Epidemiology as a Context for Teaching Confounding in Elementary Statistics

Chris Olsen
Mathematics/Assessment Facilitator
Cedar Rapids Community Schools, Cedar Rapids, IA

Abstract

Designing and critically interpreting studies are major topics in the AP statistics course. Confounding is a difficult concept at best, and epidemiological studies are replete with confounding concerns. It would seem epi examples could provide an interesting and important source of context that not only would help elementary statistics students understand confounding, but provide concepts that may help them evaluate individual and public health risks.

1. Introduction

Before I begin I must acknowledge my comparatively primitive expertise in the field of statistics. As they say in Texas, I wasn't born here, but got here as soon as I could. I am a math teacher who fell in love with statistics as soon as I could in grad school, and am still learning.

My soon-to-follow prattling about confounding is the result of three related interests: teaching the Advanced Placement Statistics high school course for 8 years, I was involved in writing curricular materials of a statistical nature as part of a set of modules, and am now involved in educational statistics as an "Assessment specialist."

My soapbox today is concerned with the treatment of confounding in the standard first semester non-calculus based college course. This course is the first, and frequently the last formal college experience with statistics.

The AP Statistics course designed to parallel this college offering contains the following in its syllabus:

Subset of the AP Statistics Syllabus:
Planning a Study

II. Sampling and Experimentation: Planning and conducting a Study

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

- p. 9

In a course such as this – or, at least in textbooks for such a course – the planning of studies chapter(s) devote the lion's share of the discussion to planning experiments, the epistemological "gold standard."

The statistical analysis techniques taught in this course are typically those common to experimental studies – randomization, inferences about proportions and means, and perhaps regression and chi square.

Surveys and observational studies are in practice given subdued treatment, perhaps because the statistics are considered too easy (surveys) or perhaps too difficult (observational studies) to present in the elementary survey course.

Beyond defining terms, little about observational studies is discussed.

2. More Attention to Observational Studies

My general view is that observational studies should be given more attention, for the following reasons:

More attention to Observational studies because:

- Statistics is the last chance that we have to develop a formal understanding of reasoning about chance behavior
- As citizens, students, professionals, and constant decision-makers in our everyday lives, we must cognitively confront and reason about chance events
- The chance events about which we must reason are rarely mirrored by the canonical "experiment" – most confrontations with chance are of an "observational" nature
- A narrow focus on experiments, with the accompanying *pro forma* mantra about randomly sampled units randomly assigned to treatments, allows the avoidance of what I believe are essential discussions of the empirical basis for everyday inference

Since human analysis and decision-making is tied to the processes of understanding and explaining, surely the most needed discussion about everyday inference is about what is probably the greatest source of inferential error – the confusion of correlation with causation, as well as a general ignorance of the concept of confounding. In my own field, I have observed the following confusion by educators who ought to know better:

- The best predictors of academic success are wealth and mother's educational level – damn, how can we change that?
- If you take the AP tests, you will do better in college
- NCLB – despite everybody from the Supreme Court to local PTAs putting pressure on schools and curriculum, we can identify the schools as the one true explanatory variable for student achievement

This confusion of correlation and causation has the following elements:

- first, the deification of a single observed correlation
- second, the absolution of other known correlations,

  and
- third, the denial of the existence of any unknown correlations.

The disentangling of a constellation of correlated variables to establish a reasonable causal chain is, in real life, a non-trivial problem.

In textbook experimental studies this problem of confounding is washed away with the flick of the random assignment wand. In observational studies problems of potential confounding must be explicitly identified and confronted as a matter of course.

Probably no discipline has refined this identification and confrontation methodology better than epidemiology. David Fraser, an accomplished epidemiologist, writing in the New England Journal of Medicine, defended the study of epidemiology as a liberal art; a significant part of his argument also applies to the study of statistics.

In judging the suitability of a discipline for undergraduate study, one should look for the essential characteristics of the liberal arts, which I take to be the fields that help free students from the limitations of prior beliefs and experiences and that teach important modes of thinking so as to prepare them to ask and answer new questions.

Five approaches to problems or modes of thinking stand out as particularly important, and although not all may be used in a particular discipline, students should seek to become competent in each during the course of liberal arts study.

The five, in no particular order, are the scientific method, analogic thinking, deductive reasoning, problem solving within constraints, and concern for aesthetic values.

-- David Fraser

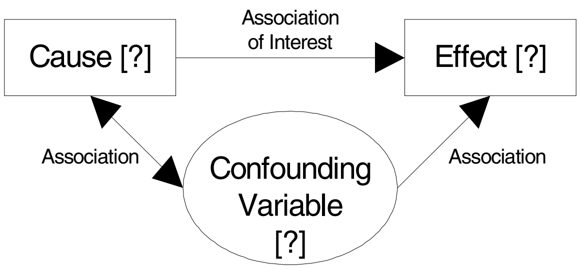
NEJM

3. Statistics as a Liberal Art

I would like to nominate the first statistics course as a candidate meeting the "liberal arts" test, and further suggest that a great public service could be performed by alleviating the confusion about correlation, confounding, and causation.

