Chapter 2
Producing Data

Introduction

2.1 Design of Experiments

Introduction

- Anecdotal Data
- Available Data & Voluntary Response
- Sample Surveys and Experiments
- Observation versus Experiment

Obtaining Data

Available data are data that were produced in the past for some other purpose but that may help answer a present question inexpensively. The library and the Internet are sources of available data.

Beware of drawing conclusions from our own experience or hearsay. **Anecdotal evidence** is based on haphazardly selected individual cases, which we tend to remember because they are unusual in some way. They also may not be representative of any larger group of cases.

Some questions require data produced specifically to answer them. This leads to **designing** observational or experimental studies.

Voluntary Response Sampling - BEWARE!

- Beware of individuals who **choose themselves** to be involved, as in polls on websites. These samples are very susceptible to being biased because different people are motivated to respond or not. Often called "public opinion polls," these are not considered valid or scientific.
- **Bias:** Sample design systematically favors a particular outcome.

Ann Landers summarizing responses of readers

* 70% of (10,000) parents wrote in to say that having kids was not worth it—if they had to do it over again, they wouldn't.

Bias: Most letters to newspapers are written by people with strong opinions. A random sample later showed that 91% of parents WOULD have kids again.

Observation versus Experiment

When our goal is to understand cause and effect, experiments are the *only* source of fully convincing data.

The distinction between observational study and experiment is one of the most important in statistics.

An **observational study** observes individuals and measures variables of interest but does not attempt to influence the responses. The purpose is to describe some group or situation.

An **experiment** deliberately imposes some treatment on individuals to measure their responses. The purpose is to study whether the treatment causes a change in the response.

In contrast to observational studies, experiments don't just observe individuals or ask them questions. They actively impose some treatment in order to measure the response.

Observational study: Record data on individuals without attempting to influence the responses.

Example: Based on observations you make in nature, you suspect that female crickets choose their mates on the basis of their health. → Observe health of male crickets that mated.

Experimental study: Deliberately impose a treatment on individuals and record their responses. Influential factors can be controlled.

Example: Deliberately infect some males with intestinal parasites and see whether females tend to choose healthy rather than ill males.

Confounding

Observational studies of the effect of one variable on another often fail because of **confounding** between the explanatory variable and other variables.

> **Confounding** occurs when two variables are associated in such a way that their effects on a response variable cannot be distinguished from each other.

Well-designed experiments take steps to avoid confounding.

2.1 Design of Experiments

- Experimental Units, Subjects, Treatments
- Comparative Experiments
- Bias
- Principles of Experimental Design
- Statistical Significance
- Matched Pairs Design
- Block Design

Individuals, Factors, Treatments

An experiment is a statistical study in which we actually do something (a **treatment**) to people, animals, or objects (the **experimental units**) to observe the **response.** Here is the basic vocabulary of design of experiments.

The **experimental units** are the smallest collection of individuals to which treatments are applied. When the units are human beings, they often are called **subjects.**

The explanatory variables in an experiment are often called **factors,** and their values are called the **levels of the factors**

A specific condition applied to the individuals in an experiment is called a **treatment.** If an experiment has several explanatory variables, a treatment is usually a combination of the levels of the factors

Comparative Experiments

Experiments are the preferred method for examining the effect of one variable on another. By imposing the specific treatment of interest and controlling other influences, we can pin down cause and effect. Good designs are essential for effective experiments, just as they are for sampling.

A high school regularly offers a review course to prepare students for the SAT. This year, budget cuts will allow the school to offer only an online version of the course. Here's picture of the design:

Students → Online Course → SAT Scores

Over the past 10 years, the average SAT score of students in the classroom course was 1620. The online group gets an average score of 1780. That's roughly 10% higher than the long-time average for those who took the classroom review course.

Is the online course more effective?
How would you know?
Are you certain the increase is due to the online course?

Comparative Experiments

Many laboratory experiments use a design like the one in the online SAT course example:

In the laboratory environment, simple designs often work well.

Field experiments and experiments with animals or people deal with more variable conditions.

