7 Great Design Problems

(AP Statistics Exam Problems) NAME___________________________________

1. Agricultural experts are trying to develop a bird deterrent to reduce costly damage to crops in the United States. An experiment is to be conducted using garlic oil to study its effectiveness as a nontoxic, environmentally safe bird repellant. The experiment will use European starlings, a bird species that causes considerable damage annually to the corn crop in the United States. Food granules made from corn are to be infused with garlic oil in each of five concentrations of garlic —0 percent, 2 percent, 10 percent, 25 percent, and 50 percent. The researchers will determine the adverse reaction of the birds to the repellant by measuring the number of food granules consumed during a two-hour period following overnight food deprivation. There are forty birds available for the experiment, and the researchers will use eight birds for each concentration of garlic. Each bird will be kept in a separate cage and provided with the same number of food granules.

   (a) For the experiment, identify

      i. the treatments

      ii. the experimental units

      iii. the response that will be measured

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?

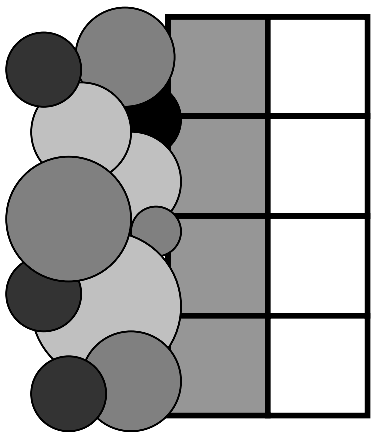
(b) Assuming a control group is added to the other two groups in the study, explain how you would assign the 300 dogs to these three groups for a completely randomized design.

(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?

3. Students are designing an experiment to compare the productivity of two varieties of dwarf fruit trees. The site for the experiment is a field that is bordered by a densely forested area on the west side. The field has been divided into eight plots of approximately the same area. The students have decided that the test plots should be blocked. Four trees, each of two varieties, will be assigned at random to the four plots within each block, with one tree planted in each plot.

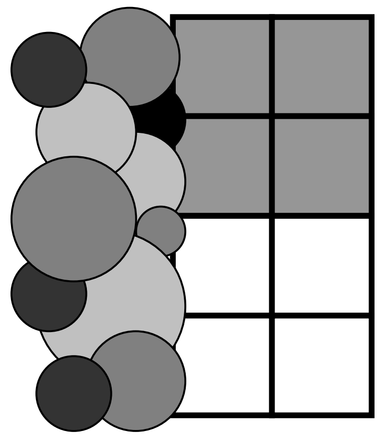
   The two blocking schemes shown below are under consideration. For each scheme, one block is identified by the white region and the other block indicated by the grey region in the figures.

Blocking Scheme A



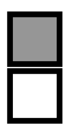
Forest

Blocking Scheme B



Forest

<u>Key</u>



Block 1

Block 2

   a. Which of the blocking schemes , A or B, is better for this experiment? Explain.

   b. Even though the students have decided to block, they must randomly assign the varieties of the trees to the plots within each block. What is the purpose of this randomization in the context of the experiment?

4. The dentists in a dental clinic would like to determine if there is a difference between the number of new cavities in people who eat an apple a day and in people who eat less than one apple a week. They are going to conduct a study with 50 people in each group.

Fifty clinics patients who report that they routinely eat an apple a day and 50 clinic patients who report that they eat less than one apple a week will be identified. The dentists will examine the patients and their records to determine the number of new cavities the patients have had over the past two years. They will then compare the number of new cavities in the two groups.

a. Why is this an observational study and not an experiment?

b. Explain the concept of confounding in the context of this study. Include an example of a possible confounding variable.

c. If the mean number of new cavities for those who ate an apple a day was statistically significantly smaller than the mean number of new cavities for those who ate less than one apple a week, could one conclude that the lower number of new cavities can be attributed to eating an apple a day? Explain.

5. A biologist is interested in studying the effect of growth-enhancing nutrients an differernt 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.

6. A manufacturer of boots plans to conduct an experiment to compare a new method of waterproofing to the current method. The appearance of the boots is not changed by either method. The company recruits 100 volunteers in Seattle, where it rains frequently, to wear the boots as they normally would for 6 months. At the end of the 6 months, the boots will be returned to the company to be evaluated for water damage.

   (a) Describe a design for this experiment that uses the 100 volunteers. Include a few sentences on how it would be implemented.

   (b) Could your design be double blind? Explain

7. When a tractor pulls a plow through an agricultural field, the energy needed to pull that plow is called the draft. The draft is affected by environmental conditions such as soil type, terrain, and moisture.

