ADMS2320.com

We Make Stats Easy.

Chapter 5

Tutorial Length
20 Minutes

Data Collection

How To Obtain Data

-Direct Observation
    We can only observe and record data

-Experiments
    We can manipulate samples to determine a cause and effect relationship

-Surveys
    -Personal Interviews
    -Telephone Interviews
    -Self-administered

Questionnaire Guidelines For Surveys

1. Keep overall questionnaire short
2. Keep questions short and worded with simple, clear language
3. Start with simple questions regarding demographics (age, gender, etc)
4. Yes/No questions are good and bad, same with multiple choice. Depends on the level of detail you want.
5. Open ended questions also good or bad. It takes more time to analyze answers, but allows for more detailed info to be obtained.
6. Don't use leading questions.
7. Try to pretest the questionnaire to uncover issues.
8. Keep in mind how you'll want to analyze the data when preparing the questionnaire.

Sampling

When trying to collect data for a given population, make sure the sample chosen to collect data from actually represents the population. In other words, the sampled population must represent the target population.

For example, sampling only wealthy people will result in bias and does not truly represent the population if you're trying to figure out who the next president will be.

Voluntary response surveys aren't usually good because only people that care deeply will respond. This is a type of "self-selected sample". Another example is the people calling in to talk radio shows to express their answer to a question by the host.

Simple Random Sample
Each sample is selected in a way that each sample has the same chance of being selected. This type of sample allows you to infer data about the population.

Stratified Random Sample
Separate the population into groups before choosing a simple random sample. In stratified random samples, each sample can belong to only ONE group (the groups are mutually exclusive). For example, separate by gender, age, religion etc. This allows a researcher to collect information on the population as well as see if there is differences based on the group they belong to.

Cluster Sample
A simple random samples of groups/clusters.
I.e. in trying to find average income of a city, we can choose different city blocks to survey.
This type of sample used when you don't have a full list of population members, or the population is widely distributed geographically, or a Simple Random Sample is not affordable.

Sampling Error

The difference between a sample statistic and a population parameter due to the observations of the selected sample.

Ex. If you want the height of all students at your university, but somehow the simple random sample consists of only the tallest students completely by chance. The sample statistics would be higher than the actual population parameter.

Nonsampling Error

Data acquisition errors – Incorrect recording of responses, or the sample giving incorrect responses.

Non response error – Bias obtained when responses are not obtained from all samples.

Selection bias – a sampling plan that prevents some members of the population from being selected.

Multiple Choice

1. A researcher counted and recorded the number of students that wore t-shirts in a given day. This type of data collection is called:

a. Direct observation.
b. An experiment.
c. A survey.
d. None of these choices.

2. When designing a good survey, which of the following are false?

a. Avoid leading questions.
b. Keep the questions as short as possible.
c. Dichotomous, multiple-choice, and open-ended questions may be used.
d. All of these choices are true.

3. A researcher selects a random sample of students and asks a list of question pertaining to food preferences. The type of data collection being used is:

a. Direct observation.
b. An experiment.
c. A survey.
d. None of these choices.

4. In order to create a good survey, a researcher should avoid which of the following?

a. Dichotomous questions.
b. Demographic questions.
c. Leading questions.
d. All of these should be avoided.

5. A study was conducted in which a researcher divided subjects into three groups. Each group was given a specific treatment and the response to the treatment was recorded. What data collection technique was used?

a. A census.
b. An experiment.
c. A survey.
d. Direct observation.

6. A math professor studied how students would calculate the result of an equation. The professor observed and recorded whether a student used a calculator, worked out the equation on paper, or simply waited to copy the answer from the blackboard. What data collection technique was used?

a. A survey.
b. Direct observation.
c. An experiment.
d. A personal interview.

7. If every possible sample is equally likely to be chosen, the type of sample is called a:

a. stratified random sample.
b. simple random sample.
c. cluster sample.
d. None of the above.

8. An electronics blog writer wants to determine how satisfied people are with their cell phones. He wishes to focus on four brands: Apple, Samsung, HTC, and Blackberry. What would be the best sampling strategy to use?

a. stratified random sample.
b. simple random sample.
c. self-selected sample.
d. cluster sample.

9. Sampling error is caused by which of the following?

a. Studying a random sample from a population instead of the entire population.
b. Error in data collection
c. Nonresponse bias.
d. All of these choices are true.

10. A researcher has selected a simple random sample from two groups, males and females. He wishes to determine how much time they spend studying for a stats course. What type of sampling was used?

a. Simple random sampling.
b. Cluster sampling.
c. Stratified random sampling.
d. None of these choices.

11. A simple random sample is drawn from a set of mutually exclusive groups that the population is divided into. This type of sampling is referred to as:

a. Selection bias.
b. Cluster sampling.
c. Simple random sampling.
d. Stratified random sampling.

12. _____ are usually biased.

a. Stratified random samples.
b. Cluster samples.
c. Self-selected samples.
d. None of the above.

13. The difference between a population and sample mean is known as a:

a. selection bias.
b. nonresponse error.
c. sampling error.
d. nonsampling error.

14. Nonsampling error can be caused by which of the following?

a. Some responses are incorrectly recorded.
b. All members of the sample did not respond fully.
c. Not all members of the target population can be chosen for the sample.
d. All of these choices are true.

15. A professor contacted 400 students selected at random to conduct a survey. 100 students were selected from first year, 100 from second year, 100 from third year, and 100 fourth years. What type of sampling was performed?

a. Cluster sample
b. Simple random sample
c. Stratified random sample
d. Systematic sample

16. Cluster samples are useful for which of the following scenarios?

a. Simple random samples are too costly.
b. Getting a hold of a complete list of population members is difficult.
c. The population members are widely dispersed.
d. All of these choices are true.

17. In order to award a prize at a charity dinner, each guests name is put in to a bowl and one name is selected at random. This type of sampling is known as:

a. Simple random sample
b. Stratified random sample
c. Systematic sample
d. None of the above