CHAPTER 15. INTRODUCTION TO DESIGN

► Experiments vs. Observational Studies, Section 15.1

► Experimental design principles; Section 15.2

► Overview of designs for experimental studies; Section 15.3

Skin all of Chapter 15 Introduction to the Design of Experimental and Observational Studies.

Important for now: completely randomized design (CRD), completely randomized factorial design, randomized complete block design (RCB or RB), and nested design.

Section 15.1: Experimental and Observational Studies

We need just a little more detail about experimental and observational studies.

In the background, we have a specific population, and we wish to compare two or more "treatments."

A *treatment* is a specific set of conditions that in principle can be applied to the population (e.g. administer vitamin C to an individual child, apply a new fertilizer to plots of land, use a new method for teaching statistics in a class, etc.)

We will talk about *observational studies* and *experiments*.

"Treatment" doesn't necessarily mean medical treatment. It can be a new method of studying for tests, a new method for teaching Statistics, a new way of training for athletic events.

Observational study A study in which we observe individuals and measure variables, but don't interfere in any way.

Example: Method 1 in the vitamin C example

Experiment A study in which we specifically impose a "treatment" on individuals in order to observe their responses to it. Purpose is to try to determine whether the treatment *causes* a change in the response.

Example: Method 2 in the vitamin C example (the actual experiment)

Advantage of well-designed experiment over observational study

From observational studies, it is difficult (or impossible) to infer causation. From a well-designed experiment, we *can* conclude causation.

Observational study. In Method 1, you select people who decided for themselves to take vitamin C.

Usefulness of Randomization in Experiments

Randomization makes the two groups similar with respect to all factors (other than that one gets the new treatment and the other gets the old treatment). This includes factors we might have predicted would affect the response *and* factors we had not ever even thought about. Thus the two groups differ only in that one got the new treatment and the other got the old.

Another way to put this:

Randomization eliminates confounding *Confounding* means a difference between the treatment and control groups–other than the treatment—which affects the responses being studied. A *confounder* is a third variable, associated with the treatment factor and with the response.

In a controlled, randomized experiment, the difference in response between the two groups is caused either by a genuine difference in the effects of the two treatments *or* by chance.

If we can essentially exclude the possibility that chance alone caused the difference in the responses, then we have established causality.

A controlled, randomized experiment is the most desirable study design; it is the "gold-standard" design.
The statistical test of hypothesis is an important tool in establishing causality in these gold-standard designs.

If by use of randomization, we have eliminated the possibility of confounding in our experiment, then the result of a statistical hypothesis test is valuable information.

Statistical Significance

When the observed difference in responses between the experimental and control groups is so large that it cannot be attributed to just a chance outcome, the difference is called "statistically significant."

If we have randomized subjects to experimental and control groups, then there are just two possible explanations for an observed difference in the mean responses for the two groups:

1. There is a real population difference between effects of treatment and placebo, or
2. The observed difference is not "real," but occurred just by chance.

If the P-value of our two-sample t test (or other appropriate procedure) is very small, this eliminates Possibility 2, and proves the "alternative hypothesis" (Possibility 1).

Example. Vitamin C (See p. 4.)

Method 1 The people taking Vit C every day may be older and more likely to catch colds because of weaker immune system than the people not taking Vit C. This would produce a bias against Vitamin C. What is the confounder here?

Or, there could be a bias in favor of Vitamin C. Some offered by students in class: Healthier people take Vitamin C as part of their health regimen. Even without the Vitamin C, they get fewer colds because they eat better food than average, stay away from sick people, etc.

Method 2 An older person is equally likely to be assigned to the placebo group, as to the Vitamin C group. Thus there is no bias either for or against Vitamin C (at least, no bias, probabilistically speaking).

What to Do if We Have Observational Studies

If you can, do a controlled (i.e. randomized) study.

If you have to do with an observational study, then you can and should try to "control for confounding factors." This means to subdivide the population into smaller but more homogeneous subgroups, within which confounding factors do not exist.

Example:

Caution

Example. Smoking and Lung Cancer. The percentage of males who smoke is higher than the percentage of females who smoke. Let's suppose males are also more likely to get lung cancer, regardless of smoking status. Then, Gender would be a confounder of the association between Smoking and Lung Cancer.

STRATEGY: Subdivide date in Male/Female groups. In each group compare the percentage of lung cancer cases between smokers and non-smokers.

Caution The problem with this is that we need to know what the confounding factors are. The advantage of randomized studies is that they automatically take care of any confounding factors, whether or not you are aware of their presence.

Section 15.2: Basic Concepts in Experimental Studies

► Factor: an explanatory variable, with *levels* or values that it can take on

  ► Experimental factor is one that is assigned at random to units
  ► Observational factor is intrinsic to the unit and cannot be assigned at random

► Multifactor studies have more than one factor. *Ex.*: Effect of three levels of temperature and two levels of concentration of solvent on yield of a chemical process.

Example. Quick bread volume Effect of baking temperature (low, medium, high, very high) on volume of bread was studied by assigning five package mixes at random to each condition.

Design Vocabulary

Experimental factor: Vitamin C in Colds experiment, temperature in Quick Bread Volume.

Observational factor: Vitamin C in hypothetical Method 1 on Vit C and colds.

