Research Design

Beyond Randomized Control Trials

Jody Worley, Ph.D.
College of Arts & Sciences
Human Relations
Introduction to the series

• Day 1: Nonrandomized Designs
• Day 2: Sampling Strategies
• Day 3: Matching Techniques for Balanced Designs
WHAT CAN HAPPEN IF YOU USE AN INCORRECT SAMPLE?

67% of your friends would like to eat pizza tonight.
Samples and Sampling Strategies

• A sample is the next best thing to collecting data from the entire target population.

• Paradox of Sampling

• Resolve the paradox via incrementalism and obtaining the best sample possible.
I just read your column in today's paper. How can you say that the number of functionally illiterate in our society is greatly exaggerated?!

I did a survey, and of all the people who completed and mailed back the questionnaire...

...not one was illiterate.
Nonprobability vs. Probability Sampling

- Nonprobability—Sample is selected using some nonrandom process (not based on chance).

- Probability—Sample is selected based on a random process.
QUESTIONS TO ASK: MAKING SENSE OF NUMBERS

SAMPLING: DO I HAVE TO COUNT EVERYTHING?

Universe (theoretical target population)

Population (empirical target population)

Original sample

Final sample (data)

sometimes called sampling frame

Loss (non-response)
Who do you want to generalize to?

What population can you get access to?

How can you get access to them?

Who is in your study?

The Theoretical Population

The Study Population

The Sampling Frame

The Sample
Nonprobability Sampling

Obtain subjects in some nonrandom way.

Typically volunteers
Types of Nonprobability Samples

- Stratified Disproportionate/Stratified Quota
- Purposive or Judgmental
- Accidental/Use of Available Subjects
- Snowball
Quota Sampling

• Divide the population into various categories

• Determine the number of people to be selected for each category

• For each category, you obtain an accidental sample until you fill the quota of people you need for that category

• Assign a weight to all respondents in a given cell according to their proportion of the total population.
Quota Sampling - Example

Population
1500 elderly living in a residential setting

Quota Sample
100 females
50 males
Purposive or Judgmental Sample

• Select your sample on the basis of your knowledge of the population, its elements and the purpose of the study.

• Community needs assessment
  • Pastor of Church
  • Local chapter leader of AARP
  • Older person who has been living community for years.
Purposive/Judgmental Sample

• Why some individuals who experience severe poverty and deprivation in childhood are successful as adults and others not.

• Why some women who experience a trauma such as a rape are devastated by the event and others are able to rebound.
Accidental/Convenience/Availability Sample

• Select people because they are readily available.

• Sample members chosen because convenient. Easy to access
Examples of Accidental Samples

• Place ad in newspaper. Those who call to volunteer to participate.

• Working at a local senior center so sample older people from Center

• People who belong to an organization
Random Sampling from a Convenient List

Study relationship between experiences of abuse and psychological well-being.

Take a random sample from Mental Health Center of Dane County
Snowball Sample

• Find someone who fits the criteria for the study (accidental/convenience sample)

• Interview person and at end of interview, you ask if he/she knows of other people who meet the study criteria and may be willing to participate
Probability Sample

Enhances likelihood of getting a representative sample.

Calculate sample error and confidence interval
Obama 52% of the vote with a +/- 3% Margin of Error
(with a 95% confidence)

If able to poll everyone, you are 95% confidence that you would find that somewhere between 49% and 55% of the population favored Obama.
So, Why do we say they are equally favored?

• Trump 49% to 55%

• Clinton 45% to 51%

October 6: Clinton 51%  Trump 49%

http://projects.fivethirtyeight.com/2016-election-forecast/iowa/
Types of Probability Samples

- Simple Random Sample
- Systematic Random Sample
- Stratified Random Sample
Simple Random Sample

Each element in the population has an equal probability (chance) of being selected for the sample.
Systematic Random Sample

• Every $k^{\text{th}}$ element in the list is chosen for inclusion in the sample

• $K$ is called the sampling interval
  
  $K = \frac{\text{population size}}{\text{sample size}}$

  population has 200 people. And you have resources to sample 50 people. $K = \frac{200}{50} = 4$

Sample every $4^{\text{th}}$ person on the list. Start at random location
Stratified Random, or Stratified Proportionate
(a. k. a. Full Probability)

• Requirements
  – Must know characteristics about the population
  – Must be mutually exclusive and homogenous subparts

• Purpose
  – Increase the probability of obtaining an accurate sample by randomly selecting participants from each significant element of the population
  – Increase the accuracy of the sample
Stratified Random Sample

- Population is first divided into two or more strata.
  Strata – is a group of people who share a common characteristic.
  Examples of strata – race, gender, marital status.

