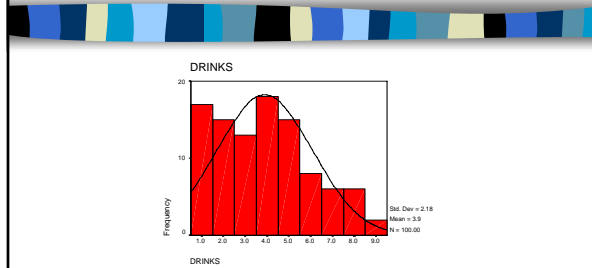


Introduction to Statistics



Analyzing Human Behavior

- Social sciences also called soft sciences
- No such thing as THE ANSWER
- Find what is MOST PROBABLE
- Rely on **statistics** instead of standard equations

A set of mathematical procedures for organizing, summarizing and interpreting information

_____ vs.


- _____
 - Def: The set of all individuals of interest
 - Ex. All teenagers in relationships
- _____
 - Def: A set of individuals selected from a population in a research study
 - Ex. 300 American teenagers in relationships

Theory of _____


- Need to be able to generalize to the population
 - Statistical analyses allow us to do this
 - Indicators in analyses tell us whether we can or not
- Must be representative of population
- Similar idea to measuring a construct
 - Sample "taps into" population accurately

Types of _____
(review)

- **Random**
- **Stratified**
 - randomly sample from groups of interest
- **Cluster**
 - sample group rather than individuals
- **Convenience**
 - sample p's who are readily available
- **Quota**
 - representative proportions




- Def: A value that describes the population
 - Ex. The average household income in the United States
 - Problem: This number is difficult to get
- Researchers obtain a **statistic** from the sample to estimate a population parameter
 - Ex. Average income of 10,000 American households from a random sample



Descriptive vs.


- **Descriptive statistics**
 - Used to simplify data
 - Ex. Average number of cigarettes smoked in a sample of EKU students
- _____ **statistics**
 - Used to make generalizations from sample to population
 - Ex. Differences in smoking between men and women at EKU → difference between men and women everywhere?




Problem with

_____ Stats

- There will always be a difference between the sample statistic and the population parameter
- Difference = _____
- Most often seen in political polls
 - Ex. "Approval rating has a sampling error of ± 3 percentage points"




A characteristic or condition that changes or has different values for different individuals



Discrete vs. _____
Variables

- **Discrete**
 - Separate categories or whole numbers
 - No "in between" numbers
- _____
 - Based on a range of values
 - Ex. Hormone levels in the sample



_____ Variables

- _____
 - Separate categories only
 - Ex. Who is still alive in the sample?
- All statistical analyses involve knowing the difference between categorical and continuous variables


Rules for _____
Variables

- Must be able to find a value BETWEEN two scores on the scale
- Rarely going to find someone who scores exactly the same as someone else
- The obtained measurement should be thought of as an interval rather than an exact point on the scale
 - Interval referred to as *real limits*

- The boundaries of scores that are represented on a continuous number line
- The real limit separating two adjacent scores is located halfway between the scores
- Upper limit at the top of the interval
- Lower limit at the bottom of the interval

Statistical _____

- Raw score = x
- Set of raw scores = x_i
 - A number next to x refers to the order in which it appears in the data set
 - So, x_1 is the first score in the set
- Sum of raw scores = $\sum x_i$
- Score multiplied by itself = x^2
- Number of people in the population = N
- Number of people in the sample = n



Order of Operations

1. Any calculation contained within parentheses
2. Raising a number by an exponent (e.g., 3^2)
3. Multiplying and/or dividing
 - A series of multiplication and/or division operations should be done in order from left to right
4. Summation using the Σ notation
5. Any other addition and/or subtraction
