

Chapter 5 § 2

Designing Experiments

Definitions:

Observational study – observes individuals and measures variables of interest but does not attempt to influence the responses

Experiment – deliberately imposes some treatment on individuals in order to observe their responses

Experimental units – the individuals on which the experiment is done

Subjects – when the experimental units are human beings

Treatment – a specific experiment condition applied to the units

Factors – the explanatory variables in an experiment

Placebo – a dummy treatment that can have no physical effect

Control group – the group of patients who received a sham treatment

Randomization – the use of chance to divide experimental units into groups

Completely randomized design – all the experimental units are allocated at random among all the treatments

Statistically significant – an observed effect too large to attribute plausibly to chance

Replication – repeat each treatment on a large enough number of experimental units or subjects to allow the systematic effects of the treatments to be seen.

Double blind experiment – neither the subjects nor the people who have contact with them know which treatment a subject received

Lack of realism – the subjects or treatments or setting of an experiment may not realistically duplicate the conditions we really want to study

Block – a group of experimental units or subjects that are similar in ways that are expected to affect the response to the treatments

Block design – the random assignment of units to treatments is carried out separately within each block

Match pairs design – a common form of blocking for comparing two treatments

A sample survey collects information about a population by selecting and measuring a sample from the population. The goal is a picture of a population, disturbed as little as possible by the act of gathering information. Sample surveys are one type of observational survey.

An observational study, even one based on a sound statistical sample, is a poor way to gauge the effect of an intervention. To see how nature responds to a change, we must actually impose the change. When our goal is to understand cause and effect, experiments are the only source of fully convincing data.

Because the purpose of an experiment is to reveal the response of one variable to changes in other variables, the distinction between explanatory and response variables are essential. The explanatory variables in an experiment are often called factors. Many experiments study the joint effects of several factors. In such an experiment, each treatment, each treatment is formed by combining a specific value (often called a level) of each of the factors.

Experimentation allows us to study the effects of the specific treatments we are interested in. Moreover, we can control the environment of the experimental units to hold constant factors that are of no interest to us.

The ideal case is a laboratory experiment in which we control all outside factors.

Comparative Experiments

The design of an experiment first describes the response variable or variables, the factors (explanatory variables), and the specific treatments.

Laboratory experiments in the sciences and engineering often have a simple design with only one treatment, which is applied to all of the units

Treatment → Observation

Here is how a before-and-after measurements are made

Observation 1 → Treatment → Observation 2

For example, suppose a teacher observes the effects of student performance in two classes. Then he/she implements a teaching strategy (such as the use of the multiple intelligence theory) with one of the class, but continues the other one without change. After a period of time, the teacher then observes the students performance in both classes again.

Experiments should compare treatments rather than attempt to assess a single treatment in isolation. The group of students who did not receive a different teaching strategy is called a control group, because it enables us to control the effects of lurking variables on the outcome. Control of the effects of lurking variables is the first principle of statistical design of experiments. Control of several treatments is the simplest form of control.

Completely randomized experiments

The first step in an experimental design is the choice of treatments, with comparison as the leading principle.

The next step is to say how we will assign the experimental units to the treatments.

Comparison of the effects of several treatments is valid only if we apply all treatments to similar groups of experimental units.

Systematic difference among the groups of experimental units in a comparative experiment are possible source of bias.
How should we allocate the available units or subjects among the treatments?

-matching

-randomization

Experimenters often try to match the treatment groups in a systematic way. (i.e. age, sex, race, etc...)

Randomization is the second major principle of statistical designs of experiments. Completely randomized designs can compare any number of treatments. To do this, assign the experimental units at random to as many groups as there are treatments.

The logic of experimental design

- Randomization produces groups of experimental units that should be similar in all respects before the treatments are applied.
- Comparative design ensures that influences other than the experimental treatments operate equally on all groups
- Therefore, differences in the response variable must be due to the effects of the treatments. That is, the treatments not only are associated with the observed differences in the response but must actually cause them.

The great advantage of randomized comparative experiments is that they can produce data that give good evidence for a cause-and-effect relationship between the explanatory and response variable.

WE KNOW that in general a strong association does not imply causation. A strong association in data from a well-designed experiment does imply causation.

The outcomes of a randomized experiment, like the result of a random sample, do depend on chance. But in both settings, the laws of probability describe how much chance variation is present. “Random” in statistics does not mean “haphazard.”

The presence of chance variation does require us to look more closely at the logic of randomized comparative experiments.

It is the business of statistical inference, using the laws of probability, to say whether the observed difference is too large to occur plausibly as a result of chance alone.

If we observe statistically significant differences among the groups after a comparative randomized experiment. We have good evidence that the treatments actually caused these differences.

-----***** experiments with many subjects are better able to detect differences among the effects of the treatments than similar experiments with fewer subjects.****-----

Three Principles of Statistical Design of Experiments:

1. **Control** of the effects of lurking variables on the response, most simply by comparing several treatments.
2. **Randomization**, the use of impersonal chance to assign subjects to treatments.
3. **Replication** of the experiment on many subjects to reduce chance variation in the results.

Cautions about experimentation

Randomized comparative experiments are the best means of gaining knowledge about the effects of explanatory variables on a response. You should nonetheless examine even experimental evidence with a critical eye.

The way the experiment is conducted may produce hidden bias despite the use of comparison and randomization.

- Unequal conditions introduce bias.
- Therefore the use of a double blind experiment may be needed in this scenario.

The lack of realism is another caution to be aware of in an experiment.

- The subjects or treatments or setting of an experiment may not realistically duplicate the conditions we really want to study.
- Lack of realism can limit our ability to apply the conclusions of an experiment to the settings of greatest interest.

Other Experimental Designs

Completely randomized designs are the simplest statistical designs for experiments.

They are the analog of simple random samples.

In fact, each treatment group is an SRS drawn from the available subjects.

Completely randomized designs illustrate clearly the principles of control, randomization, and replication.

However, just as in sampling, more elaborate statistical designs are often superior. In particular, matching the subjects in various ways can produce more precise results than simple randomization.

A simple and common special type of block design is the matched pairs design. Matched pairs designs compare just two treatments. Each block consists of just two units, as closely matched as possible. These units are assigned at random to the treatments, by tossing a coin or reading odd and even digits from Table B.