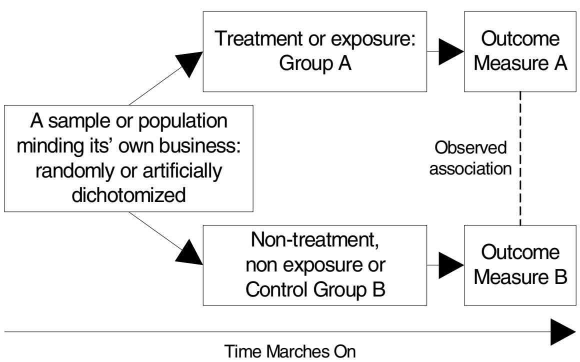
The concept of confounding is usually presented in textbooks with something like the following diagram:

Figure 1. Triangle Diagram



After even a careful study of this diagram one might opine (in agreement with General Custer at the Little Big Horn): this diagram has entirely too many arrows, going in entirely too many directions, to make much sense. It is frequently not clear whether the arrows indicate association, a causal relation, an "arrow of time," or some combination of these. My key frustration with the diagram, however, is that it suggests that assessing confounding is a simple matter of measuring associations and correlations. It seems to me that assessing confounding is a logical task that is better placed as an issue of experimental design. And, as is not-well-enough known, the existence of an association is not completely probative.

Figure 2. Experimental Design Diagram



A diagram something like Figure 2 suggests the association we are looking for and in addition points the student to focus on the study design rather than searching for variables that might be associated. The logic of establishing a causal link between exposure/treatment and outcome is also supported – though not expressed completely – in the diagram in Figure 2. (The arrow of time is meant to express the order of the existence, rather than measurement, of the attributes in question – in a case-control study the data might be gathered at the "same" point in time, but the explanatory attribute predates the response.)

1. We somehow create two populations or samples that are more or less equivalent, so that no characteristic (i.e. variable) can be associated with the outcome variable in advance.

2. We prepare for an association by executing a treatment and a non-treatment or by classifying subjects as exposed or unexposed and then finding that our measure on the outcome of interest differs in the two populations or samples.

3. We measure the outcome to see if there is an association between exposure/treatment and outcome. (If so, there might be a case to be made for causation)

4. We note in passing the establishment of the appropriate time sequence,

and

5. We anticipate the successful explication of a plausible mechanism before, or possibly after, the study

There are, of course, two clouds on our horizon. First, the treatment/exposure groups might differ prior to treatment/exposure, on a variable that plausibly accounts for the difference in outcome measure. Second, some different event might occur between the groups during the treatment, a difference plausibly related to the difference in outcomes. In either case, we have a problem of potential confounding.

Treating the problem of assessing confounding as an experimental design problem thus evolves into a task of identifying and evaluating the possible effects of variables other than the treatment/exposure variable.

The student is not permitted to devalue a study simply because there are identified or suspected associations between the outcome and some other variables – they must consider the relevance and import of those associations in the context of the study design.

In an elementary statistics class it might be argued that the "context" of a study design is frightening to students who may not have a great deal of background in the topic of the study. This is especially problematic when one considers that the "real-life" context of experimentation is – content-wise – typically beyond the level of scientific expertise assumed in typical survey courses. The student may be stretched to the conceptual limit once he or she has an understanding of the explanatory and response variables and why they might be related in the experiment under discussion – searching for possible confounding variables may be asking too much of some students.

4. Teaching Epidemiological Studies

Given that, there is much to recommend epidemiological studies as fruitful examples for use in an elementary statistics class:

- The treatment/exposure and outcome measures are understandable to students without imposing a significant science background
- The timelines between "cause" and "effect" are clear
- The mechanisms are arguably simple
- Potential confounders are fairly easily identified and understood, due to student familiarity with the variables in the study
- Relevance to the students!

Here are some examples from recent epidemiological studies that could be used when discussing and analyzing possible confounding:

- Risk of overweight among adolescents who were breastfed as infants.
- Time of birth and the risk of neonatal death.
- Vaccination and allergic disease: A birth cohort study.
- Onset of adolescent eating disorders: population based cohort study over 3 years.
- Does parental disapproval of smoking prevent adolescents from becoming established smokers?
- Television and adolescent use of over-the-counter analgesic agents

5. Bottom Line

So, here are our bottom lines:

1. Understanding correlation, confounding, and causation is important to intellectual and cognitive development

2. Elementary statistics can (and should) play an important role in this development

3. Epidemiology is a fruitful source of good and useful and important and relevant examples.

Thank you for your kind attention.

References

College Board. AP Statistics: Course Description. Go to http://apcentral.collegeboard.com/repository/statistics_cd_0506_4328.pdf

College Board. The Young Epidemiology Scholars Program. Sponsored by the Robert Wood Johnson Foundation. www.collegeboard.com/yes

Easterling, R. G. (2004), Teaching experimental design. *The American Statistician*, 58, 244-52.

Fraser, D. W. (1987). Epidemiology as a liberal art. *New England Journal of Medicine*, 316, 309-314.

Gillman, M. W., et al. (2001). Risk of overweight among adolescents who were breastfed as infants. *JAMA*, 285, 2461-2467.

Gould, J. B., Qin, C., & Chavez, G. (2005). Time of birth and the risk of neonatal death. *Obstetrics & Gynecology*, 106, 352-358.

McKeever, T. M., et al. (2004). Vaccination and allergic disease: A birth cohort study. *American Journal of Public Health*, 94, 985-9.

Patton, G. C., et al. (1999). Onset of adolescent eating disorders: population based cohort study over 3 years. *BMJ*, 318,765-8.

Sargent, J. D., & Dalton, M. (2001). Does parental disapproval of smoking prevent adolescents from becoming established smokers? *Pediatrics*, 108, 1256-1261.

Stroup, D. F., et. al. (2004), Teaching Statistical Principles Using Epidemiology: Measuring the Health of Populations. *The American Statistician*, 58, 77-83.

Van den Bulck, J., Leemans, L., & Laekeman, G. M. (2005). Television and adolescent use of over-the-counter analgesic agents. *The Annals of Pharmacotherapy*, 39, 58-62.