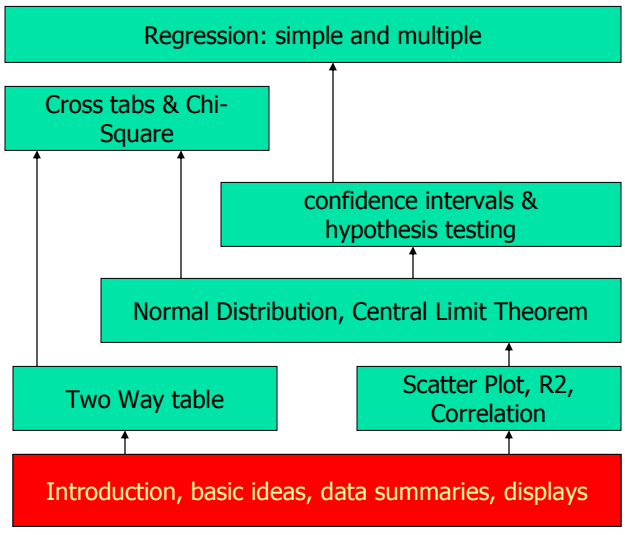
PUB – POS 316
Week 2

How do we know what we know? Data Collection

Navid Ghaffarzadegan
navidg@gmail.com
Last updated – Jan 1, 10

Course Road Map



Agenda

- 1. Introduction
- 2. Systematic methods of data gathering
  - Experiment
  - Survey
  - Interview
  - Observation
- 3. Population vs Sample
- 4. Sampling Distribution
- 5. Ethics

1. Introduction

Anecdotal evidence:

Haphazardly selected cases

Not representative

We need systematic data gathering.

1. Introduction

A simple example:

Last term John woke up at 7 AM every day. How do you find out that this influenced his grades or not?

→ We need to have two John's!

Not possible. What can we do to solve this problem?

Experiments

1. Introduction

How do you conduct an experiment for evaluation of a specific policy for homeless people?

<u>There are different methods to gather data and study data.</u>

Agenda

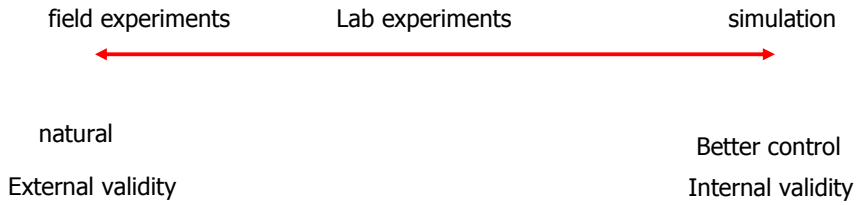
- 1. Introduction
- 2. Systematic methods of data gathering
  - Experiment
  - Survey
  - Interview, Observation
- 3. Population vs Sample
- 4. Sampling Distribution
- 5. Ethics

2. Systematic methods of data gathering

- 2.1. Experiment
- 2.2. Survey
- 2.3. Qualitative methods: Interview, Observation

2.1. Experiment

- The logic of an experiment:
  - Treatment → response



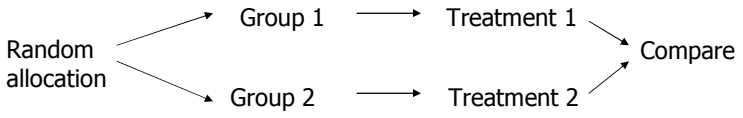
- Let's review some examples.

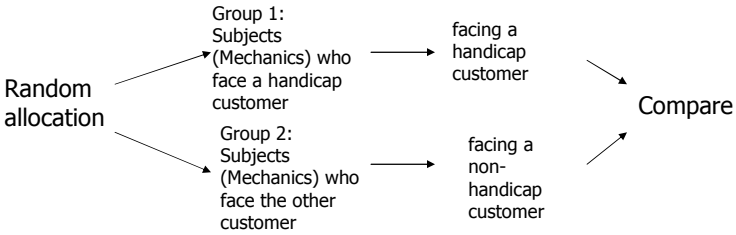
2.1. Experiment

- Simulation as experiment
  - Economic policy making
- Lab experiment
  - The Stanford Prison experiment
  - http://www.prisonexp.org/
- Field experiment
  - John List's experiment on discrimination against handicap

2.1. Experiment

- Experiment procedure.





