INTRODUCTION TO STATISTICS

Frances Chumney, PhD

What are statistics?

UNDERSTANDING STATISTICS

UNDERSTANDING STATISTICS

Statistics is a set of mathematical procedures for organizing, summarizing, & interpreting information

- Purpose
  - Organize & understand information
  - Facilitate communication
  - Answer research questions by indicating what conclusions are justified given the data collected

group vs. subgroup

POPULATIONS VS. SAMPLES

POPULATIONS

- **Population:** Entire target group we would like to study
  - Example Research Question: *How do political views of men & women differ?*
    - Population 1: Men (all men)
    - Population 2: Women (all women)
  - Populations can be large or small
    - Example Populations
      - Men (implies *all men in existence*)
      - Adult men in the U.S.
      - Adult men in the U.S. who are registered voters
      - Adult men in the U.S. who are registered voters in the state of Georgia
  - Researcher is responsible for specifying population(s) of interest

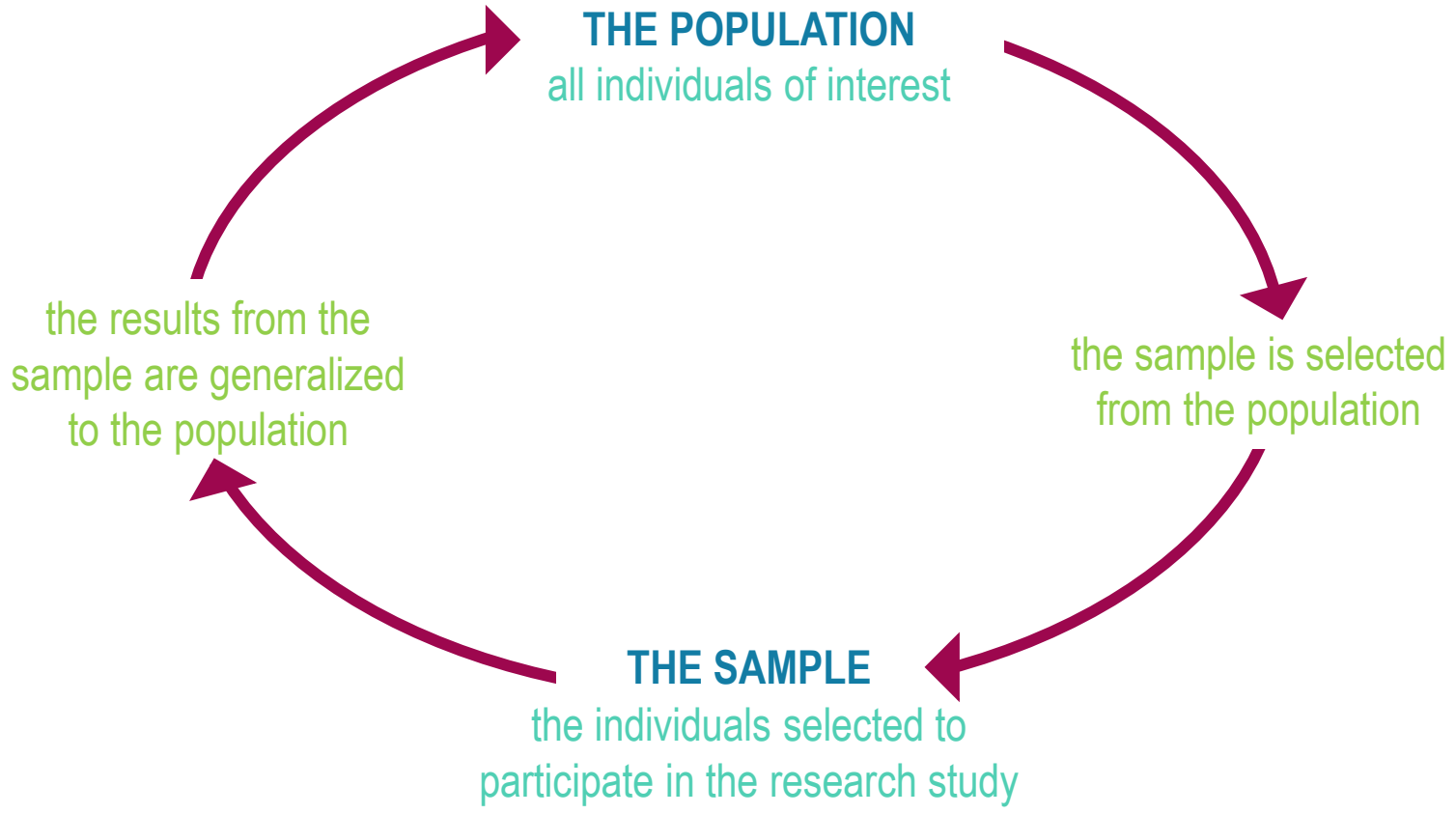
SAMPLES

- **Sample:** Subgroup or subset of units/individuals from the population meant to represent the population we would like to study
  - Motivation for Sampling
    - Populations vary in size
    - Impossible to have complete, accurate list of members (sampling frame) for some populations
    - Not feasible to collect data from all members of some populations (even with a complete, accurate sampling frame)

Statistics allow us to generalize from a sample to the population

POPULATIONS ↔ SAMPLES

statistics allow us to generalize from a sample to the population



parameter vs. statistic, descriptive vs. inferential

TYPES OF STATISTICS

PARAMETERS & STATISTICS

- **Parameter:** Value that describes a population
- **Statistic:** Value that describes a sample
- Every parameter has a corresponding statistic
  - Notation differs
  - Calculations may differ

DESCRIPTIVE & INFERENTIAL

- **Descriptive Statistics**

  statistical procedures used to summarize, organize, simplify data

- **Inferential Statistics**

  techniques that allow us to study a sample & make generalizations about the populations from which the sample was selected

