Observational Studies, Sampling Strategies, and Experiments Applications Solutions

Lt Col Kris Pruitt Professor Bradley Warner

25 August, 2020

Exercises

1. **Propose a sampling strategy**. A large college class has 160 students. All 160 students attend the lectures together, but the students are divided into 4 groups, each of 40 students, for lab sections administered by different teaching assistants. The professor wants to conduct a survey about how satisfied the students are with the course, and he believes that the lab section a student is in might affect the student's overall satisfaction with the course.

a. What type of study is this? Observational study.
b. Suggest a sampling strategy for carrying out this study. Stratified sample, sample randomly within each section.

2. **Flawed reasoning**. Identify the flaw in reasoning in the following scenarios. Explain what the individuals in the study should have done differently if they wanted to make such strong conclusions.

a. Students at an elementary school are given a questionnaire that they are required to return after their parents have completed it. One of the questions asked is, *Do you find that your work schedule makes it difficult for you to spend time with your kids after school?* Of the parents who replied, 85% said *no*. Based on these results, the school officials conclude that a great majority of the parents have no difficulty spending time with their kids after school.

Solution
Non-responders may have a different response to this question. The parents who returned the surveys are probably those who do not have difficulty spending time with their kids after school. Parents who work might not have returned the surveys since they probably have a busier schedule.

b. A survey is conducted on a simple random sample of 1,000 women who recently gave birth, asking them about whether or not they smoked during pregnancy. A follow-up survey asking if the children have respiratory problems is conducted 3 years later, however, only 567 of these women are reached at the same address. The researcher reports that these 567 women are representative of all mothers.

Solution
It is unlikely that the women who were reached at the same address 3 years later are a random sample. These missing responders are probably renters (as opposed to homeowners) which means that they might be in a lower socio-economic status than the respondents.

3. **Sampling strategies**. A Math 377 student who is curious about the relationship between the amount of time students spend on social networking sites and their performance at school decides to conduct a survey. Four research strategies for collecting data are described below. In each, name the sampling method proposed and any bias you might expect.

a. He randomly samples 40 students from the study's population, gives them the survey, asks them to fill it out and bring it back the next day.

b. He gives out the survey only to his friends, and makes sure each one of them fills out the survey.

c. He posts a link to an online survey on his Facebook wall and asks his friends to fill out the survey.

d. He stands outside the QRC and asks every third person that walks out the door to fill out the survey.

Solution
a. Simple random sample. Non-response bias, if only those people who have strong opinions about the survey responds his sample may not be representative of the population.
b. Convenience sample. Under coverage bias, his sample may not be representative of the population since it consists only of his friends. It is also possible that the study will have non-response bias if some choose to not bring back the survey.
c. Convenience sample. This will have a similar issues to handing out surveys to friends.
d. Convenience sample. Same.

4. **Vitamin supplements**. In order to assess the effectiveness of taking large doses of vitamin C in reducing the duration of the common cold, researchers recruited 400 healthy volunteers from staff and students at a university. A quarter of the patients were assigned a placebo, and the rest were evenly divided between 1g Vitamin C, 3g Vitamin C, or 3g Vitamin C plus additives to be taken at onset of a cold for the following two days. All tablets had identical appearance and packaging. The nurses who handed the prescribed pills to the patients knew which patient received which treatment, but the researchers assessing the patients when they were sick did not. No significant differences were observed in any measure of cold duration or severity between the four medication groups, and the placebo group had the shortest duration of symptoms.

a. Was this an experiment or an observational study? Why?

b. What are the explanatory and response variables in this study?

c. Were the patients blinded to their treatment?

d. Was this study double-blind?

e. Participants are ultimately able to choose whether or not to use the pills prescribed to them. We might expect that not all of them will adhere and take their pills. Does this introduce a confounding variable to the study? Explain your reasoning.

Solution
a. Experiment, since the researchers randomly assigned different treatments to the participants.
b. Response variable: Duration of the cold.
Explanatory variable: Treatment, with 4 levels; placebo, 1g, 3g, 3g with additives.
c. The patients were blinded as they did not know which treatment they received.
d. The study was double-blind with respect to the researchers evaluating the patients, but the nurses who briely interacted with patients during the distribution of the medication were not blinded. (It was partially double-blind.)
e. Since the patients were randomly assigned to the treatment groups and they are blinded we would expect about an equal number of patients in each group to not adhere to the treatment. While this means that final results of the study will be based on fewer number of participants, non-adherence does not introduce a confounding variable to the study.

5. **Exercise and mental health**. A researcher is interested in the effects of exercise on mental health and she proposes the following study: Use stratified random sampling to ensure representative proportions of 18-30, 31-40 and 41-55 year olds from the population. Next, randomly assign half the subjects from each age group to exercise twice a week, and instruct the rest not to exercise. Conduct a mental health exam at the beginning and at the end of the study, and compare the results.

a. What type of study is this?

b. What are the treatment and control groups in this study?

c. Does this study make use of blocking? If so, what is the blocking variable?

d. Does this study make use of blinding?

e. Comment on whether or not the results of the study can be used to establish a causal relationship between exercise and mental health, and indicate whether or not the conclusions can be generalized to the population at large.

f. Suppose you are given the task of determining if this proposed study should get funding. Would you have any reservations about the study proposal?

Solution
a. This is an experiment since we assigned subjects to the exercise program.
b. The treatment is exercise twice a week and control is no exercise.
c, Yes, the blocking variable is age.
d. No, the study is not blinded since the patients will know whether or not they are exercising.
e. Since this is an experiment, we can make a causal statement. Since the sample is random, the causal statement can be generalized to the population at large. However, we should be cautious about making a causal statement because of a possible placebo effect.
f. It would be very difficult, if not impossible, to successfully conduct this study since randomly sampled people cannot be required to participate in a clinical trial.

File Creation Information

- File creation date: 2020-08-25
- R version 3.6.3 (2020-02-29)
- `mosaic` package version: 1.7.0
- `tidyverse` package version: 1.3.0