2.1. Experiment

- Possible problems?
  - Bias
    - Use randomization
    - Control group
    - Repetition

2.2. Survey

- We can conduct an observational study.
  - No treatment. e.g, survey
- How should we design a survey?
- How to send out?
- Public management question: Compare job satisfaction between public employees and non-profit employees. (five questions)
  - http://www.surveymonkey.com/

2.3. Qualitative methods: Interview, Observation

- There are fantastic qualitative methods out there!
  - Interview
  - Observation
- All methods have their own problems.
  - Subjectivity
  - Enough evidence?
- We don't deal with qualitative methods in this course.

2. Systematic methods of data gathering

- We may use second hand data.
- Organizational reports, etc.

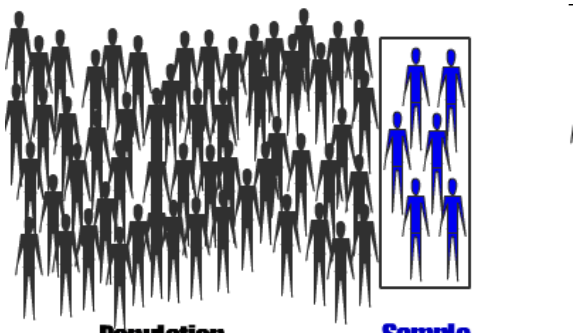
Agenda

- 1. Introduction
- 2. Systematic methods of data gathering
  - Experiment
  - Survey
  - Interview, Observation
- 3. Population vs Sample
- 4. Sampling Distribution
- 5. Ethics

3. Population vs Sample

- A sample is a part of the population that we "actually" examine to say something about the population.
  - Reasons:
    - It is expensive to study the whole population
    - It is impossible.
    - It is time consuming.
    - It doesn't matter that much!
    - ….

3. Population vs Sample



Population	Sample
quantity (count) = N	quantity (count) = n
mean = μ	mean = $\bar{x}$
variance = σ^2	variance = s^2
standard deviation = σ	standard deviation = s

3. Population vs Sample

- We want to know these information about UAlbany students:
  - How many hours a day do they study?
  - Do they prefer to work for the government or the private sector?
  - How long do they spend in their way from the school to their apartment?
- Let's use our class as a sample of the university.
- Voluntary.

3. Population vs Sample

- We should differentiate population from a sample. [seems easy, but it is usually a problem!]
- How should we sample?
  - Voluntary sampling
    - Problem: biased
    - How can we solve this issue?

3. Population vs Sample

- Simple Random Sample (SRS)
  - Consists of *n* individuals from the population chosen in such a way that every set of n individuals has an equal chance to be the sample actually selected.
    - In a simple word: it is purely random!

3. Population vs Sample

- More careful way of sampling: divide the population into groups of similar individuals, and make sure you are choosing your sample from each group. [stratified random sample]
  - Example: Is the university safe?
    - pick 50 male, and 50 female students for your survey.

Agenda

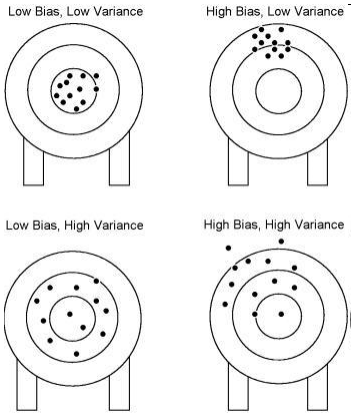
- 1. Introduction
- 2. Systematic methods of data gathering
  - Experiment
  - Survey
  - Interview, Observation
- 3. Population vs Sample
- 4. Sampling Distribution
- 5. Ethics

4. Sampling Distribution

- We want to use sample statistics to make statements about unknown population
- Problems with sampling

4. Sampling Distribution

- Sampling bias, variability
- In shooting arrows!



Agenda

- 1. Introduction
- 2. Systematic methods of data gathering
  - Experiment
  - Survey
  - Interview, Observation
- 3. Population vs Sample
- 4. Sampling Distribution
- 5. Ethics

5. Ethics

- What kinds of ethical concerns we may have?
- Institutional review board.
- Informed consent.
  - Example: Relation between ones personality and the place she sits!
- Confidentiality
  - Example: Course evaluation

Summery

- We finished chapter 3 – What I expect:

- I expect you to know the basics of data collection.
- You should know
  - about designing an experiment.
  - about conducting surveys.
  - about how to differentiate Population from Sample
  - the basics of finding a good random sample
  - about Sampling Distribution
  - ethics of conducting research

- Submit your assignment on time.
- Visit me at the office hours.
- Next we will talk about how to use data.