AP Statistics

Review #2
Producing Data

Name _____

II. Sampling and Experimentation: Planning and conducting a study (10%-15%)
Data must be collected according to a well-developed plan if valid information on a conjecture is to be obtained. This plan includes clarifying the question and deciding upon a method of data collection and analysis.

A. Overview of methods of data collection

1. Census
2. Sample survey
3. Experiment
4. Observational study

B. Planning and conducting surveys

1. Characteristics of a well-designed and well-conducted survey
2. Populations, samples, and random selection
3. Sources of bias in sampling and surveys
4. Sampling methods, including simple random sampling, stratified random sampling, and cluster sampling

C. Planning and conducting experiments

1. Characteristics of a well-designed and well-conducted experiment
2. Treatments, control groups, experimental units, random assignments, and replication
3. Sources of bias and confounding, including placebo effect and blinding

4. Completely randomized design

5. Randomized block design, including matched pairs design

D. Generalizability of results and types of conclusions that can be drawn from observational studies, experiments, and surveys

Tips and Hints:

- Know what is required for a sample to be a simple random sample (SRS). If each individual in the population has an equal probability of being chosen for a sample, it doesn't necessarily follow that the sample is an SRS. Consider a class of six boys and six girls. I want to randomly pick a committee of two students from this group. I decide to flip a coin. If "heads," I will choose two girls by a random process. If "tails," I will choose two boys by a random process. Now, each student has an equal probability (1/6) of being chosen for the committee. However, the two students are not an SRS of size two picked from members of the class. Why not? Because this selection process does not allow for a committee consisting of one boy and one girl. To have an SRS of size two from the class, each group of two students would have to have an equal probability of being chosen as the committee.
- SRS refers to how you obtain your sample; random assignment (allocation) is what you use in an experiment to assign subjects to treatment groups. They are not synonyms.
- Well-designed experiments satisfy the principles of control, randomization, and replication.
  * Control for the effects of lurking variables by comparing several treatments in the same environment.
  *Note: Control is not synonymous with "control group."
  * Randomization refers to the random allocation of subjects to treatment groups, and not to the selection of subjects for the experiment. Randomization is an attempt to "even out" the effects of lurking variables across the treatment groups. Randomization helps avoid bias.
  * Replication means using a large enough number of subjects to reduce chance variation in a study. Note: In science, replication often means, "do the experiment again."
- Distinguish the language of surveys from the language of experiments. Stratifying is for sampling and Blocking is for experiments
- It is not enough to memorize the terminology related to surveys, observational studies, and experiments. You must be able to apply the terminology in context. For example:
  Blocking refers to a deliberate grouping of subjects in an experiment based on a characteristic (such as gender, cholesterol level, race, or age) that you suspect will affect responses to treatments in a systematic way. After blocking, you should randomly assign subjects to treatments within the blocks. Blocking reduces unwanted variability.
- An experiment is double blind if neither the subjects nor the experimenters know who is receiving what treatment. A third party can keep track of this information.
- Suppose that subjects in an observational study who eat an apple a day get significantly fewer cavities than subjects who eat less than one apple each week. A possible confounding variable is overall diet. Members of the apple-a-day group may tend to eat fewer sweets, while those in the non-apple-eating group may turn to sweets as a regular alternative. Since different diets could contribute to the disparity in cavities between the two groups, we cannot say that eating an apple each day **causes** a reduction in cavities.

Experiments vs. Samples
Many students confuse experimentation with sampling or try to incorporate ideas from one into the other. This is not totally off-base since some concepts appear in both areas, but it is important to keep them straight. The purpose of sampling is to estimate a population parameter by measuring a representative subset of the population. We try to create a representative sample by selecting subjects randomly using an appropriate technique. The purpose of an experiment is to demonstrate a cause and effect relationship by controlling extraneous factors. Experiments are rarely performed on random samples because both ethics and practicality make it impossible to do so. For this reason, there is always a concern of how far we can generalize the results of an experiment. Generalizing results to a population unlike the subjects in the experiment is very unreliable.

Blocking vs. Stratifying
Students often ask, "What is the difference between blocking and stratifying?" The simple answer is that blocking is done in experiments and stratifying is done with samples. There are similarities between the two, namely the dividing up of subjects before random assignment or selection, but the words are definitely not interchangeable.

Blocking
In blocking we divide our subjects up in advance based on some factor we know or believe is relevant to the study and then randomly assign treatments within each block. The key things to remember:

- You don't just block for the heck of it. You block based on some factor that you think will impact the response to the treatment
- The blocking is not random. The randomization occurs within each block essentially creating 2 or more miniature experiments.
- Blocks should be homogenous (i.e. alike) with respect to the blocking factor. For example, I want to find out if playing classical music during tests will result in higher mean scores. I could randomly assign half my students to the room with the music and the other half to the normal room, but I know that my juniors consistently score higher than my seniors, and I want to account for this source of variation in the results. I block according to grade by separating the juniors and seniors first and then randomly assigning half the juniors to the music room and the other half to the normal room. I do the same with the seniors. For this design to be valid, I have to expect that each grade will respond to the music similarly. In other words, I know that juniors will score higher, but I expect to see a similar improvement or decline in both groups as a result of having the music. At the end of my study I can subtract out the effect of grade level to reduce the unaccounted for variation in the results.

