Regarding Confounding

Tiger Shrimp Problem (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.

From the scoring guide:

Part (c) is essentially correct (E) if

- the statistical advantage of reduced variability is identified
  AND
- an appropriate explanation that relates reduced variability to increasing the likelihood of determining differences among treatments is clearly provided.

Part (c) is partially correct (P) if only one of the two components is correct.

Part (c) is incorrect (I) if neither of the two components is present.

Notes:

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

From the Chief Reader's Report:

Part (c):

- The most common error was the improper use of "confounding variable" or "lurking variable." When students identified the advantage of reduced variability, they often did not express why this was an advantage.

Based on your experience of student responses at the AP Reading, what message would you like to send to teachers that might help them to improve the performance of their students on the exam?

When answering statistical design questions, students should not use terminology that is not appropriate for the situation. For example, in this question many students referred to 'confounding'; however, in a completely randomized design, confounding is not possible. While students seemed to know that some type of randomization was necessary, the level of their understanding of that process was often very minimal. They either were unable to provide additional details or described randomizations of treatments that were incorrect or not possible. Helping students to understand exactly what the treatments are in an experiment and the reasons (and details) of the related randomization(s) may help them to improve their performance on design questions.

Friends,

I'd like to speak to a couple of the issues that have been hot on the list in these days on the cusp of two years.

First, the 2006 shrimps problem, part c, asks about the advantage of using only one species of shrimps, tiger shrimps. The correct answer is that it reduces within-treatment variability, since shrimp from the same species are likely to respond more uniformly than shrimps from a mixture of species. This thereby makes it easier to determine the differences in effects of the treatments of nutrient and salinity. This within-treatment variability is _not_ confounding, and responses that used this term were penalized for mixing up two different concepts of experimental design. Students who just mixed up these ideas could not get full credit.

Other students wrote something like "well, if they put all the tiger shrimps in one treatment and all the other species in another treatment, there would be confounding." This is a true statement about confounding, but the problem stem said that shrimps were randomly assigned to tanks, so this is hypothesizing a situation that the stem of the problem rules out. This is again incorrect as an answer to the question posed, even though it includes a correct description of confounding.

Confounding occurs when there is a _systematic_ way in which one treatment is favored. This is nicely defined by George Cobb in his Design and Analysis of Experiments when he says "Two influences on the response are confounded if the design makes it impossible to isolate the effects of one from the effects of the other." Note that in an experimental situation, confounding has to involve the _design_. There are more complicated experimental design situation in which certain factors or interactions are deliberately confounded to put greater power on the main effects. You can read about this in the classic Statistics for Experimenters by Box, Hunter, and Hunter. Thus, in an experiment, it's a design issue. With a completely randomized design, this is not possible because the randomization precludes a designed-in way in which one outcome would be favored. This is the root of the statement from the problem rubric that Mark quoted in a previous post that confounding is not possible in this situation.

Lurking variables occur most commonly in regression-type situations, in which you are trying to model a response variable with one or more explanatory variables. We all know that old saw that "correlation doesn't imply causation" and the reason is the potential lurking variables which may actually be instrumental in linking the explanatory and response variable. Both Cobb and Box, Hunter, Hunter only talk about lurking in their regression chapters.

In my opinion, confounding and lurking variables are primarily in the domain of observational studies, at least at the AP Stat level, and are best confined to that domain. In that domain I'm not sure it's critically important to distinguish the two, and I can't recall an AP problem that has ever asked students to distinguish the two.

A couple years ago I wrote an article for STATS magazine about the many different kinds of variables. If you'd like a copy, send me a private note.

Happy New Year to all!

Peter Flanagan-Hyde
Phoenix Country Day School
Paradise Valley, AZ

Hi Nancy,

This issue has arisen before -- think tiger shrimp -- and there are a couple of perspectives. Here's mine:

Confounding refers to a systematic way in which one outcome is favored, and is the bane of observational studies. A confounding variable is one which affects the response, but is not controlled for.

In an experiment, you can control for the effects of variables which otherwise might be confounding variables. This is the idea behind the statement in the rubric for the tiger shrimp problem of 2005 that confounding isn't possible in a completely randomized experiment.

The most basic way to control for potentially confounding variables is randomization, which removes the systematic way in which one outcome is favored. But if there are different-looking subjects in your experimental group that you believe will produce different responses, then just randomizing may leave more of one in treatment A than in treatment B, or vice versa. This possibility increases the random within-group variability.

You can reduce this random variability by blocking on the observable characteristic. Then you guarantee that the two treatment groups are equally balanced with regard to this variable. This generally works better than just randomizing, hence the statement that one aim of blocking is to reduce variability in an experiment. It's specifically the within-group variability that's reduced, making it easier to see the between-group variability -- which is what your experiment is trying to measure.

Peter

Peter Flanagan-Hyde
Phoenix Country Day School
Paradise Valley, AZ

Dentists Problem: 1999 #3:

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.

Solution to 1999 #3 (Dentists):

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).