University of Texas at Austin

Problem Set # 1

Producing data.

Provide your complete solutions for the following problems.

Problem 1.1. In an experiment on a new drug, subjects were randomly assigned to either a placebo or the active drug. In addition, the method of delivery of the drug (pill, skin patch, or nasal mist) was considered. In this experiment, there were how many factors?

Solution: 2
A factor is an explanatory variable of interest in an experiment. This is, in essence, a general category for what is being manipulated. In this experiment, we randomized whether or not the patient got the drug and how it was given.

Problem 1.2. A researcher is interested in the cholesterol levels of adults in the city in which she lives. A cholesterol screening program is set up in the downtown area during the lunch hour. Individuals can walk in and have their cholesterol determined for free. The service is used by 173 people, and their average cholesterol is 217.8. The sample obtained is an example of what kind of sample? Is it: *simple, stratified, biased*, e.g?

Solution: This is a sample probably containing **bias** and **undercoverage.**
Individuals "volunteered" for the screening. Individuals who knew their cholesterol was fine may not have bothered to have it checked, which would cause the average from the screening to be too high. The downtown area was not equally accessible to all adults in the population. Those who work in the downtown area are more likely to be in the sample, while other groups may not have known about it or may not have been able to reach the screening center, which would lead to undercoverage.

Problem 1.3. Two amateur gardeners are interested in comparing the yields of two varieties of tomatoes. They each have small backyard gardens. Each gardener is going to plant three plants of each variety in his garden. The first gardener will select six small areas in his garden for planting, then choose three of these at random for the three plants of the first variety and then use the remaining three for the second variety. The second gardener will follow the same procedure with his own randomization in his garden. At the end of the growing season they will compare the yields of the two varieties. What do the gardens represent in this example? Are they: *randomizers, samples, lurking variables, explanatory variables, blocks, . . .* ?

Solution: Blocks.
A block is a group of experimental units that are known to be similar in some way that is expected to affect the response to the treatments. In this case, the six small areas within each garden should be similar to each other, so the gardens are the blocks. In a block design, the random assignment of the units (the small areas) to the treatments (varieties) is carried separately within each block (the garden).

Problem 1.4. In order to assess the opinion of students at the University of Montana on campus snow removal, a reporter for the student newspaper interviews the first 12 students he meets who are willing to express their opinion. What is the population in this case?

Solution: All students at the University of Montana.
The group of individuals we want information about is the population. In this case it would have been better to have obtained a simple random sample rather than sampling the first 12 students who are willing to give their opinion, but that doesn't change the population.

Problem 1.5. Consider a study performed by a medical center to determine which of two heart surgeries is most effective: angioplasty (running plastic tubes through the arteries) or bypass (rerouting arteries). The purpose of either procedure is to prolong the life of the patient. The study records the survival time of each patient (measured from the time of the surgery). Identify the response and explanatory variables in this study!

Solution: Response: survival time, explanatory: type of surgery.
The purpose of the study is to observe the effect of two kinds of operations on prolonging the lifetime of patients. Therefore, the response variable is the survival time, and the explanatory variable is the kind of procedure that was used.

Problem 1.6. A sociologist wants to study the attitudes of American male college students toward marriage and husband-wife relations. She gives a questionnaire to 25 of the men enrolled in Sociology 101 at her college. All 25 complete and return the questionnaire. Identify the sample in this situation!

Solution: The 25 respondents.

Problem 1.7. A statistics instructor wants to know which route will get her to school the fastest. Each day from October 2 to November 15, when she gets to the turn point she checks the odometer on her car. If it shows an even number, she takes the freeway; if it shows an odd number, she takes the in-town route. She records the total time each day. What is the explanatory variable in this study? What is the response variable in this study? What kind of a study is this (an *observations study* or an *experiment* or . . .)?

Solution: Explanatory: Route, Response: time to school, Type: Experiment.

Problem 1.8. A researcher is studying the effects of a new drug on reducing high blood pressure. He recruits 250 men to test the new active drug against a current standard. At the end of six weeks, the decrease in systolic blood pressure will be evaluated. He believes the drug will be more effective for black men than for white men. What kind of experimental design should he use to properly test his belief?

Solution: Blocked.
In a blocked experiment, the subjects are randomly assigned treatments within their blocks (like subgroups). This ensures that each treatment is replicated within the blocks and will allow us to make conclusions about the effectiveness of each treatment in each block.