You have learned how to analyze the results of one special type of blocked design, namely, matched pairs. In matched pairs you subtract each pair of values which eliminates the variation due to the subject. Similar techniques are available for fancier blocked designs.

Stratified Sampling vs. Cluster Sampling
Many students confuse stratified and cluster sampling since both of them involve groups of subjects. There are 2 key differences between them. First, in stratified sampling we divide up the population based on some factor we believe is important, but in cluster sampling the groups are naturally occurring (I picture schools of fish). Second, in stratified sampling we randomly select subjects from each stratum, but in cluster sampling we randomly select one or more clusters and measure every subject in each selected cluster. (Note: There are more advanced techniques (multistage) in which samples are taken within the cluster(s))

Final Thoughts
It is especially important to stay focused when answering questions about design. Too many students get caught up in minor details but miss the big ideas of randomization and control. Always remember that your mission in responding to questions is to demonstrate your understanding of the major concepts of the course.

CLEAR COMMUNICATION IS KEY!!!

2011B #2

2. People with acrophobia (fear of heights) sometimes enroll in therapy sessions to help them overcome this fear. Typically, seven or eight therapy sessions are needed before improvement is noticed. A study was conducted to determine whether the drug D-cycloserine, used in combination with fewer therapy sessions, would help people with acrophobia overcome this fear.

   Each of 27 people who participated in the study received a pill before each of two therapy sessions. Seventeen of the 27 people were randomly assigned to receive a D-cycloserine pill, and the remaining 10 people received a placebo. After the two therapy sessions, none of the 27 people received additional pills or therapy. Three months after the administration of the pills and the two therapy sessions, each of the 27 people was evaluated to see if he or she had improved.

   (a) Was this study an experiment or an observational study? Provide an explanation to support your answer.

   (b) When the data were analyzed, the D-cycloserine group showed statistically significantly more improvement than the placebo group did. Based on this result, would the researchers be justified in concluding that the D-cycloserine pill and two therapy sessions are as beneficial as eight therapy sessions without the pill? Justify your answer.

   (c) A newspaper article that summarized the results of this study did not explain how it was determined which people received D-cycloserine and which received the placebo. Suppose the researchers allowed the therapists to choose which people received D-cycloserine and which received the placebo, and no randomization was used. Explain why such a method of assignment might lead to an incorrect conclusion.

2010 #4

4. An automobile company wants to learn about customer satisfaction among the owners of five specific car models. Large sales volumes have been recorded for three of the models, but the other two models were recently introduced so their sales volumes are smaller. The number of new cars sold in the last six months for each of the models is shown in the table below.

Car Model	A	B	C	D	E	Total
Number of new cars sold in the last six months	112,338	96,174	83,241	3,278	2,323	297,354

   The company can obtain a list of all individuals who purchased new cars in the last six months for each of the five models shown in the table. The company wants to sample 2,000 of these owners.

   (a) For simple random samples of 2,000 new car owners, what is the expected number of owners of model E and the standard deviation of the number of owners of model E?

   (b) When selecting a simple random sample of 2,000 new car owners, how likely is it that fewer than 12 owners of model E would be included in the sample? Justify your answer.

   (c) The company is concerned that a simple random sample of 2,000 owners would include fewer than 12 owners of model D or fewer than 12 owners of model E. Briefly describe a sampling method for randomly selecting 2,000 owners that will ensure at least 12 owners will be selected for each of the 5 car models.

2009 #3

3. Before beginning a unit on frog anatomy, a seventh-grade biology teacher gives each of the 24 students in the class a pretest to assess their knowledge of frog anatomy. The teacher wants to compare the effectiveness of an instructional program in which students physically dissect frogs with the effectiveness of a different program in which students use computer software that only simulates the dissection of a frog. After completing one of the two programs, students will be given a posttest to assess their knowledge of frog anatomy. The teacher will then analyze the changes in the test scores (score on posttest minus score on pretest).

   (a) Describe a method for assigning the 24 students to two groups of equal size that allows for a statistically valid comparison of the two instructional programs.

   (b) Suppose the teacher decided to allow the students in the class to select which instructional program on frog anatomy (physical dissection or computer simulation) they prefer to take, and 11 students choose actual dissection and 13 students choose computer simulation. How might that self-selection process jeopardize a statistically valid comparison of the changes in the test scores (score on posttest minus score on pretest) for the two instructional programs? Provide a specific example to support your answer.

2008B #4

4. A researcher wants to conduct a study to test whether listening to soothing music for 20 minutes helps to reduce diastolic blood pressure in patients with high blood pressure, compared to simply sitting quietly in a noise-free environment for 20 minutes. One hundred patients with high blood pressure at a large medical clinic are available to participate in this study.

   (a) Propose a design for this study to compare these two treatments.

