

Statistical Principles of Design

Bret Larget

Departments of Botany and of Statistics
University of Wisconsin—Madison
Statistics 371

1st November 2005

- ▶ We make a distinction between an **experiment** in which the researchers intervene in the experimental conditions and an **observational study**, in which the researchers merely observe an existing situation.
- ▶ The distinction is important **in the interpretation** of the results of an analysis.
- ▶ Consider an analysis to compare two groups. In an experiment, **the researchers assign the groups**.
- ▶ In an observational study, **the groups are simply observed**.



Principles of Design

- ▶ Previous chapters have dealt with methods of **analyzing data** to make inferences about a single population or to make comparisons between two populations.
- ▶ Future chapters will focus on methodology for the analysis of data that arises in additional settings.
- ▶ The present chapter looks instead at the problem of **design**, the methods for collecting data.
- ▶ How should data be collected so that analysis of the data leads to valid inferences?
- ▶ What are the pitfalls of poor design choices, and how can these affect inference?
- ▶ What are the main statistical principles of design?



Case Study: Cigarettes and Smoking

- ▶ In a study, pregnant women were questioned about their smoking habits, diet, and other variables. The babies were followed up for some time.
- ▶ There was strong statistical evidence that **the mean birth weight of smokers' babies was lower than the mean birth weight of nonsmokers' babies**.
- ▶ Low birth weight is associated with a number of health problems in babies, which makes the problem of finding causes of low birth weight important.
- ▶ We say that smoking and low birth weight are **associated with one another**.



Case Study: Cigarettes and Smoking

- ▶ However, this single study alone is **insufficient** to support the conclusion that **smoking caused the lower birth weight**.
- ▶ The smokers and nonsmokers differed on a number of possible **explanatory variables**, and it is unclear which of these variables may have caused the difference.
- ▶ We say that the possible effects of smoking on birth weight are **confounded** with many other possible explanations.
- ▶ Confounding has the potential to mislead
- ▶ **Association is not causation!**



More on smoking and low birth weight

- ▶ In one study, a large number of variables were measured. A complex statistical method that simultaneously estimates the effects of several explanatory variables found that even after making adjustments for these other variables, **smoking still had an effect on birth weight**.
- ▶ A second study found differences in the placenta between smokers and nonsmokers, and that some of the differences were **associated with chemicals found in cigarettes**.
- ▶ This same study also found that having smokers not smoke for three hours **caused a change in blood flow to the placenta**.



Comparison

- ▶ To attribute a causal relationship between an **explanatory variable** and a **response variable** (such as smoking and low birth weight), we would like to be able to make a comparison between two groups that differ only in the explanatory variable under study with all other possible explanatory variables the same between the two groups.
- ▶ While in experimental settings it is possible for the researcher to create groups in which a single explanatory variable is the largest difference between two groups, there are many settings for which experiments are either impossible, too expensive, or unethical.
- ▶ It is not impossible to attribute a causal relationship based on observation studies alone, but it is far more difficult because the researchers essentially need to identify and rule out or control for the effects of the other possible explanatory variables.



Example (cont.)

- ▶ A third study identified 159 women who smoked during a first pregnancy but not during a second pregnancy.
- ▶ These women were matched with 159 other women who had smoked during both pregnancies and for whom other explanatory variables were similar.
- ▶ This study found that **the second babies of the women who quit smoking were heavier than the second babies of their matched controls** who continued to smoke.



The First Study

- ▶ The first study attempts to make comparisons of similar groups by **a statistical analysis that attempts to adjust for the effects of other explanatory variables** and so leave a comparison where the only important difference is the explanatory variable of interest.
- ▶ Interpretation of causality from such a study, however, assumes that the statistical model for the joint effects of all the variables is an accurate description of reality, often a dubious assumption.



The Third Study

- ▶ The third study attempts to make a comparison between like groups by constructing groups that are as similar as possible on the basis of other explanatory variables that are thought to also have an effect.



The Second Study

- ▶ The second study attempted to establish a link between smoking and low birth weight by establishing a link between smoking and the placenta.
- ▶ The link between the placenta and birth weight is perfectly plausible without statistical justification on biological grounds alone.
- ▶ This study even included an experiment in which the blood flow to the placenta could be compared within the same woman after she had smoked and when she had abstained from smoking.



