Week 3: Sampling and Surveys

Jennifer Lin

2021-10-06

So you want to Conduct a Survey?

What you need

To run a good survey, you need some of the following

1. A **RANDOM** Sample
2. Good Survey Questions
3. Skills to analyze survey data

Today, we will focus on the **Random** Sample component

What is Random? What is a Sample?

First, what is a **sample**? To figure that out, we need to know what is part of the **population**, or the entire entity of the thing that we are interested in studying. Individuals or single unites in the population constitute our **unit of analysis**. This can be people (a person), states (a state), or households (house).

When designing a study, your **sample** is a selection of the units in the population and it should resemble the entire group of interest. There are many ways to pick a sample and there are positives and negatives of each method

1. **Sampling of Convenience**: Select cases that are easily accessible
    a. EXAMPLE: Stand outside of dining hall and interview everyone trying to go in
    b. EXAMPLE: Polls on FOX, CNN, etc.
2. **Quota Samples** are where you modify the convenience sample to make it more like the population
    a. This is also known as **non-probability sampling**
    b. This is often used in online survey marketplaces and Amazon Mechanical Turk
    c. EXAMPLE: The COVID States Project: `https://covidstates.org/`
        - We want to know political attitudes and behaviors of people in each US State
        - We conduct online samples to get respondents from each US State
        - Each state has a quota of 300-500 per wave, depending on size

- Within each state, we do race, gender and age cross tabs so that each bucket gets a fix number equivalent to their percent in the state.

3. **Random Sampling** is where each case has an equal and known probability of being selected
   a. Also known as **Probability Sampling**
   b. Often done by Random Digit Dialing
   c. EXAMPLE: Kaiser Family Foundation (Random Digit Dialing)
   d. **Law of Large Numbers**: Random Sampling gives us a good approximation of the population with a large enough population
   e. Types of Random Samples:
      1. **Simple Random Sample**: Each person has the same probability of being included
      2. **Cluster Sample**: Group participants into meaningful subgroups and select the subgroup to its entirety
      3. **Stratified Random Sampling**: Group participants into meaningful subgroups and select samples from the subgroup where EACH member of the subgroup has the same probability of selection

Issues with Surveys

Surveys may have multiple issues related to measurements. For the sampling front, however, we are interested in issues with the sample recruitment. Some issues here might include

1. **Non-response**: People selected for the survey just don't want to engage with the survey
2. **Hard to Reach Populations**: Some people who are selected for surveys might not be able to access the survey OR it can be hard to reach some people since they are difficult to identify.

There are some ways to address these issues. But to do this, we need to be creative with our sampling methods.

1. **Snowball Sampling**: Getting potential participants from current participants
2. **Respondent-driven Sampling**: Go where the participants are (or might be)