Outside the laboratory, badly designed experiments often yield worthless results because of ***confounding*****.**

Randomized Comparative Experiments

The remedy for confounding is to perform a **comparative experiment** in which some units receive one treatment and similar units receive another. Most well-designed experiments compare two or more treatments.

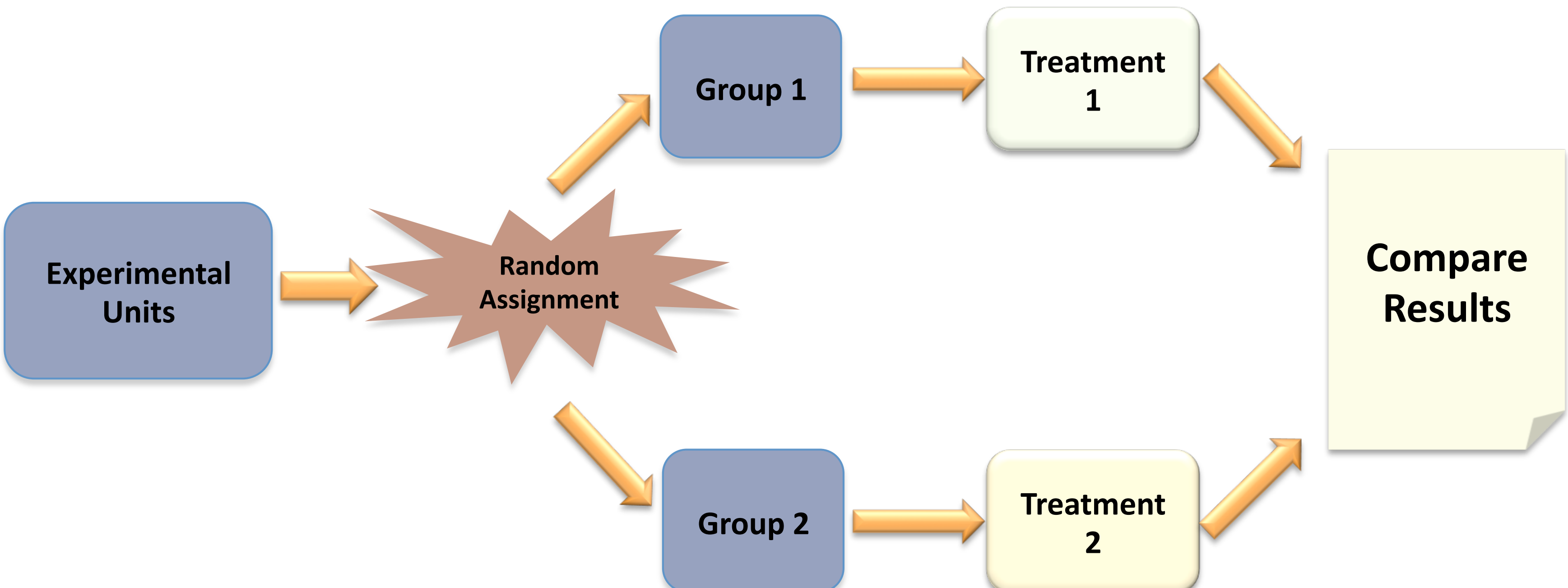
Comparison alone isn't enough. If the treatments are given to groups that differ greatly, **bias** will result. The solution to the problem of bias is **random assignment.**

In an experiment, **random assignment** means that experimental units are assigned to treatments at random, that is, using some sort of chance process.

Randomized Comparative Experiments

In a **completely randomized design,** the treatments are assigned to all the experimental units completely by chance.

Some experiments may include a **control group** that receives an inactive treatment or an existing baseline treatment.



In Comparative Experiments, we compare the response to a treatment to:

Another treatment

No treatment (a control)

A placebo

Or any combination of the above

A **control** is a situation where no treatment is administered. It serves as a reference mark for an actual treatment (e.g., a group of subjects does not receive any drug or pill of any kind).