   (b) The null hypothesis for this study is that there is no difference in the mean reduction of diastolic blood pressure for the two treatments and the alternative hypothesis is that the mean reduction in diastolic blood pressure is greater for the music treatment. If the null hypothesis is rejected, the clinic will offer this music therapy as a free service to their patients with high blood pressure. Describe Type I and Type II errors and the consequences of each in the context of this study, and discuss which one you think is more serious.

2008 #2

2. A local school board plans to conduct a survey of parents' opinions about year-round schooling in elementary schools. The school board obtains a list of all families in the district with at least one child in an elementary school and sends the survey to a random sample of 500 of the families. The survey question is provided below.

   A proposal has been submitted that would require students in elementary schools to attend school on a year-round basis. Do you support this proposal? (Yes or No)

   The school board received responses from 98 of the families, with 76 of the responses indicating support for year-round schools. Based on this outcome, the local school board concludes that most of the families with at least one child in elementary school prefer year-round schooling.

   (a) What is a possible consequence of nonresponse bias for interpreting the results of this survey?

   (b) Someone advised the local school board to take an additional random sample of 500 families and to use the combined results to make their decision. Would this be a suitable solution to the issue raised in part (a) ? Explain.

   (c) Suggest a different follow-up step from the one suggested in part (b) that the local school board could take to address the issue raised in part (a).

2007 #2

2. As dogs age, diminished joint and hip health may lead to joint pain and thus reduce a dog's activity level. Such a reduction in activity can lead to other health concerns such as weight gain and lethargy due to lack of exercise. A study is to be conducted to see which of two dietary supplements, glucosamine or chondroitin, is more effective in promoting joint and hip health and reducing the onset of canine osteoarthritis. Researchers will randomly select a total of 300 dogs from ten different large veterinary practices around the country. All of the dogs are more than 6 years old, and their owners have given consent to participate in the study. Changes in joint and hip health will be evaluated after 6 months of treatment.

   (a) What would be an advantage to adding a control group in the design of this study?

   (b) Assuming a control group is added to the other two groups in the study, explain how you would assign the 300 dogs to these three groups for a completely randomized design.

   (c) Rather than using a completely randomized design, one group of researchers proposes blocking on clinics, and another group of researchers proposes blocking on breed of dog. How would you decide which one of these two variables to use as a blocking variable?

2006 #5

5. A biologist is interested in studying the effect of growth-enhancing nutrients and different salinity (salt) levels in water on the growth of shrimps. The biologist has ordered a large shipment of young tiger shrimps from a supply house for use in the study. The experiment is to be conducted in a laboratory where 10 tiger shrimps are placed randomly into each of 12 similar tanks in a controlled environment. The biologist is planning to use 3 different growth-enhancing nutrients (A, B, and C) and two different salinity levels (low and high).

   (a) List the treatments that the biologist plans to use in this experiment.

   (b) Using the treatments listed in part (a), describe a completely randomized design that will allow the biologist to compare the shrimps' growth after 3 weeks.

   (c) Give one statistical advantage to having only tiger shrimps in the experiment. Explain why this is an advantage.

   (d) Give one statistical disadvantage to having only tiger shrimps in the experiment. Explain why this is a disadvantage.

2005B #3

3. In search of a mosquito repellent that is safer than the ones that are currently on the market, scientists have developed a new compound that is rated as less toxic than the current compound, thus making a repellent that contains this new compound safer for human use. Scientists also believe that a repellent containing the new compound will be more effective than the ones that contain the current compound. To test the effectiveness of the new compound versus that of the current compound, scientists have randomly selected 100 people from a state.

   Up to 100 bins, with an equal number of mosquitoes in each bin, are available for use in the study. After a compound is applied to a participant's forearm, the participant will insert his or her forearm into a bin for 1 minute, and the number of mosquito bites on the arm at the end of that time will be determined.

   (a) Suppose this study is to be conducted using a completely randomized design. Describe a randomization process and identify an inference procedure for the study.

   (b) Suppose this study is to be conducted using a matched-pairs design. Describe a randomization process and identify an inference procedure for the study.

   (c) Which of the designs, the one in part (a) or the one in part (b), is better for testing the effectiveness of the new compound versus that of the current compound? Justify your answer.

2005 #5

5. A survey will be conducted to examine the educational level of adult heads of households in the United States. Each respondent in the survey will be placed into one of the following two categories:

   - Does not have a high school diploma
   - Has a high school diploma

   The survey will be conducted using a telephone interview. Random-digit dialing will be used to select the sample.

   (a) For this survey, state one potential source of bias and describe how it might affect the estimate of the proportion of adult heads of households in the United States who do not have a high school diploma.

   (b) A pilot survey indicated that about 22 percent of the population of adult heads of households do not have a high school diploma. Using this information, how many respondents should be obtained if the goal of the survey is to estimate the proportion of the population who do not have a high school diploma to within 0.03 with 95 percent confidence? Justify your answer.

   (c) Since education is largely the responsibility of each state, the agency wants to be sure that estimates are available for each state as well as for the nation. Identify a sampling method that will achieve this additional goal and briefly describe a way to select the survey sample using this method.