- From each stratum, a simple random sample is taken.

- The subsamples (samples from each stratum) are then joined to form the total sample.
Stratified Random Sample

<table>
<thead>
<tr>
<th>Variables</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>Min./majority</td>
</tr>
<tr>
<td>Religion</td>
<td>Cath./Prot./Jew</td>
</tr>
<tr>
<td>Education</td>
<td>HS/Col./Grad-Prof.</td>
</tr>
<tr>
<td>Income</td>
<td>Low/Mod/High</td>
</tr>
<tr>
<td>Age</td>
<td>Young / Old</td>
</tr>
</tbody>
</table>

The researcher must decide in advance which variables are to be stratified.

Know information about distribution of population according to strata.

Keep the number of stratification variables (strata) to a minimum because otherwise the sample size requirements to achieve a representative sample becomes unwieldy.

Minimum strata is 10 cases.
Stratified Random Sample

Conditions for Stratified Random Sample

• The researcher must decide in advance which variables are to be stratified.

• Know information about distribution of population according to strata.

• Keep the number of stratification variables (strata) to a minimum because otherwise the sample size requirements to achieve a representative sample becomes unwieldy.

• Minimum strata is 10 cases.

• To determine the number of strata, multiply the number of attributes, and then multiply that sum by 10 to get total number of strata.
## Stratified Random Sample

To determine the number of strata, multiply the number of attributes, and then multiply that sum by 10 to get total number of strata.

**STEPS**
1. Multiply # attributes to get # of strata \((2 \times 3 \times 3 \times 3 \times 2) = 108\)

2. Strata \(\times 10 = \# \) needed for sample \(= (108 \times 10) = 1080\)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Attributes</th>
<th># of Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>Min./majority</td>
<td>2</td>
</tr>
<tr>
<td>Religion</td>
<td>Cath./Prot./Jew</td>
<td>3</td>
</tr>
<tr>
<td>Education</td>
<td>HS/Col./Grad-Prof.</td>
<td>3</td>
</tr>
<tr>
<td>Income</td>
<td>Low/Mod/High</td>
<td>3</td>
</tr>
<tr>
<td>Age</td>
<td>Young / Old</td>
<td>2</td>
</tr>
</tbody>
</table>
EXAMPLE: STRATIFIED RANDOM SAMPLE

Population: 70,000,000 U.S. voters in the Presidential election
Sample Size: 2,000
Sample Size as a proportion of the population: 2000/70,000,000 = .0000285 (sampling rate) = sample as a proportion of the population
Theoretically significant variables and their attributes:
1. Sex: Male, female
2. Age: Young (18-30 yr.); Middle Age (31-50 yr.); Aged (over 60 yr.)
3. Income: Poverty (< $15,000); Nonpoverty ($15,000 and over)

STEPS:
1. Name the Strata
2. Count the number in the population found in each strata
3. Calculate the proportion each stratum is of the population

Sex: Male
   Age: Young  Middle  Aged
   Income: Poor  Rich  Poor  Rich  Poor  Rich  Poor  Rich
   Number (in millions) in population, by stratum
   4.2  5.6  10.5  7  5.6  1.4  2.1  6.3  7 10.5  5.3  6.3

Stratum as a percent of population
   6  8  15  10  8  2  3  9  10 15  5  9

STRATIFIED RANDOM SAMPLE
Stratum as a percent of Sample Size 2,000
   6  8  15  10  8  2  3  9  10 15  5  9

Number in each Stratum of Sample:
Technique I:
   Stratum n * % sample is of population
   120 160 299 200 160 40 60 160 200 299 100 100

Technique II:
   n * % Stratum is of population
   120 160 300 200 160 40 60 160 200 300 100 100

QUOTA Sample:
Stratum as percent of Sample Size**
   6  8  12  10  8  5  5  9  10 13  5  9

Number in each Stratum of Sample***
   120 160 240 200 160 100 100 160 200 260 100 100

* These strata are judged more significant than others, thus these quotas are fixed.
** Compare these percentages with those percentages for "Stratified Random Sample"
*** Compare these numbers with those numbers for "Stratified Random Sample"