

POLS 5377 Scope & Method of Political Science

Week 4 Introduction to Statistics

Statistics in Social Research

Healey. (2016) *Statistics: A Tool for Social Research*, Chapter 1

Key Question:

- * Why statistics for public managers?
- * What are roles of statistics in social research?
- * What are the discrete and continuous variables?
- * What are the different levels of measurement?

Outline

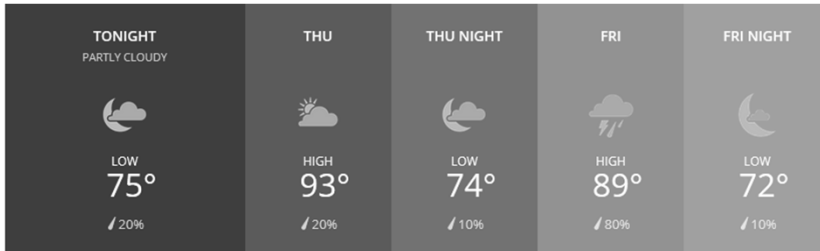
- * The use of statistics
- * Statistics in social research
- * Discrete and continuous variables
- * Level of measurement

What is Statistics?

- * Statistics is the science of **learning from data**, and of measuring, controlling, and communicating **uncertainty**; and it thereby provides the navigation essential for controlling the course of scientific and societal advances. (Davidian & Louis; American Statistical Association)
- * We go through life making choices based on incomplete information
- * Most of us live comfortably with some level of uncertainty

What is Statistics?

- * What makes statistics unique is its ability to quantify uncertainty, to make it precise.



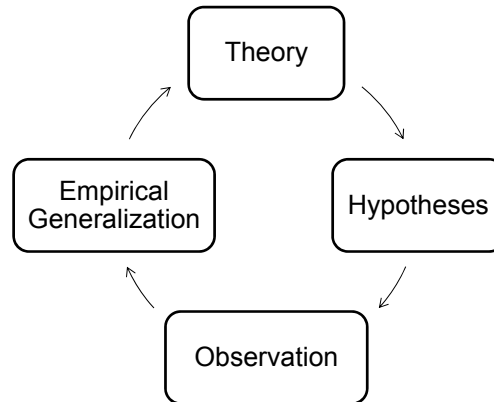
- * This allows analysts/researchers to make statements with complete assurance about their level of uncertainty.

Use of Statistics

- * As a public and nonprofit manager, there are common uses of data and analysis:
 - * To describe and analyze problems
 - * To describe policies and programs
 - * To monitor progress and prevent fraud
 - * To guide and improve program operations
 - * To evaluate outcomes
- * Data does not explain itself. Always need appropriate organization, analysis, interpretation and communication.

Role of Statistics in Social Inquiry

- Logic of reasoning (Deductive reasoning model)



Role of Statistics in Social Inquiry

- Operationalize concepts
- **Variables**: the traits that can change values from case to case
 - Example: Age, Gender, Race, Social class
 - Independent variables → Dependent variables
Cause → Effect
- **Case**: the entity from which data are gathered (unit of analysis)
 - People, Groups, States, or Countries

Role of Statistics in Social Inquiry

- Hypothesis
 - The statement about the relationship of specific variables, which is derived from theory
 - Education and income
 - People who receive higher education are more likely to have higher income levels
 - A hypothesis is not a question. Rather, it is a statement that a researcher aims to test in the study.
 - Hypotheses often include statements about a presumed impact of an independent variable (IV) on a dependent variable (DV)
 - Education (IV) → Income (DV)

Role of Statistics in Social Inquiry

- Example of using statistics
 - Variable: age
 - Cases: citizens
 - Data: the actual ages of citizens (8, 18, 30, 41, 76... etc.)
 - Statistics:
 - Average – average age of citizens in this community is 41.7 years old
 - Percentage – 15% of citizens are older than 70

Descriptive & Inferential statistics

- Descriptive statistics
 - **Univariate** descriptive: statistics describe or summarize variables one at a time
 - Percentage, averages and various charts & graphs
 - **Bivariate** and **Multivariate** descriptive: statistics describe or summarize the relationship between two or three (or more) variables
 - People with higher education level have higher income (bivariate)
 - Income levels increases with education for males but not for females (multivariate)
 - **Measure of association**: descriptive statistics designed for understand the relationship between two or more variables.

