Good and Bad Ways to Design Experiments

Example: **Aspirin and Heart Attacks**

- $22,000$ **male physicians were randomly divided into two groups.**
- **One group took an aspirin every other day**
- **Other group took a placebo every other day**
- **After several years, total number of heart attacks was measured for each group**
- **Group taking aspirin had significantly fewer heart attacks**

Things to Think About

- **Why was the assignment into groups *random*?**
- **Why did the second group need to take a *placebo*?**
- **What does "*significantly fewer*" heart attacks mean?**

Concepts in Experiments

- ***Response Variable*: Measures main outcome of interest in a study**
- ***Explanatory Variable*: A variable that we think may *explain* the pattern of variation in the response variable**
- ***Subjects* (or *units*): The individuals in an experiment on whom (or on which) we measure the response**
- ***Treatment*: A specific experimental condition applied to the subjects. Could be a value of the explanatory variable (or a combination of values of several explanatory variables)**

Online Learning Example

- **Study done at Nova Southeastern University in Florida**
- ***Subjects:* Sample of college students at Nova SE**
- ***Response:* Score on end-of-course test**
- ***Explanatory Variable*: Setting of class (in-person or online)**
- ***Goal*: Compare average response for the two treatments (in-person instruction and online instruction)**

Online Learning Example (continued)

- **Study found students enrolling online performed "as well" on end-of-class test as students taking in-person classes.**
- ***Problem:*** **Students were *not* assigned to one of the two treatments *at random*.**
- **Beware of *lurking variables*: Variables *not measured in the study* that affect the relationship between the response and the explanatory variables.**

- **The *lurking variable* in the Nova study was: "previous ability" on course material.**
- **Turns out the students who signed up for online class were better on the material *initially*.**
- **The *"previous ability"* lurking variable and the "class setting" variable were *confounded*: It was impossible to separate the effect on the response *of one variable* from the effect on the response *of the other variable*.**
- **In general: Two explanatory variables can be *confounded*, or an explanatory variable can be *confounded* with a lurking variable.**

Clicker Quiz 1

A study found a sizable relationship between air-conditioning bills and lemonade purchases for a sample of households. What is a likely lurking variable in this situation?

A. Household annual income

B. Household size

C. Outdoor air temperature

D. City population

Randomized Comparative Experiments

- **A *randomized comparative experiment* will minimize the effect of lurking variables.**
- **Often we compare performance for two (or several) competing treatments.**
- **Sometimes we compare one *treatment* under consideration to a *control*.**
- **Control group in a drug study might get a placebo (fake treatment), or some previous standard drug.**
- **Major concept: Subjects are assigned to each group *at random*. This way no lurking variable can *systematically favor* one group over another.**

- ***Double-blind experiment*: Neither the experimenter nor the subject know which treatment the subject is receiving.**
- **Otherwise, for example: Doctor might be tempted to assign sickest patients to treatment group and healthier patients to placebo group.**

Three Major Principles of Experimental Design

- ***Control*: Limit effects of lurking variables by making groups as equal as possible (except for the fact that they get different treatments).**
- ***Randomization*: Use chance to assign subjects to treatments, to eliminate bias.**
- ***Replication*: Use lots of subjects, to reduce variation due to chance in the results.**

Statistical Significance

- **If the difference in average response among treatment groups is *large enough* that such a difference is *unlikely to arise from chance*, then this is called a *statistically significant* difference.**

- **Whether a difference is *statistically significant* depends on:**

  1. **the actual size of the difference**

  2. **the number of subjects in the study**

- **If the sample size is *large*, a difference among treatments may be *statistically significant* but not really *practically significant*.**

Clicker Quiz 2

What result of a designed experiment is most likely to be statistically significant?

A. **Average response for treatment group = 14.5, Average response for control group = 10.5 (based on 10 subjects per group)**

B. **Average response for treatment group = 10.6, Average response for control group = 10.5 (based on 10 subjects per group)**

C. **Equally likely, because they are based on the same sample size.**

Clicker Quiz 3

What result of a designed experiment is most likely to be statistically significant?

A. **Average response for treatment group = 10.6, Average response for control group = 10.5 (based on 100 subjects per group)**

B. **Average response for treatment group = 10.6, Average response for control group = 10.5 (based on 10 subjects per group)**

C. **Equally likely, because they show the same difference in average response.**

Making Conclusions from Observational Studies

- ***Example*: Do people who regularly attend church services live longer on average?**
- **A randomized comparative experiment is the best way to answer such a question . . . However:**
- **We can't randomly assign certain people to attend church and assign others to not attend.**
- **We must use an *observational study* instead.**

Making Conclusions from Observational Studies (Continued)

- **An observational study should still be *comparative*.**
- **We examine a random sample of churchgoers and a random sample of non-churchgoers and observe some response variable (like life length) for each.**
- **A good idea is to use *matching*: Compare groups that are similar in terms of potential *lurking variables* (age, gender, education, etc.)**
- **This way these lurking variables won't affect the comparison of interest too much.**

Making Conclusions from Observational Studies (Continued more)

- **Another approach is to directly *measure (and adjust for)* potential confounding variables in our analysis.**
- **We can measure smoking habits, weight, physical activity, etc., for both the churchgoers and non-churchgoers.**
- **Use statistical methods to adjust for (account for) these other variables.**
- **Then, if we still find a significant difference in average life length for churchgoers and non-churchgoers, it's *probably* because of the effect of attending religious services as opposed to the physical health factors.**

- ***Warning*: We still have to consider the possibility of another lurking variable that we didn't think of or adjust for.**
- **The conclusions from a well-designed experiment will usually be more certain than those from an observational study.**