In regression, factors are usually called predictors or explanatory variables. Note that if we have an observational factor in an experiment, we still have an experiment wrt the primary factor of interest. However, we can't infer causality wrt the observational factor.

Steps in designing an actual experiment

Proper design requires consideration of what factors and how many factors to include, number and value of levels of each included factor, range of levels for quantitative factors, what control treatment to include.

It would be a good preview to set up the design matrix for the Quick Bread example.

Section 15.2 continued: Crossed vs. nested factors

► In multifactor studies, factors can be *crossed* or *nested*. *Ex. Yield of chemical process* If all 3×2 possible treatments (factor combinations) are included in the experiment, then the factors are *crossed*.

Temp | Concentration

		L	H
		L	H
L		X	X
M		X	X
H		X	X

"X" means treatment occurs in the experiment

► The levels of a *nested* factor are unique to a particular level of another factor. Ex. of study with *nested* factor: Effects of operators on production yield in three manufacturing plants. The operators are four different people in each of the three plants.

Sketch the designs

Plant

Operator
(w/in Plant)

	1	2	3	4	5	6	7	8	9	10	11	12
1	X	X	X	X								
2					X	X	X	X				
3									X	X	X	X

Levels 1, 2, 3, 4 of Operator only occur
in Plant 1, etc.

Nested design : observational study
Plant, operator are observational

► *Experimental unit*: the smallest unit to which a treatment can be assigned. Thus the method of randomization determines what is the experimental unit.
Example Quick Bread Volume. The twenty package mixes are the experimental units.

► Sample size and replication. *Ex* Quick bread; sample size is 20, and there are five (5) replicates of the basic experiment.

. **Experimental Units**

Ex. Vit C/Colds The child is the experimental unit.

Ex. Educational methods study Three high schools participated in a study to evaluate the effectiveness of a new computer-based math curriculum. In each school, four $24-$student sections of freshman algebra were available for the study. The two types of instruction (standard curriculum, computer-based curriculum) were randomly assigned to the four sections in each of the three schools. At the end of the term, a standard math achievement test was given to each of the 24 students in each section.

If the treatment is applied to all the students in the classroom, then the class is the unit, not the individual student.

► You do collect data (test scores) on each individual student separately. The unit on which you collect data may not be the same as the experiment unit.

► The **experimental unit** is that entity to which the treatment can be applied at random.

Important special case:

Sometimes *time units* are the experimental units.

Perhaps you want to run an experiment on effectiveness of two different hay fever medicines on yourself, over a ten-day period. The experimental units are the ten days.

So, for a CRD, you would randomly allocate five days to Medicine A, and the remaining five days to Medicine B.

How to randomize (conceptually) We will use a "random number generator" on the computer (in R) in practice.

Ex. Vit C/Colds Say there are 20 children in the experiment, for simplicity.

► Number the 20 children, 1 to 20.
► Generate "random permutation" of these numbers.
► Assign the children with the first 10 id numbers to Vitamin C, and the remaining children to Placebo.

Constrained randomization, or blocking

Eg. Say you want to control the number of males and females in the experiment, because you know gender is associated with frequency of colds.

Steps in a blocked experiment (one with constrained randomization):

► Select 434 males and 434 females for the experiment.
► Randomize the 434 males to Vit C and Placebo groups (as described above)
► Randomize the 434 females to Vit C and Placebo groups (as described above)

Section 15.3: EXPERIMENTAL DESIGNS

► Completely randomized design (CRD)
To compare two treatments, so far our experimental design has been a CRD:
   All experimental subjects are allocated at random among the two treatments.
More generally, the Quick Bread Volume experiment is an example of a CRD with four treatments, provided the twenty package mixes are randomly allocated, five apiece to each of the four treatments.

► Completely randomized factorial design
Uses the same randomization scheme as the CRD. There are two or more crossed factors which define the set of treatments.
Example. Chemical yield p. 43 Experimental unit might be a single run of the chemical process, using standard company equipment. In this 3×2 factorial design, if there are 18 runs assigned at random to the six treatments, this is a CR-factorial.

replicates ?   - - -   3

3×2 & 3 = 18

► Randomized complete block design. In a blocked design, first the
RCB experimental units are divided into homogenous groups, each with the same number of units as there are treatments. Then the treatments are assigned at random within each of the blocks separately.

Example Quick Bread Volume , where you have eight mixes, but four mixes were produced in Plant A and four mixes were produced in Plant B. You decide to run the experiment in two blocks of size four:

Note: The treatments can be defined by crossing of two or more factors in a factorial design; this would be called a factorial experiment run in blocks.

Skip

Randomized blocks: Although the mixes are in principle identical, there is concern that different plants have somewhat different processes and even raw materials. To control this source of variation, you include every treatment in each of the two blocks (the two manufacturing plants).

► Repeated measures design
This is a variation on randomized block design.
A repeated measures design is a randomized blocks experiment, where the subject is the block. The order in which a subject receives the treatments is randomized.

Example Compare three cereal formulations using $n = 12$ consumer volunteer taste-testers. Compare CRD and repeated measures design.

► Split-plot design
Example Does perceived wholesomeness affect rating? Six consumers are told that the cereals are formulations of a new organic, nutritious cereal; the other six consumers are told just that the formulations are new formulations of a cereal.