14–1 Common Sampling Techniques READ pp. 718-728 and fill in notes

Population – all subjects (human or otherwise) under study

Sample – a subgroup of the population

What must the sample be? for researchers to make valid inferences about population characteristics, the sample must be **random**.

Random sample – a sample in which **every member** of the population must have an **equal chance of being selected**

Unbiased sample – a sample is chosen at random from a population

Biased sample – a sample chosen in which some type of systematic error has been made in the selection of the subjects

Reasons for using a sample to get information about a population:

1. It saves the researcher time and money.
2. It enables the researcher to get information that he or she might not be able to obtain otherwise.
3. It enables the researcher to get more detailed information about a particular subject.

To obtain unbiased samples, statisticians have developed several basic sampling methods, list 4.

1. random
2. systematic
3. stratified
4. cluster sampling

Random Sampling

What is the basic requirement for *random sampling*?

For a sample of size *n,* all possible samples of this size have an equal chance of being selected from the population.

List 3 incorrect methods commonly used.

1. **"the person on the street"** Many people will be at home or at work when the interview is being conducted and therefore do not have a chance of being selected.

2. **Ask a question by either radio or television and have the listeners or viewers call the station to give their responses or opinions.** Smaple is not random, since only those who feel strongly for or against the issue may respond and people may not have heard or seen the program.

3. **Ask people to respond by mail.** Only those who are concerned and who have the time are likely to respond.

Work through the examples, you will have an assignment where you will need to complete a similar exercise.

Systematic Sampling

systematic sample – a sample obtained by **numbering each element** in the population and then **selecting every** third or fifth or tenth, **nth,** etc., number from the population to be included in the sample. This is done after the first number is selected at random.

What is an advantage of systematic sampling? the ease of selecting the sample elements

What must you be careful of when using a systematic sampling method? You must be careful how the items are arranged on the list. Example: If I arrange students Senior, Junior Sophomore, and pick every third student, I will have a group of just Sophomores.

Stratified Sampling

Stratified Sampling – a sample obtained by dividing the population into subgroups, called ***strata***, according to various homogeneous characteristics and then selecting members from each stratum for the sample.

Example: A researcher wants to test previous studies that link higher academic performance with playing a musical instrument. A high school population consists of male and female students who play or do not play a musical instrument. The researcher will want to include in the sample

people from each group—males who play a musical instrument, males who do not play a musical instrument, females who play a musical instrument, and females who do not play a musical instrument. To accomplish this selection, the researcher divides the population into four subgroups and then selects a random sample from each subgroup. This method ensures that the sample is representative on the basis of the characteristics of gender and playing a musical instrument.

What is an advantage of stratified sampling?

It ensures representation of all population subgroups that are important to the study.

What are two disadvantages of stratified sampling?

1. If there are many variables of interest, dividing a large population into representative subgroups requires a great deal of effort.
2. If the variables are somewhat complex or ambiguous (such as beliefs, attitudes, or prejudices), it is difficult to separate individuals into the subgroups according to these variables.

Cluster Sampling

cluster sample – a sample obtained by selecting a preexisting or natural group, called a *cluster,* and using the members in the cluster for the sample.

List three advantages to using a cluster sample.

1. A cluster sample can reduce costs.
2. It can simplify fieldwork.
3. It is convenient. For example, in a dental study involving X-raying fourth-grade students' teeth to see how many cavities each child had, it would be a simple matter to select a single classroom and bring the X-ray equipment to the school to conduct the study. If other sampling methods were used, researchers might have to transport the machine to several different schools or transport the pupils to the dental office.

Explain the major disadvantage of cluster sampling.

The elements in a cluster may not have the same variations in characteristics as elements selected individually from a population.

The reason is that groups of people may be more homogeneous (alike) in specific clusters such as neighborhoods or clubs. For example, the people who live in a certain neighborhood tend to have similar incomes, drive similar cars, live in similar houses, and, for the most part, have similar habits.

Other Types of Sampling Techniques

sequence sampling – used in quality control, successive units taken from production lines are sampled to ensure that the products meet certain standards set by the manufacturing company.

double sampling – a very large population is given a questionnaire to determine those who meet the qualifications for a study. After the questionnaires are reviewed, a second, smaller population is defined. Then a sample is selected from this group.

multistage sampling – the researcher uses a combination of sampling methods.