A study was conducted to determine whether a newly developed hitch would be able to reduce draft compared to the standard hitch. (A hitch is used to connect the plow to the tractor.) Two large plots of land were used in this study. It was randomly determined which plot was to be plowed using the standard hitch. As the tractor plowed that plot, a measurement device on the tractor automatically recorded the draft at 25 randomly selected points in the plot.

After the plot was plowed, the hitch was changed from the standard one to the new one, a process that takes a substantial amount of time. Then the second plot was plowed using the new hitch. Twenty-five measurements of draft were also recorded at randomly selected points in this plot.

a) What was the response variable in this study?

Identify the treatments.

What were the experimental units?

b) Given that the goal of the study is to determine whether a newly developed hitch reduces draft compared to the standard hitch, was randomization used properly in this study? Justify your answer.

c) Given that the goal of the study is to determine whether a newly developed hitch reduces draft compared to the standard hitch, was replication used properly in this study? Justify your answer.

d) Plot of land is a confounding variable in this experiment. Explain why.

SOLUTIONS to 7 Great Experimental Design Problems

1. (2010 #1a)

i. The treatments are the different concentrations of garlic in the food granules. Specifically, there are five treatments: 0 percent, 2 percent, 10 percent, 25 percent and 50 percent.
ii. The experimental units are the birds (starlings), each placed in an individual cage.
iii. The response is the number of food granules consumed by the bird.

2. (2007 #2)

Part (a):

A control group gives the researchers a comparison group to be used to evaluate the effectiveness of the treatments. The control group allows the impact of the normal aging process on joint and hip health to be measured with appropriate response variables. The effects of glucosamine and chondroitin can be assessed by comparing the responses for these two treatment groups with those for the control group.

Part (b):

Each dog will be assigned a unique random number, 001–300, using a random number generator on a calculator, statistical software, or a random number table. The numbers will be sorted from smallest to largest. The dogs assigned the first 100 numbers in the ordered list will receive glucosamine. The dogs with the next 100 numbers in the ordered list will be assigned to the control group. Finally, the dogs with the numbers 201–300 will receive chondroitin.

Part (c):

The key question is which variable has the strongest association with joint and hip health. The goal of blocking is to create groups of homogeneous experimental units. It is reasonable to assume that most clinics will see all kinds and breeds of dogs so there is no reason to suspect that joint and hip health will be strongly associated with a clinic. On the other hand, different breeds of dogs tend to come in different sizes. The size of a dog is associated with joint and hip health, so it would be better to form homogeneous groups of dogs by blocking on breed.

3. (2001 #4)

Part (a):

Blocking scheme A is preferable because it creates homogeneous blocks with respect to forest exposure. That is, plots in the same block have similar exposure to the forest.

Part (b):

Randomization of varieties of trees to the plots within each block should reduce any possible bias due to confounding variables, such as fertility or moisture, on the productivity of the two types of dwarf trees.

OR

Randomization of varieties of trees to the plots within each block should even out (or equalize) the effect of other characteristics of the plots that might be related to the productivity of the trees.

4. (1999 #3)

a. The student can appeal to any of three reasons in judging this study not an experiment:

   1. there is no random assignment of subjects to treatments;
   2. there are no treatments imposed;
   3. existing data is being used.

b. Two variables are confounded if their effect on the number of new cavities cannot be distinguished from one another. The student must mention not only that the confounding variables may affect the outcome but that they have differential effects within the two groups. For instance: confounding would occur if patients who eat an apple a day differ from those who eat less than one apple a week **on some variable that is related to dental health**. In this example, diet or general level of health are examples of what might be confounding variables. For example, it is possible that people who eat an apple a day are more nutrition conscious and have a more healthy diet in general than those who eat one or fewer apples per week, and this might explain the observed difference in dental health.

Note:
There are many possible examples of confounding variables. Any reasonable example of a confounding variable is acceptable, as long as a good explanation is given and the connection between the confounding variable and group membership is clear. Lack of a definition here can be rectified by a response in (c) that demonstrates a clear understanding of the concept of confounding variable.

c. No, because it is not an experiment, and cause-and-effect conclusions cannot be drawn from an observational study.

   OR

   No, because there are possible confounding variables.

Notes:

1. In b), a good definition of confounding with a bad example should be regarded as temporarily weak. An example that does not mention group affiliation can be recovered in part c). To recover the definition of confounding in c) they must connect the term with the definition. To recover the group affiliation, they may do by example in c).

2. If the student, in attempting to discuss group differentiation, only mentions one of the groups, that is OK - we will consider the other implied. For example, it is counted correct if the student says, "The apple-eating group may be more health-conscious…" they need not explicitly deny health-consciousness to the one-apple-a-week group.

