STA304/1003 H1 F - Summer 2014: Surveys, Sampling, and Observational Data

Lecture 2 - Part I: Questionnaire Design

Ramya Thinniyam

May 20, 2014

REVIEW - Why Bother Sampling?

- Surveys can be done very poorly ...
- People are skeptical about survey results - “My opinion was not asked so how can the results represent me?”

A census is not always better - may not be practical, expensive, time-consuming, may not eliminate error.

Advantages of Sampling

1. Provides reliable information at less cost:

   - Can quantify sampling error when using probability samples
   - Does not destroy population in cases where element must be destroyed to be measured (ex. cookie fat content, chemical content, flame ionization detection)

2. Faster data collection:

   - Results can be published in a reasonable time frame so they are useful

3. Estimates often more accurate than those based on census:

   - Census requires lots of organization, many people - prone to errors if results need to be produced in timely fashion
   - Samples can give more attention to data quality by training investigators, following up on non-respondents, etc.
   - Better to get good measurements on a (representative) sample than unreliable ones on the whole population

‘Sampling’ is not just taking a subset of the population.

“Sampling is the science and art of controlling and measuring the reliability of useful statistical information through the theory of probability”.

GUIDELINES for Constructing a Good Questionnaire

- Decide what you want to find out:
  Specific goals/parameters to be estimated, write questions that will get accurate answers to questions of interest and that will get responses
- Test questions before actual survey:
  Find mistakes, ways to reword/clarify questions based on feedback, test on different members of the target population
- Clear and simple questions:
  Words that are simple to you may be misinterpreted. Ex. “proportion”.
- Specific questions, not general
- Relate questions to the concept of interest

Types of Questions

- Open vs. Closed-Ended questions:
  Open questions: respondent is not given categories, they answer freely.

  Obtain their actual opinions, can get answers investigator hadn't thought of

  Closed questions: respondent must choose from categories/multiple choice questions.

  Prompts ideas that respondent may have forgotten, usually more accurate responses, investigator can give clear questions, always have "Other" category, all options should be mutually exclusive, be aware of "Don't know" option - people can be lazy but "No opinion" is okay.
- Report the actual question asked when stating results:
  Do not reword question
- Avoid leading/loaded questions
  Leading/Loaded questions: questions that prompt respondent to say what investigator wishes to hear

- Use forced-choice rather than agree/disagree questions:
  People tend to agree, people tend to choose middle of "Likert scales"
- Ask one concept per question
  Double-barreled questions: question based one more than one issue but only allows one answer.

  Disagreeing to one opinion/concept leads to disagree on entire question, cannot distinguish which one
- Be careful with wording and sentence structure:
  Avoid complex sentence structure, order of questions matter (usually better to state general then specific question), order of words matter, avoid double negatives, use gender neutral terms
- Avoid having answer to one question affect answer to another
- Do not shorten questions that refer to previous ones:
  Display full reference
- Check respondent's consistency by using similar questions widely spaced

Planning a Survey:

1. State Objectives
2. Define Target Population
3. Find Sampling Frame
4. Choose Sample Design
5. Determine Method of Measurement
   (phone, mail, in-person, online, etc.)
6. Obtain or Construct Measurement Instrument
7. Train Field Workers
8. Conduct a Pre-test
9. Organize Field-work
10. Organize Data
11. Analyze Data

Example: Bad Surveys

Identify why each of the following survey questions are ‘bad’:

(1) Moore (2000) reports an uncited survey about assisting the poor: Only 13% of American’s think too much is spent on assistance to the poor. However, 44% believe too much is spent on welfare.
Why is there a discrepancy?

1. Change wording, define specific term for ex. 'poor' assistance' means?

(2) Have you taken any steps to alleviate water problems around the world?

2. Leading question survey's opinion is clear, more specific wording need

(3) An internet survey in Orange County, California, USA, surveyed 437 potential homebuyers.
One of the conclusions - All but 2% of the buyers have at least one computer at home, and 62 percent have two or more. Of those with a computer, 99 percent are connected to the internet.

3. Survey conducted on Internet -results are obvious.

Example: Bad Surveys (con'd...)

(4) Do you feel that eliminating taxes by law is an effective way to stop the government from picking your pocket every day?

4. Loaded/leading question

(5) Do you agree or disagree with this statement: A chemical used in food production that has a negligible cancer risk should be prohibited even though it delays spoilage, prevents rancidity, or prolongs storage time?

5. Confusing wording, leading question vague 'negligible'

(6) Did you vote in the federal election?

6. Too general, better question 'did you vote for a prime minister in the election that took place in May 2011?'

(7) How much water do you drink everyday?

7. Units of measurement to be defined. Define 'water'

(8) Do you favour the use of capital punishment?

8. Loaded/leading. Better: 'do you favor/oppose/neutral the use of capital punishment?'

Example: Good Survey - HOMEWORK ASSIGNMENT

Search the internet (or some other source) for an example of what you think is a good survey. Describe what makes it a good survey and propose further improvements.

COURSE PROJECT - Instructions

Group project - 5/6 students per group. Graduate students (STA1003) may work individually instead of in a group.

Aim: Conduct a real survey to answer questions of interest. The project requires you to:

- Choose a target population and question(s) of interest
- Create a questionnaire (mini-questionnaire 3-5 questions, but you may have more questions)
- Collect data from a sample
- Analyze the data and report statistics
- Make conclusions/inferences based on the data
- Discuss findings and comment on methods

Parts of the Project

1. Proposal

- Define your target population (it cannot be the STA304 class) and question(s) of interest
- Describe what you will estimate
- Sampling frame OR Explain why you are not using a sampling frame
- Propose sample size
- Name the sampling method (SRS, STRS, Systematic, Cluster, etc). If necessary, describe clusters and/or strata you will use
- Outline data collection methodology - cannot use pre-existing data, collect data yourself/train field workers to collect data
- Describe how variables will be measured. If you are using a questionnaire, provide a copy of it

2. Written Report

Formal written report of project (5 pages maximum) - word-processed, appropriately titled, headings/sections, tables/figures, etc.

Include all of the following in your report:

- Description of your goal/questions of interest, target population, sampling method, and data collection methodology.
- Questionnaire and statistics/summary of data. The mean, proportion, or total that you estimated. Include your sample size and provide appropriate confidence interval(s). Interpret the results in practical terms.
- Sources of bias in your data.
- Any other relevant information
- Limitations and Possible Improvements
- Overall conclusion

- In addition to the 5 page report, please include:
  - Cover sheet signed by all group members
  - Proposal (that was submitted before)
  - An appendix containing your questionnaire, raw data, code, and output (code should be commented)

DEADLINES

- Student Groups - Thur May 22nd (by email to TA)
- Proposal - Tue June 3rd at 6:10pm (in lecture)
- Written Report - Thur June 19th at 6:10pm (in lecture)

*** Read Course Project Instructions posted on website for full details and follow instructions ***