- Samples = Problems
  - Provide limited information about populations
  - Representative of population ≠ microcosm of population
  - **Sampling Error**

    the discrepancy, or amount of error, that exists between a sample statistic & its corresponding population parameter

DESCRIPTIVE & INFERENTIAL

- Example: Presence of Sampling Error

Population of 1000 college students

Population Parameters
Average Age = 21.3 years
Average IQ = 112.5
65% Women, 35% Men

Sample #1

Eric, Jessica, Laura
Karen, Brian

Sample Statistics
Average Age = 19.8
Average IQ = 104.6
60% Women, 40% Men

Sample #2

Tom, Kristen, Sara
Andrew, John

Sample Statistics
Average Age = 20.4
Average IQ = 114.2
40% Women, 60% Men

standardized system for communicating mathematical functions

STATISTICAL NOTATIONS

STATISTICAL NOTATION

- Summation Notation
  - Most statistics computations involve adding (summing)
  - Σ = Summation (*sigma*) = read as "the sum of"
  - ΣX = "the sum of *X* values" or "sum of scores"
  - Example

$$X\text{: } 10, 6, 7, 4$$

    - $\Sigma X = 10+6+7+4 = 27$, $N = 4$

STATISTICAL NOTATION

- Summation Notation
  - Order of Mathematical Operations

Please Excuse My Dear Aunt Sally

    - Parentheses
    - Exponents
    - Multiply & Divide in order from left to right
    - Addition indicated using Σ
    - Other addition & subtraction, in order from left to right

STATISTICAL NOTATION

- Summation Notation Examples

X: 8, 3, 5, 1, 6

- ΣX = sum of scores

$8+3+5+1+6 = 23$

- ΣX^2 = sum of squared scores

$8^2 + 3^2 + 5^2 + 1^2 + 6^2 = 64+9+25+1+36 = 135$

- $(\Sigma X)^2$ = squared sum of scores; sum of scores, squared

$(8+3+5+1+6)^2 = (23)^2 = 529$

STATISTICAL NOTATION

- Summation Notation Examples

X: 8, 3, 5, 1, 6

- $\Sigma(X-1)$ = sum of X - 1 values

(8-1) + (3-1) + (5-1) + (1-1) + (6-1) = 7+2+4+0 +5 = 18

- $\Sigma(X\text{-}1)^2$ = sum of squared X - 1 values

$(8-1)^2 + (3-1)^2 + (5-1)^2 + (1-1)^2 + (6-1)^2 =$

$7^2 + 2^2 + 4^2 + 0^2 + 5^2 =$

49+4+16+0+25 = 94

STATISTICAL NOTATION

- Summation Notation Examples

X: 8, 3, 5, 1, 6

Y: 6, 2, 4, 3, 5

- ΣY = sum of Y values

  6+2+4+3+5 = 20

- ΣXY = sum of XY products

  (8×6) + (3×2) + (5×4) + (1×3) + (6×5) = 48+6+20+3+30= 107