Statistics for Social Sciences

- Change in gov't
- Better economics
- Technology
- Prevention programs
- Better law enforcement
  - Cameras
- Tougher penalties

Freakonomics, Steven Levitt

Roe v. Wade (1972)
- Legalization of abortion

Descriptive statistics vs. Inferential statistics

Descriptive statistics
- Describing data
  - Averages, mean, "center", variation, deviation.
  - e.g. how would you describe your classmates?

Inferential statistics
- Relationship between two variables.

Data
- Observations
- Variables

Observations - are items that are observed.
- People
- Countries
- Atoms, planets
- Nouns
- People at specific times

Variables - descriptors of the observations.
- People in class
  - Gender
  - Race
  - Education
  - Major
  - Weight*
  - Married, divorce, single
  - Parental Education
  - Income*
  - Job
  - Neighborhood
  - Height*
  - Age*

557 are male

Types of variables

- Continuous (e.g. numeric)
  - Integers, numbers, . . .
  - e.g. Age, weight, height, income, years of education
- Categorical (Discrete)
  - Categorical descriptors
  - e.g. race/ethnicity, gender, highest level of educatio
  - Binary variable
    - Either/or
    - Two options
    - Dummy variables

- Excel

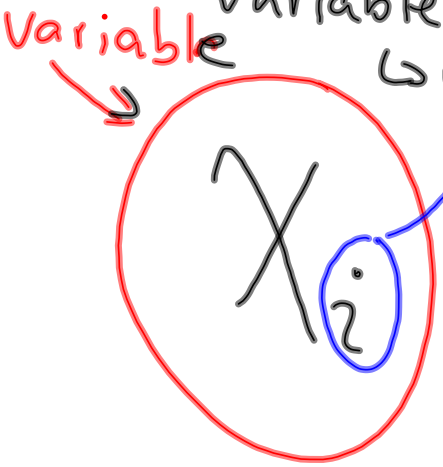
- SPSS

  - Grand View

Mathematical Notation

Data → Observation
Data → Variable

i.e., observation → Prof. Schenk
variable → age
↳ value → 27



X = age
i = observation

$X_1 = 27$

1st person = $X_1 = 27$
2nd person = $X_2 = 20$
3rd person = $X_3 = 33$

e.g. X = age
Y = income
Z = race

Average (mean)

$x_1 = 27,\ x_2 = 20,\ x_3 = 33$

$$\text{Average age} = \frac{27+20+33}{3}$$

$$= 26\tfrac{3}{3} = 26.66\overline{6}$$

$$\frac{x_1 + x_2 + x_3}{\text{number of observations}} \equiv$$

$$x_1 + x_2 + x_3 + \cdots + x_{10{,}000}$$

$$\sum \longrightarrow \sum_{i=1}^{N} x_i$$

Sigma (sum of)

N: End at last person

i=1: Start at 1st person

x_i: value of the x variable for ith observation

$$\sum_{i=1}^{3} x_i = x_1 + x_2 + x_3$$

$$\text{Average (mean)} = \sum_{i=1}^{N} \frac{x_i}{N} = \overline{x}$$

N: constant

$$\frac{x_1 + x_2 + \cdots + x_N}{N}$$

$$\text{ex: } \sum_{i=1}^{3} \frac{x_i}{N} = \frac{27+20+33}{3}$$

$$\sum_{i=1}^{N} \frac{x_i}{N} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$\text{ex: } \frac{1}{N} \sum_{i=1}^{N} x_i = \frac{1}{3}(27+20+33)$$

$$= \frac{1}{3}(80) = \frac{(80)}{3} = 26.66\overline{6}$$

$x_1 = 27, x_2 = 20, x_3 = 33$

$\sum_{i=1}^{3} x_i = 27 + 20 + 33$ Sum of x_i

$\sum_{i=1}^{3} x_i^2 = 27^2 + 20^2 + 33^2$ Sum of x_i-squared

$= 27 \times 27 + 20 \times 20 + 33 \times 33$

$\left(\sum_{i=1}^{3} x_i\right)^2 = (27 + 20 + 33)^2$

square of the sum of x_i

Descriptive Statistics

- Measures of centrality
- Measures of variability

→ Mean, median, & mode

- B average: A, B, C
- B average: B, B, B