Chapter 13: Experiments and Observational Studies

Designing Experiments:

If we want to observe individuals and record data without any intervention, we conduct an observational study. Observational studies can look at data that has already been collected (called a retrospective study) or it can look at data as it occurs (called a prospective study). Observational studies are valuable for discovering trends and possible relationships, but they cannot demonstrate causal relationships.

> If we want to examine a cause-and-effect relationship, we must conduct a controlled experiment rather than an observational study.

The individuals on which the experiment is done are called experimental units. If the units are people, they are called subjects.

The experimental condition we apply to the units is called the treatment. The explanatory variables (causing a change in the other variables) are called factors. These factors may be applied in different levels.

When designing an experiment we want to minimize the effect of lurking variables so that our results are not biased. Because we may not be able to identify and eliminate all lurking variables, it is essential that we use a control group. The control group gets either no treatment, or a fake treatment to counter the placebo effect and/or any other lurking variables present. Having a control group allows us to compare the results of the treatments.

<u>Experimental Design</u>

Step 1: Choose treatments
- identify factors and levels
- control group

Step 2: Assign the experimental units to the treatments
- matching (similar units in each group)
- randomization

Remember, if we want to examine a cause-and-effect relationship, we conduct an experiment. If an experiment is well-designed, a strong association in the data does imply causation, since any possible lurking variables are controlled.

Principles of Experimental Design:

1. Control the effects of lurking variables by comparing several treatments (include a control group if possible/applicable).
2. Use Randomization to assign subjects/units to treatments. Without randomization, we do not have a valid experiment and will not be able to draw conclusions from your study.
3. Replicate the experiment on many subjects/units to reduce chance variation in the results. If the experimental group is not representative of the population of interest, we may need to replicate the entire experiment for different groups in different places and/or at different times.

An effect is called statistically significant if it is too great to be caused simply by chance.

Even a well-designed experiment can contain hidden bias, so it is extremely important to handle the subjects/units in exactly the same way. One way to avoid hidden bias is to conduct a blind experiment. In a double-blind experiment, neither the subjects nor the people who have contact with them know which treatment a subject has received.

Types of Experimental Design:

1. In a completely randomized design all subjects are randomly assigned to treatment groups.
2. In a block design, subjects are first split into groups called blocks. Subjects within each block have some common characteristic (*for example*: gender, age, education, ethnicity, etc.) Then, within each block, subjects are randomly assigned to treatment groups.
3. In a matched pairs design, there are only two treatments. In each block, there is either: a single subject receiving both treatments or a pair of subjects each receiving a different treatment.

Lurking and Confounding Variables:

A lurking variable creates an association between two other variables that tempts us to think that one may cause the other. A lurking variable is usually thought of as a prior cause of both *y* and *x* that makes it appear that *x* may be causing *y*.

Confounding can arise in experiments when some other variables associated with a factor have an effect on the response variable. Unlike a lurking variable, a confounding variable is associated in a noncausal way with a factor and affects the response. Because of the confounding, we find that we can't tell whether any effect we see was caused by our factor or by the confounding factor (or by both working together).