MATH 1342

Section 6.1

Sampling:

Terms:

Population – each element (or person) from the set of observations that can be made

Sample – a subset of the population

Census – systematically getting information about an entire population

Sampling – studying a part (a sample) in order to gain information about an entire group

Sampling Frame – the list of individuals from which a sample is actually selected

Types of Sampling –

Voluntary Response sample – people who choose themselves by responding to a general appeal (over represents people with strong opinions)

A simple random sample (SRS) consists of individuals from the population chosen in such a way that every set of individuals has an equal chance to be the sample actually selected.

A ***probability sample*** gives each member of the population a known chance to be selected.

    Probability is known but not necessarily equal

A stratified sample divides the population into groups of similar individuals, called strata, and chooses a SRS in each stratum and combines these to form the full sample.

    This guarantees representation in each of these different groups

In multistage sample design samples are taken from various subsets of the population until a manageable number of samples to interview are arrived upon.

Convenience sampling is a non-probability type of sample where the sample is chose based on their convenient accessibility and proximity.

Popper 18:

1. A sample of 50 was selected by randomly drawing tickets from a hat. *A*

2. A sample of 100 was selected by subjects filling out a survey on a website obtained from a posted sign. *D*

3. A sample of 200 was obtained by surveying individuals in one classroom for a population of the entire university. *C*

4. A sample of 100 selected by randomly drawing names from a hat (women had their names entered twice, and men entered once). *B*

5. The population was divided into BA and BS majors and random samples were selected from each group, and then combined. *E*

a. Simple Random Sample

b. Probability Sample

c. Convenience Sample

d. Voluntary Response

e. Stratified Sample

Random Digit Table:

Table of Random Digits

Line								
101	98360	26534	47384	94612	88666	14170	10847	05567
102	55556	59863	86607	00094	77213	35711	52851	42108
103	31634	15399	73476	77412	06186	16636	54307	14947
104	13785	11509	54891	98375	68377	50572	08453	80376
105	80376	73842	95465	59746	38078	25727	78502	95324

Appendix B of your textbook

Line 101: 98 36 02 65 34

Random Digits

A table of random digits is a long string of the digits 0 – 9 where each entry in the table is equally likely to be any of the 10 digits and the entries are independent of each other.

Example: You have 15 employees and you wish to select 5 of them for a training program. Create a selection process utilizing a Random Digit Table.

$$\frac{100}{15} = 6.7$$

Assign each of the 15 employees 6 2-digit numbers.
Andy: 00, 01, 02, 03, 04, 05
Bobby: 06, 07, 08, 09, 10, 11
Cathy: 12, 13, 14, 15, 16, 17
Etc....
Any unassigned numbers your skip.

Continue until you have all 5 selected.

Experiment – actively impose some treatment in order to observe the response

Observational study – investigators observe subjects and measure variables of interest without assigning treatments to the subjects.

Two elements are *confounded* when their effects on a response variable cannot be distinguished from one another.

Statistical inference provides ways to answer specific questions from data with some guarantee that the answers are good ones. In inference we must think about how to produce data as well as analyze data.

The design of sample refers to the method used to choose the sample from the whole population.

       * Voluntary response and convenience sampling are examples of bad sample design.

The design of a study is biased if it systematically favors certain outcomes.

Undercoverage occurs when some groups in the population are left out of the process of choosing the sample.

Nonresponse occurs when an individual chosen for the sample can't be contacted or refuses to cooperate.

Things to watch out for in interviewing technique:

- Response Bias – when an interviewer's attitude suggests that some answers are more desirable than others gives the interviewer specific answers more often

     General term that includes many things: wording, body language, tone of voice, dress

- Wording of Questions – confusing or misleading questions can introduce strong bias

     Specific type of response bias where the wording of the questions in the concern

Examples:
1. Identify the population and the sample then describe the sampling method that was used.

To conduct a pre-election opinion poll on a proposed city ordinance, a random sample of telephone numbers from the city phone book were chosen and called. (Assume all who were called answered).

Examples:
1. Identify the population and the sample then describe the sampling method that was used.

To conduct a pre-election opinion poll on a proposed city ordinance, a random sample of telephone numbers from the city phone book were chosen and called. (Assume all who were called answered).

Population: Residents of the city

Sample: The people contacted

Sampling Method: Simple Random Sampling

Sampling Frame: Phone Book

2. Determine if the study is an experiment or an observational study. Give a reason for your answer.

a. A personnel director at a large company studied the eating habits of employees by watching the movements of a selected group of employees at lunchtime. The purpose of the study was to determine the proportion of employees who buy lunch in the cafeteria, bring their own lunches, or go out to lunch.

Observational Study: the researcher is just watching people in their day-to-day activities. There is no treatment or control imposed.

b. A pharmacy student would like to know if there is a difference in results from a specific brand of drug and its equivalent generic prescription. She randomly selects 50 people who take the drug and has them complete a questionnaire regarding their symptoms and improvements after taking the specific brand or generic equivalent.

Experiment: The researcher is imposing a treatment (brand name vs generic) and there is a degree of control involved. (The researcher decides who received which medication)

3. It is believed that 75% of all apartment dwellers in a large city deadbolt their doors in addition to locking them as an added precaution against burglary. Describe how you would select an SRS of 20 apartment dwellers to survey if there are 50 complexes in the city and each complex has 250 residents. (Use the random digit table)

Begin by selecting 10 out of the 50 apartment complexes:
Assign each of the 50 complexes 2 2-digit numbers: 00 01, 02 03, 04 05, …. 98 99
Read a line of the table until you have your first 10 selections.

Assign each resident 4 3-digit numbers: 000 001 002 003, 004 005 006 007, … 996 997 998 999
The first two apartments selected will be in the sample.

Popper 18: (a) Response Bias (b) Wording of Questions (c) Sampling Bias

Under/Coverage Bias

Bias is present in each of the following sample designs in the situations below. In each case, describe the type of bias involved and state whether you think the sampling frequency obtained is lower or higher than the actual population parameter.

6. A political pollster seeks information about the proportion of American adults that oppose gun control. He asks a SRS of 1000 American adults, "Do you agree or disagree with the following statement: Americans should preserve their constitutional right to keep and bear arms." A total of 910 or 91% agreed.

Wording of Questions (Choice B): By using the phrase "preserve their constitutional rights"

Popper 18: (a) Response Bias (b) Wording of Questions (c) Sampling Bias

Bias is present in each of the following sample designs in the situations below. In each case, describe the type of bias involved and state whether you think the sampling frequency obtained is lower or higher than the actual population parameter.

7. A restaurant chain wants to know what percentage of American families go out to eat for dinner at least 3 nights per week. They call a SRS of 1000 households between the hours of 6:00 p.m. and 8:00 p.m. and talk to 400 people. 300 of those 400 said they do not eat out 3 or more times a week.

Sampling Bias (undercoverage) (Choice C): By calling homes between 6:00 and 8:00 pm you are excluding the population that is eating out for dinner.