Comments on Confounding

- ▶ Notice that **there are several possible ways to address confounding** with the objective to eventually establish a causal relationship through a series of observational studies.
- ▶ Several large observational studies are often necessary to present a convincing case for establishing causality.



Example: A Common Cold

- ▶ Researchers invited college students to volunteer in an experiment to test the effectiveness of a vaccine for preventing the common cold.
- ▶ The volunteers were randomly assigned to two groups. One group received the vaccine, the other group took a placebo.
- ▶ The study was **blinded**. The subjects did not know what group they were in.
- ▶ Both groups reported dramatic decreases in the number of colds.

Group	n	mean number of colds	
		previous year	current year
Vaccine	201	5.6	1.7

Importance of Control Groups

- ▶ This experiment shows the importance of a **control group**. Without a control group, it may have been thought that the vaccine was effective.
- ▶ The real reason for the difference in both groups is most likely in the quality of measurement.
- ▶ Previous year cold counts were based on the students' memories.
- ▶ Current year colds were measured by how often the subjects went to the health center.

Example: A Common Cold

- ▶ Researchers invited college students to volunteer in an experiment to test the effectiveness of a vaccine for preventing the common cold.
- ▶ The volunteers were randomly assigned to two groups. One group received the vaccine, the other group took a placebo.
- ▶ The study was **blinded**. The subjects did not know what group they were in.
- ▶ Both groups reported dramatic decreases in the number of colds.

Group	n	mean number of colds	
		previous year	current year
Vaccine	201	5.6	1.7
Placebo	203	5.2	1.6

Randomization

- ▶ In an experiment the researcher has the opportunity **to assign treatment groups**.
- ▶ There are substantial advantages to **randomization**.
- ▶ If the groups are randomly determined, then any explanatory variables that might be confounded with the treatment will tend to be equally balanced in the two groups so that whatever effects they might have will cancel out.
- ▶ (Note that this tendency is only true on average. It is much more likely to be true for a particular randomization when there are **large samples**.)
- ▶ Randomization has this effect for both explanatory variables that are suspected or known and those for which the researcher is completely unaware.
- ▶ The greatest justification for randomization is that **it controls for unsuspected effects**.

Blocking

- ▶ Blocking is an alternative design consideration.
- ▶ In a **randomized block design**, individuals are grouped initially into **blocks** or **strata** based on explanatory variables that are thought to be important, but are not the focus of the study.
- ▶ Treatments are then randomly allocated within each block.
- ▶ The advantage of blocking is that the effects of explanatory variables known to be important can be made essentially to balance, because equal numbers of individuals can be assigned to each treatment within each block.



Comparing blocking and randomization

- ▶ The advantage of randomization is that it helps to control for effects whether their sources are known or not.
- ▶ The advantage to blocking is that it enforces balance for effects known or thought to be important.
- ▶ Blocking is limited to only a few variables — if you keep subdividing based on other variables, it doesn't take long for each subject to be in a unique block.
- ▶ It is almost always a good idea to randomize within blocks.
- ▶ Blocking can be a better choice than **complete randomization** (treating the entire sample as a single block) in terms of yielding more precise estimates or having more powerful tests.
- ▶ This improvement in statistical power depends on the variable being blocked having an association with the response variable and on the blocking resulting in a separation into more homogeneous groups.



Blocking (cont.)

- ▶ In a field experiment, regions of a field might be expected to have different conditions. The field can be divided into large blocks where conditions are more similar and then plots within each block can be randomly assigned the different treatments.
- ▶ In a medical study, we may stratify human subjects on the basis of sex and age (female/young, female/old, male/young, male/old) to ensure that the groups are evenly represented in each treatment.



Replication

- ▶ Replication involves having more than one observation under the same conditions.
- ▶ The importance is that differences within replicates are due to individual differences and not the treatment.
- ▶ This allows for assessment of individual variability.
- ▶ Methods of statistical inference depend on comparing variability in estimated effects due to explanatory variables with estimated effects due to individual variability.
- ▶ Without replication, individual variability cannot be measured and there the methods for inference don't work.
- ▶ The more replication there is, the more accurate the estimation of individual variability is, the more powerful the methods of inference will be. (**Statisticians almost always want more data...**)

