Chapter 5, Experiments and Observational Studies

Sanford Weisberg

Univ. of Minnesota

February 9, 2009

Beer Tasting

Three beers are to be compared in a taste test to find the preferred beer.
The research design:

1. Recruit subjects then:
   1. Each subject tastes all three beers in random order
   2. we get three times as many subjects and each tastes just one beer
   3. Each subject does three trials, first comparing A to B, then B to C and A to C.
2. The experiment should be
   1. **double blind**
   2. **single blind**
   3. **blinding is unimportant**

More beer...

1. Order of presentation is
   1. **randomized**
   2. **always fixed**
   3. **something else**
2. Is there a control in this experiment? Is the control a placebo?
3. The response should be
   1. Rating (1 to 7, say) of each beer.
   2. Rank order of the three beers, e.g, $C \succ B \succ A$ (C is preferred to B which is preferred to A).
   3. Pick your favorite.
   4. If two at a time, pick the winner each time.

Q-Comp

- Q-Comp is a Minnesota state mandated program that provides incentive pay for teachers, with the eventual goal of increasing learning of students.
- School districts will enroll in the program, provide mentoring, additional training and other tools for teachers to apply in the classroom.
- "Participating school districts and charter schools are eligible to receive additional revenue of up to $260 per student. In fiscal year 2009, Q Comp participants will receive an estimated $49 million in state Q Comp funds."

`http://www.auditor.leg.state.mn.us/ped/2009/qcomp.htm`.

Experiment? Observational Study?

1. How could rolling out Q-Comp be done as an experiment? What is the **experimental unit** (EU)?
2. How could it be done as an observational study?
3. Could this be done as a **blocked design**, where Period 1, the past, had no Q-Comp and Period 2, the present, has Q-Comp? What is the problem? New technical term: **confounder**. Another technical term: **quasi-experiment**.
4. Are there other explanatory variables that can account for differences between Q-Comp and non Q-Comp schools (more important for observational studies than for experiments, why).

Results of Q-comp

The results of Q-Comp, according to the Legislative Auditor:

- Too little time: "Q Comp has only been in place for three full school years."
- No randomization "Q CompŠs voluntary nature also makes it difficult to draw conclusions about the effectiveness of the program."
- Confounding and Bias: "Further, it is difficult to disentangle the effects of Q Comp from other initiatives. Q Comp is one of many programs aimed at improving student achievement."

Berkeley Admissions

Are women and men treated equally in admission to graduate programs? At UC Berkeley in 1973, the data, from P. J. Bickel, E. A. Hammel and J. W. O'Connell, *Science*, 187, (Feb. 7, 1975), pp. 398-404, were:

Applicants	# Admitted	# Applied	Admission rate
Men	3738	8442	44.2%
Women	1494	4321	34.5%

The admission rate for men is 44.2/34.5=1.28 times the rate for women. Does this prove that women are discriminated against in admissions? Is this an experiment? Why or why not?

Simplified Example

Department of machismatics

Applicants	# Admitted	# Applied	Admission rate
Men	300	400	75%
Women	75	100	75%

Dept. of social warfare

Applicants	# Admitted	# Applied	Admission rate
Men	50	150	33%
Women	150	450	33%

Add together the two departments

Applicants	# Admitted	# Applied	Admission rate
Men	350	550	64%
Women	225	550	41%

Huh? **Simpson's Paradox**.

Highway Accidents

Do lower speed limits lead to fewer accidents? For 39 segments of major MN highways, here are the observed accident rates per million vehicle miles in 1973 (from S. Weisberg (2005), *Applied Linear Regression*, Third Ed, Wiley):

Speed limit	Accidents per MVM
≥ 60 mph	2.6
≤ 55 mph	4.6

1. Is this an experiment or an observational study?
2. What is treatment and what is control?
3. Is there any blinding?
4. What is the outcome?
5. Can we infer that *higher speed limits cause fewer accidents*.

Two more...

- Does smoking cause cancer? How could you do experiments to learn about this issue?
- Does the amount of homework assigned in a class effect how well students do in the course? How could you design an experiment for this? What is the experimental unit? What are possible confounders?