Descriptive & Inferential statistics

- Inferential statistics
 - Generalize from a sample to a population
 - Population includes **all** cases in which the research is interested
 - Samples include carefully chosen **subsets** of the population
 - Example: Presidential election polls
 - Conduct survey with a few thousands of carefully selected voters
 - This information is used to estimate the voting attention of **all** voters (millions of people)
 - The Republican candidate will receive 41% of the vote

Types of Variables

- Independent or dependent
 - Cause → Effect
 - Independent → Dependent
- Discrete or continuous
- Nominal, ordinal, or interval-ratio

Types of Variables

- Discrete variables:
 - are measured in units that cannot be subdivided (whole numbers)
 - Example: Number of children or number of cars
- Continuous variables:
 - are measured in a unit that can be subdivided infinitely (fractional numbers)
 - Example: Age or income

Level of Measurement

- Nominal
 - Scores are different from each other but cannot be treated as numbers
 - Example
 - Sex
1 = Female, 2 = Male
 - Race
1 = White, 2 = Black, 3 = Hispanic, 4 = Asian, 5 = Other
 - Political party
1 = Republican, 2 = Democrat, 3 = Other

Level of Measurement

- Ordinal
 - Scores can be ranked from high to low or from more to less
 - Survey items that measure opinions and attitudes are typically ordinal
 - To what extent do you agree or disagree that your jurisdiction has enough funding to meet its current fire protection needs?
 - A respondent that agree would be more in favor than a respondent who disagreed
 - If you can distinguish the options of the variable using terms such as more, less, higher, lower, the variable is ordinal

Level of Measurement

- Interval-ratio
 - Scores are actual numbers and have a true zero point and equal intervals between scores
 - Examples:
 - Age (in years), interval level
 - Income (in dollars) ratio level
 - Number of children, ratio level
 - A true zero point (0 = no children)
 - Equal intervals: each child adds one unit

Level of Measurement

(A)	What is your marital status? Are you presently:	(B)	Do you support the death penalty for persons convicted of homicide?																								
	<table border="1"> <thead> <tr> <th>Score</th> <th>Category</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Married</td> </tr> <tr> <td>2</td> <td>Divorced</td> </tr> <tr> <td>3</td> <td>Separated</td> </tr> <tr> <td>4</td> <td>Widowed</td> </tr> <tr> <td>5</td> <td>Single</td> </tr> </tbody> </table>	Score	Category	1	Married	2	Divorced	3	Separated	4	Widowed	5	Single		<table border="1"> <thead> <tr> <th>Score</th> <th>Category</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Strongly support</td> </tr> <tr> <td>2</td> <td>Somewhat support</td> </tr> <tr> <td>3</td> <td>Neither support nor oppose</td> </tr> <tr> <td>4</td> <td>Somewhat oppose</td> </tr> <tr> <td>5</td> <td>Strongly oppose</td> </tr> </tbody> </table>	Score	Category	1	Strongly support	2	Somewhat support	3	Neither support nor oppose	4	Somewhat oppose	5	Strongly oppose
Score	Category																										
1	Married																										
2	Divorced																										
3	Separated																										
4	Widowed																										
5	Single																										
Score	Category																										
1	Strongly support																										
2	Somewhat support																										
3	Neither support nor oppose																										
4	Somewhat oppose																										
5	Strongly oppose																										
(C)	<table border="1"> <thead> <tr> <th>Score</th> <th>Income Range</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Less than \$24,999</td> </tr> <tr> <td>2</td> <td>\$25,000 to \$49,999</td> </tr> <tr> <td>3</td> <td>\$50,000 to \$99,999</td> </tr> <tr> <td>4</td> <td>\$100,000 or more</td> </tr> </tbody> </table>	Score	Income Range	1	Less than \$24,999	2	\$25,000 to \$49,999	3	\$50,000 to \$99,999	4	\$100,000 or more		Level of measurement?														
Score	Income Range																										
1	Less than \$24,999																										
2	\$25,000 to \$49,999																										
3	\$50,000 to \$99,999																										
4	\$100,000 or more																										

Level of Measurement

- Answer:
 - (A) Nominal level
 - (B) Ordinal level
 - (C) Ordinal level

The End