or raising to other exponents) is done second.
- **Multiplying** and/or **dividing** is done third. A series of multiplication and/or division operations should be done in order from left to right.
- **Summation** using the  $\Sigma$  notation is done next.

# Order of Mathematical Operations

- Any calculation contained within **parentheses** is done first.
- **Squaring** (or raising to other exponents) is done second.
- **Multiplying** and/or **dividing** is done third. A series of multiplication and/or division operations should be done in order from left to right.
- **Summation** using the  $\Sigma$  notation is done next.
- Finally, any other **addition** and/or **subtraction** is done.

## A simple example

- Suppose the following data set

$$\begin{array}{r} \hline X \\ \hline 3 \\ 1 \\ 7 \\ 4 \\ \hline \end{array}$$

- Calculate:  $\Sigma X$ ,  $\Sigma X^2$ ,  $(\Sigma X)^2$ ,  $\Sigma(X - 1)$ , and  $\Sigma(X - 1)^2$
- The results:  $\Sigma X = 15$ ;  $\Sigma X^2 = 75$ ;  $(\Sigma X)^2 = 225$ ;  $\Sigma(X - 1) = 11$ ;  $\Sigma(X - 1)^2 = 49$ .

Questions?