For example, suppose a research organization wants to conduct a nationwide survey for a new product being manufactured. A sample can be obtained by using the following combination of methods. First the researchers divide the 50 states into four or five regions (or clusters). Then several states from each region are selected at random. Next the states are divided into various areas by using large cities and small towns. Samples of these areas are then selected. Next, each city and each town are divided into districts or wards. Finally, streets in these wards are selected at random, and the families living on these streets are given samples of the product to test and are asked to report the results.

List the 9 Steps in Conducting a Sample Survey

1. **Decide what information is needed.**
2. **Determine how the data will be collected (phone interview, mail survey, etc.).**
3. **Select the information-gathering instrument or design the questionnaire if one is not available.**
4. **Set up a sampling list, if possible.**
5. **Select the best method for obtaining the sample (random, systematic, stratified, cluster, or other).**
6. **Conduct the survey and collect the data.**
7. **Tabulate the data. (Put the data in table form.)**
8. **Conduct the statistical analysis.**
9. **Report the results.**

14–2 Surveys and Questionnaire Design READ pp. 734-736 and fill in notes

survey – when a sample of individuals is asked to respond to questions about a particular subject

List and explain two types of surveys.

1. **Interviewer administered surveys** – surveys that require a person to ask the questions of the subject. The interview can be conducted face to face in an office, on a street, or in the mall, or via telephone.

2. **Self-administered surveys** – surveys that are administered by the subject. Self-administered surveys can be done by mail or in a group setting such as a classroom.

What must you be careful of when analyzing the results of surveys?

When analyzing the results of surveys, you should be very careful about the interpretations. **The way a question is phrased can influence the way people respond.**

For example, when a group of people were asked if they favored a waiting period and background check before guns could be sold, 91% of the respondents were in favor of it and 7% were against it. However, when asked if there should be a national gun registration program costing about 20% of all dollars spent on crime control, only 33% of the respondents were in favor of it and 61% were against it.

What are 5 mistakes to be avoided when writing questions for a questionnaire?

1. ***Asking biased questions.*** By asking questions in a certain way, the researcher can lead the respondents to answer in the way he or she wants them to. For example, asking a question such as "Are you going to vote for the candidate Jones even though the latest survey indicates that he will lose the election?" instead of "Are you going to vote for candidate Jones?" may dissuade some people from answering in the affirmative.
2. ***Using confusing words.*** In this case, the participant misinterprets the meaning of the words and answers the questions in a biased way. For example, the question "Do you think people would live longer if they were on a diet?" could be misinterpreted since there are many different types of diets—weight loss diets, low-salt diets, medically prescribed diets, etc.
3. ***Asking double-barreled questions.*** Sometimes questions contain compound sentences that require the participant to respond to two questions at the same time. For example, the question "Are you in favor of a special tax to provide national health care for the citizens of the United States?" asks two questions: "Are you in favor of a national health care program?" and "Do you favor a tax to support it?"
4. ***Using double negatives in questions.*** Questions with double negatives can be confusing to the respondents. For example, the question "Do you feel that it is not appropriate to have areas where people cannot smoke?" is very confusing since *not* is used twice in the sentence.

5. ***Ordering questions improperly.*** By arranging the questions in a certain order, the researcher can lead the participant to respond in a way that he or she may otherwise not have done. For example, a question might ask the respondent, "At what age should an elderly person not be permitted to drive?" A later question might ask the respondent to list some problems of elderly people. The respondent may indicate that transportation is a problem based on reading the previous question.

What are some other factors that can bias a survey?

- The participant may not know anything about the subject of the question but will answer the question anyway to avoid being considered uninformed.
- Many people will make responses on the basis of what they think the person asking the questions wants to hear.
- Participants will, in some cases, respond differently to questions depending on whether their identity is known. This is especially true if the questions concern sensitive issues.
- The time and place of the survey
- The type of questions asked influences the responses – whether the question is open-ended or closed-ended.

Give an example of an open ended question.

- List three activities that you plan to spend more time on when you retire.
- List three things you would purchase if you won a million dollars.
- What are the three most important problems facing the world today.

Give an example of a closed ended question.

- Select three activities that you plan to spend more time on after you retire: traveling; eating out; fishing and hunting; exercising; visiting relatives.
- Select two things you would purchase if you won a million dollars: a home, a new car, a boat, a vacation, jewelry, food for the poor.
- Of the following which are the three most important problems facing the world today: global warming, hunger, disease epidemics, terrorism, nuclear annihilation.