A **placebo** is a fake treatment, such as a sugar pill. This is to test the hypothesis that the response to the actual treatment is due to the actual treatment and not the subject's apparent treatment.

Check out this article: http://people.uncw.edu/frierson/S215%5CMigranes.pdf

About the placebo effect

The "placebo effect" is an improvement in health not due to any treatment, but only to the patient's belief that he or she will improve.

- The "placebo effect" is not totally understood, but it is believed to have therapeutic results on up to a whopping 35% of patients.
- It can sometimes ease the symptoms of a variety of ills, from asthma to pain to high blood pressure, and even to heart attacks.
- An opposite, or "negative placebo effect," has been observed when patients believe their health will get worse.

Designing "controlled" experiments

Sir Ronald Fisher—The "father of statistics"—was sent to Rothamsted Agricultural Station in the United Kingdom to evaluate the success of various fertilizer treatments.

Fisher found that the data from experiments that had been going on for decades was basically worthless because of poor experimental design.

- Fertilizer had been applied to a field one year and not another, in order to compare the yield of grain produced in the two years. BUT
  - It may have rained more or been sunnier during different years.
  - The seeds used may have differed between years as well.

- Or fertilizer was applied to one field and not to a nearby field in the same year. BUT
  - The fields might have had different soil, water, drainage, and history of previous use.

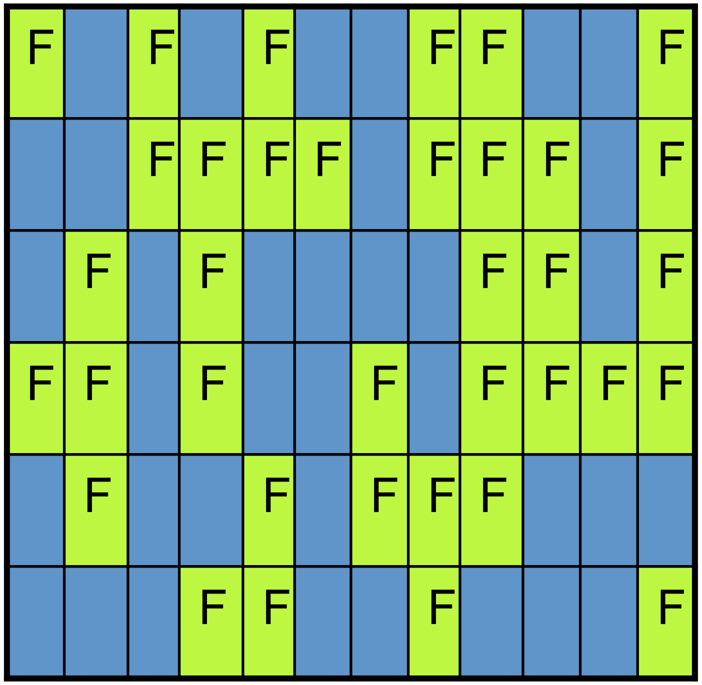
➔ Too many factors affecting the results were "uncontrolled."

Fisher's solution:

"Randomized comparative experiments"

- In the same field and same year, apply fertilizer to randomly spaced plots within the field. Analyze plants from similarly treated plots together.

- This minimizes the effect of variation within the field, in drainage and soil composition on yield, as well as controls for weather.



A Table of Random Digits can be used to Randomize an Experiment – Look at Table B in the Resources Tab…

- any digit in any position in the table is as equally likely to be 0 as 1 as 2 as … as 9
- the digits in different positions are independent in the sense that the value of one has no influence on the value of any other
- any pair of random digits has the same chance of being picked as any other (00, 01, 02, … 99)
- any triple of random digits has the same chance of being picked as any other (000, 001, … 999)
- and so on…
- EXAMPLE: Use Table B to randomly divide the 72 plots in the previous slide into the two groups, treatment ("Fertilizer") and control ("No Fertilizer")
  - Step 1: Label the experimental units with as few digits as possible
  - Step 2: Decide on a protocol for how you will place the chosen units into the groups
  - Step 3: Start anywhere in the Table and begin reading random digits. Matching them with labeled experimental units and following the protocol creates the groups – Do it with this example…

