Study Session Week of 2/26

- I will review terms from the experimental design unit

- I will apply that knowledge to solve problems involving experimental design

Agenda:

- Review

- MC practice problems

- FR practice problem

How are we going to choose (or assign to treatments) the subjects or experimental units?

- Simple random sample
- Stratified random sample
- Systematic random sample
- Cluster sample
- Voluntary response sample
- Convenience sample
- Completely randomized design
- Blocked design
- Paired design

What are we going to do once we have our sample?

- Survey – what will you ask and exactly how will the question be worded?

- Observational study – what data will you collect? What type of data is it?

- Experiment – exactly how will you carry your experiment out?

  - Explanatory variable
  - Response variable
  - Treatment
  - Placebo
  - Blinding
  - Double blind
  - Control group

What can go wrong?

- Bias – types?

- Confounding variables

- Undercoverage
- Nonresponse
- Response
- Voluntary response

Practice FRAPPY – 2013 #2

2. An administrator at a large university wants to conduct a survey to estimate the proportion of students who are satisfied with the appearance of the university buildings and grounds. The administrator is considering three methods of obtaining a sample of 500 students from the 70,000 students at the university.

(a) Because of financial constraints, the first method the administrator is considering consists of taking a convenience sample to keep the expenses low. A very large number of students will attend the first football game of the season, and the first 500 students who enter the football stadium could be used as a sample. Why might such a sampling method be biased in producing an estimate of the proportion of students who are satisfied with the appearance of the buildings and grounds?

(b) Because of the large number of students at the university, the second method the administrator is considering consists of using a computer with a random number generator to select a simple random sample of 500 students from a list of 70,000 student names. Describe how to implement such a method.

(c) Because stratification can often provide a more precise estimate than a simple random sample, the third method the administrator is considering consists of selecting a stratified random sample of 500 students. The university has two campuses with male and female students at each campus. Under what circumstance(s) would stratification by campus provide a more precise estimate of the proportion of students who are satisfied with the appearance of the university buildings and grounds than stratification by gender?

Scoring

An administrator at a large university wants to conduct a survey to estimate the proportion of students who are satisfied with the appearance of the university buildings and grounds. The administrator is considering three methods of obtaining a sample of 500 students from the 70,000 students at the university.

Intent of Question

The primary goals of this question were to assess a student's ability to (1) recognize and explain why a particular sampling method is likely to be biased; (2) describe a method for selecting a simple random sample from a population using a computer random number generator; and (3) demonstrate an understanding of the principle of stratification by describing circumstances in which one stratification variable would be better than another.

Scoring

Because of financial constraints, the first method the administrator is considering consists of taking a convenience sample to keep the expenses low. A very large number of students will attend the first football game of the season, and the first 500 students who enter the football stadium could be used as a sample. Why might such a sampling method be biased in producing an estimate of the proportion of students who are satisfied with the appearance of the buildings and grounds?

The first 500 students who enter the football stadium were not likely to be representative of the population of all students at the university. In other words, these 500 students were likely to differ systematically from the population with regard to many variables. For example, these 500 students might have more school pride than the population of students as a whole, which might be related to their opinions about the appearance of university buildings and grounds. Perhaps their school pride is related to having more positive opinions about the appearance of university buildings and grounds, in which case the sample proportion of students who were satisfied would be biased toward overestimating the population proportion of students who were satisfied.

Essentially correct (E) if the response correctly includes the following three components:

1. Provides a reasonable explanation for why the sample might not be representative of the population;
2. Mentions a link between the nonrepresentative nature of the convenience sample and the variable of interest (opinion about appearance of university buildings and grounds);
3. Indicates a plausible direction for the bias of the estimator by:
   o Explicitly identifying the direction of the bias in the estimate of the population proportion of students satisfied with the appearance of the buildings and grounds,
   OR
   o Stating or implying that the students in the sample were more (or less) likely to be satisfied with the appearance of the buildings and grounds than those not in the sample.

Partially correct (P) if the response correctly provides exactly two of the three components listed above.

Incorrect (I) if the response correctly provides one or none of the three components listed above.

Scoring

Because of the large number of students at the university, the second method the administrator is considering consists of using a computer with a random number generator to select a simple random sample of 500 students from a list of 70,000 student names. Describe how to implement such a method.

Obtain a list of all 70,000 students at the university. Assign an identification number from 1 to 70,000 to each student.

Then use a computer to generate 500 random integers between 1 and 70,000 without replacement. The students whose ID numbers correspond to those numbers were then selected for the sample.

Essentially correct (E) if the response correctly includes the following three components:

1. Assigns numbers to the student names;
2. Uses a computer random number generator to randomly generate 500 distinct/unique numbers between 1 and 70,000;
3. Selects students whose names correspond to the 500 random numbers for the sample.

Partially correct (P) if the response correctly includes two of the three components listed above (with the exception of the second reason given for an (I) below).

Incorrect (I) if the response correctly includes no more than one of these three components;
   OR
if the response proposes implementing a sampling method other than simple random sampling (for example, systematic sampling).

Scoring

Because of the large number of students at the university, the second method the administrator is considering consists of using a computer with a random number generator to select a simple random sample of 500 students from a list of 70,000 student names. Describe how to implement such a method.

Stratifying by campus would be more advantageous than stratifying by gender provided that opinions about appearance of university buildings and grounds between the two campuses differ more than the opinions about appearance of university buildings and grounds between the two genders.

Essentially correct (E) if the response correctly notes that the circumstance described requires more variability in opinions about appearance of university buildings and grounds between the two campuses than between the two genders.

Partially correct (P) if the response says that the circumstance described requires considerable variability in opinions about appearance of university buildings and grounds between the two campuses without explicitly comparing to variability between the two genders,
OR
if the response only says that the circumstance described requires more variability between the two campuses than between the two genders without referring to opinions about appearance of university buildings and grounds,
OR
if the response notes that the circumstance described requires homogeneity of opinions about appearance of university buildings and grounds within the two campuses.

Incorrect (I) if the response does not meet the criteria for E or P.

Scoring

4	**Complete Response**
	All three parts essentially correct
3	**Substantial Response**
	Two parts essentially correct and one part partially correct
2	**Developing Response**
	Two parts essentially correct and one part incorrect
OR	
	One part essentially correct and one or two parts partially correct
OR	
	Three parts partially correct
1	**Minimal Response**
	One part essentially correct and two parts incorrect
OR	
	Two parts partially correct and one part incorrect

Other Questions/Comments?