List one problem with open ended questions and one problem with closed ended questions.

Open ended – the results may be so varied that attempting to summarize them might be difficult, if not impossible.

Closed ended – the respondent is forced to choose the answers that the researcher gives and cannot supply his or her own.

Using ANOTHER source, define the following words at the end of your notes.

Population – The entire group of objects or individuals considered for a survey.

Sample – A portion, piece, or segment that is representative of a whole. A set of elements drawn from and analyzed to estimate the characteristics of a population.

Census – The procedure of systematically acquiring and recording information about the members of a given population.

Experimental units – A unit is a person, animal, plant or thing which is actually studied by a researcher; the basic objects upon which the study or experiment is carried out. For example, a person; a monkey; a sample of soil; a pot of seedlings; a postcode area; a doctor's practice.

Subjects – The entities that are the object of study

Treatment – The substance or remedy so applied. What is being done to the subject.

Placebo effect – a beneficial health outcome resulting from a person's anticipation that an intervention—pill, procedure, or injection, for example—will help them.

Double blind experiment – an experimental method used to ensure impartiality, and avoid errors arising from bias.

Double-blind describes an especially stringent way of conducting an experiment, usually on living, conscious, human subjects. **A blind experiment is designed so that individuals do not know whether they are so-called "test" subjects or members of an "experimental control" group.**

In a double-blind experiment, neither the individuals nor the researchers know who belongs to the control group. Only after all the data are recorded (and in some cases, analyzed) may researchers be permitted to learn which individuals are which. Performing an experiment in double-blind fashion is a way to lessen the influence of prejudices and unintentional physical cues on the results.

Random digit table – A table of random digits is a listing of the numbers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. But what sets any listing of these digits apart from a table of random digits? There are two features of a table of random digits. **The first property is that every digit from 0 to 9 is just as likely to appear in every entry of the table. The second feature is that the entries are independent of each other.**

These properties imply that there is **no pattern to a table of random digits**. Information about some of the table will not help at all to determine the other entries of the table.

For convenience these digits can be arranged in rows of blocks. But any arrangement is really just for ease of reading. There is no pattern to the digits in the above row.

Simple random sample – any sampling method that has the following properties.

- The population consists of *N* objects.
- The sample consists of *n* objects.
- **If all possible samples of *n* objects are equally likely to occur, the sampling method is called simple random sampling.**

An important benefit of simple random sampling is that it allows researchers to use statistical methods to analyze sample results.

Types of flaws with surveys

Confounding – occurs when the experimental controls do not allow the experimenter to reasonably eliminate plausible alternative explanations for an observed relationship between independent and dependent variables. (There could be **lurking** variable.)

Voluntary response sample – one of the main types of non-probability sampling methods; A voluntary sample is made up of **people who self-select into the survey**. Often, these folks have a strong interest in the main topic of the survey.

Convenience sample – one of the main types of non-probability sampling methods; A convenience sample is made up of people who are easy to reach.

Example: A pollster interviews shoppers at a local mall. If the mall was chosen because it was a convenient site from which to solicit survey participants and/or because it was close to the pollster's home or business, this would be a convenience sample.

Undercoverage – In survey sampling, undercoverage is a type of selection bias . It occurs when some members of the population are inadequately represented in the sample.

Non-response – Sometimes, in survey sampling, individuals chosen for the sample are unwilling or unable to participate in the survey. Nonresponse bias is the bias that results when respondents differ in meaningful ways from nonrespondents. Nonresponse is often problem with mail surveys, where the response rate can be very low.

Types of bias

Wording of questions – **The wording of the question may be loaded in some way to unduly favor one response over another.** For example, a satisfaction survey may ask the respondent to indicate where she is satisfied, dissatisfied, or very dissatified. By giving the respondent one response option to express satisfaction and two response options to express dissatisfaction, this survey question is biased toward getting a dissatisfied response.

Response – bias that results from problems in the measurement process.

http://stattrek.com/survey-research/survey-bias.aspx

Given an example of when this class would be a sample and when it would be a population.

Sample – a nationwide survey of high school Statistics classes asking about intended college majors.

Population – a survey of Mrs. Cooper's students asking about AP exams this Spring.

Assignment: pp. 729-731 (#11-14, 17, 18)
p. 736 (1–10)