Principles of Experimental Design

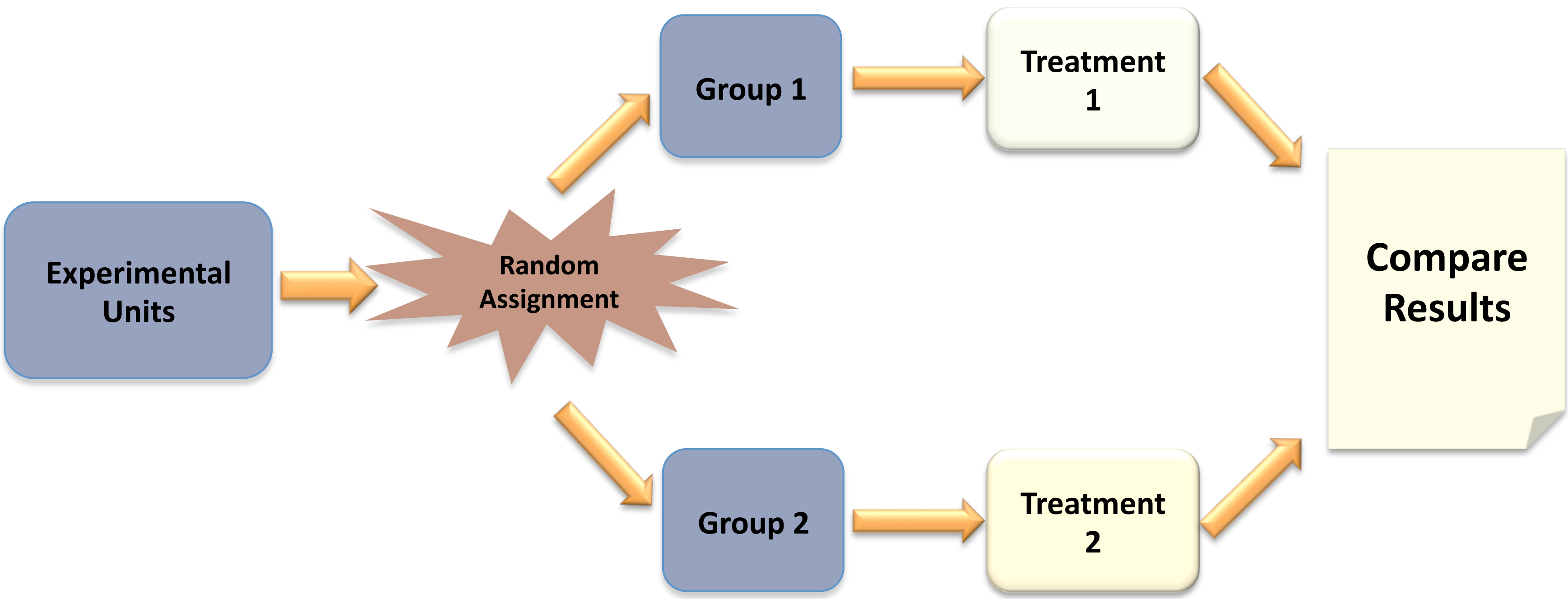
Randomized comparative experiments are designed to give good evidence that differences in the treatments actually cause the differences we see in the response.

Principles of Experimental Design

1. **Compare** two or more treatments and **Control** for "lurking" variables.
2. **Randomize:** Use impersonal chance to assign experimental units to treatments.
3. **Repeat or Replicate** each treatment on many units to reduce chance variation in the results.

An observed effect so large that it would rarely occur by chance is called **statistically significant.**

A statistically significant association in data from a well-designed experiment does imply causation.