3. The constructions "Some people" and "A person in the apple-eating group may …" are NOT enough to establish group differentiation; this construction suggests only that some subset of the group may differ from the rest of the group. This is just natural variation. The construction "A person who is an apple-eater may …" can establish group differentiation if it is clear that this is describing a representative member of the group.

4. Mentioning initial non-equivalence of groups without tying that non-equivalence to the outcome is not correct. Mentioning concepts such as self-reporting bias, social desirability, etc. may constitute measurement error in the study but is not confounding.

5. In (c), appealing to the definition of confounding variables in (b) would get a "correct" for (c) if the definition in (b) is correct. If the definition in (b) is weak, that appeal alone would not get credit.

6. If the definition in (b) is "there are other variables that affect the outcome measure, such as age, health, etc," this is **not** regarded as a correct definition for purposes of appealing to the definition from (c). For purposes of reading part (b), this definition would be regarded as weak.

7. If in (c), they give an example which is the equivalent of confounding, and **refer to this as confounding**, they would get credit for (b).

5. (2006 #5)

Part (a):

The three different growth-enhancing nutrients (A, B, and C) and two different salinity levels (low and high) yield a total of $3 \times 2 = 6$ different treatment combinations for this experiment.

Treatment Combination	Nutrient	Salinity Level
1	A	Low
2	A	High
3	B	Low
4	B	High
5	C	Low
6	C	High

Part (b):

Since 10 tiger shrimps have already been randomly placed into each of 12 similar tanks in a controlled environment, we must randomly assign the treatment combinations to the tanks. Each treatment combination will be randomly assigned to 2 of the 12 tanks. One way to do this is to generate a random number for each tank. The treatment combinations are then assigned by sorting the random numbers from smallest to largest.

Treatment Combination	Nutrient	Salinity Level	Tanks with
1	A	Low	Smallest and second smallest random numbers
2	A	High	Third and fourth smallest random numbers
3	B	Low	Fifth and sixth smallest random numbers
4	B	High	Seventh and eighth smallest random numbers
5	C	Low	Ninth and tenth smallest random numbers
6	C	High	Next to largest and largest random numbers

After three weeks the weight gain (after – before) is computed for each tank, and the treatments are compared using appropriate averages.

Part (c):

Using only tiger shrimp will reduce a source of variation in the experimental units, the tanks of shrimp in this experiment. By eliminating this possible source of variation, type of shrimp, we are better able to isolate the variability due to the factors of interest to us (nutrient and salinity level). This will make it easier to identify any treatment effects that may be present.

Notes:

- In this completely randomized design, confounding is not possible. Therefore a reference to confounding or lurking variables always incurs a penalty.

Part (d):

Using only tiger shrimp will limit the scope of inference for the biologist. Ideally, the biologist would like to identify the treatment combination that leads to the most growth for all shrimp. However, the biologist will only be able to identify the best treatment combination for tiger shrimp because other types of shrimp may respond differently to the treatments.

6. (2002 #2)

Part (a):

A **paired design** is used in which each subject receives a pair of boots where one boot is treated with the new method and the other with the current method.

Subjects should be randomly assigned to one of two groups. Group 1 would have the new method applied to the right boot; group 2 would have the new method applied to the left boot.

OR

For each subject, whether the new method is applied to the right or left boot is determined at random.

OR

A **crossover design** is used in which each subject receives a pair of boots, both of which were treated with one treatment. The boots are used for three months and then exchanged for a second pair of boots, both of which were treated with the other treatment. These boots are then used for the next three months.

Subjects should be randomly assigned to one of two groups. One group receives boots with the new treatment first and the other group receives boots with the current method first.

NOTE: Additional appropriate blocking schemes are considered extraneous.

Part (b):

The design could be double blind, as long as both the *subjects* and the person *evaluating* the boots for water damage do not know which boots were treated with the new method and which were treated with the current method.

NOTE: If the student does something unexpected in part (a) and gives a design that actually cannot be double blind, then part (b) could be considered correct provided the response explains why the design could not be double blind.

7. 2006B, #5

Part (a):

The response variable was the amount of draft. The two treatments were the standard hitch and the new hitch. The experimental units were the two large plots of land.

Part (b):

Yes, the two hitches (treatments) were randomly assigned to the two plots (experimental units).

Part (c):

No, each treatment (type of hitch) was applied to only one experimental unit (plot of land). Replication is used to repeat the treatments on different experimental units so general patterns can be observed. There is no replication in this study.

Part (d):

Although 25 measurements were taken at different locations in the two plots, each hitch was used in one plot (experimental unit) only. Thus, if a difference in the draft is observed we will not know whether the difference is due to the hitch or the plot. In statistical language, the treatments (hitches) are confounded with the plots.