Experimental Units
Random Assignment
Group 1
Group 2
Treatment 1
Treatment 2
Compare Results

Cautions about Experimentation

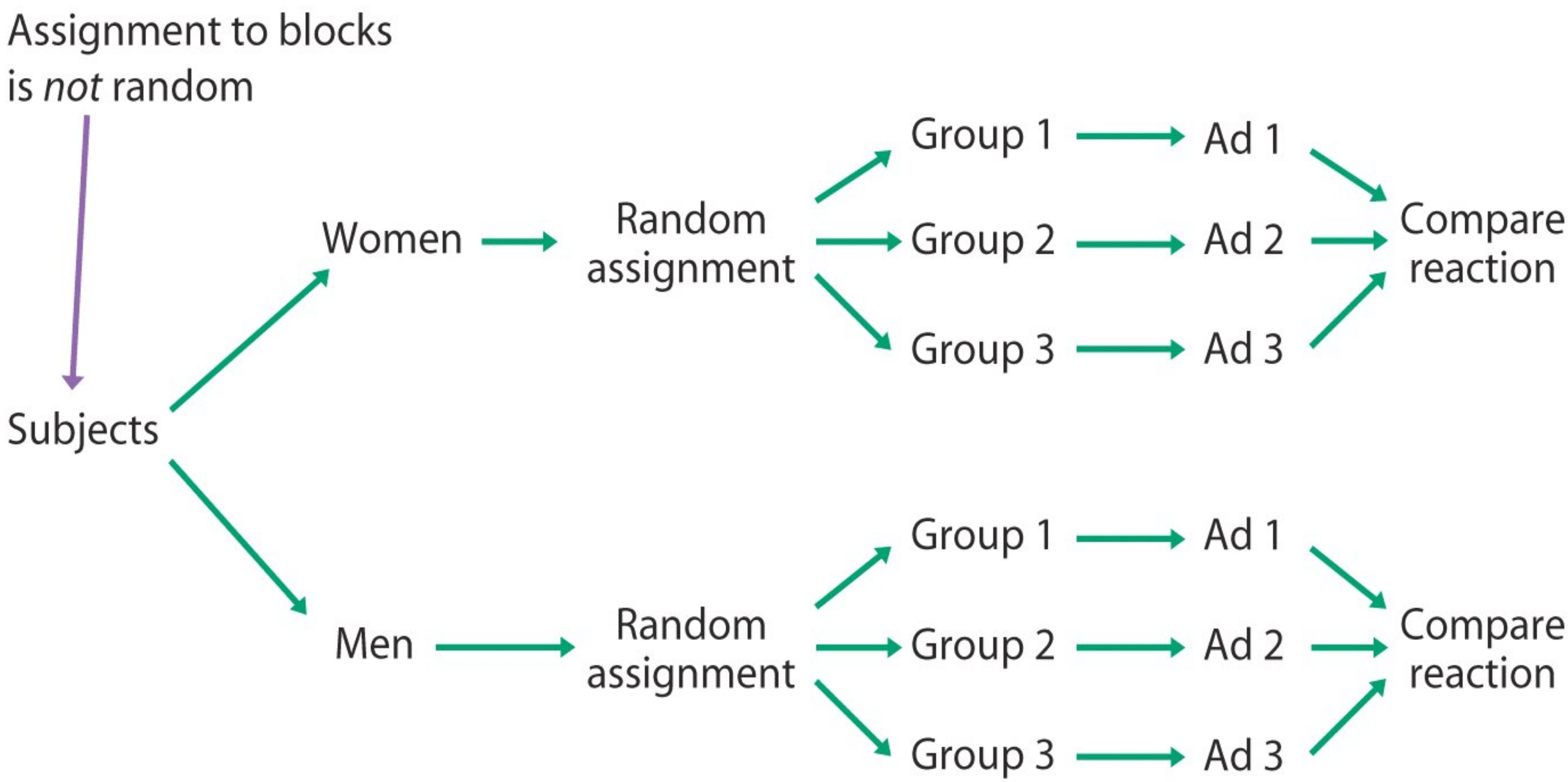
The logic of a randomized comparative experiment depends on our ability to treat all the subjects the same in every way except for the actual treatments being compared.

In a **double-blind experiment,** neither the subjects nor those who interact with them and measure the response variable know which treatment a subject received.

The most serious potential weakness of experiments is **lack of realism.** The subjects or treatments or setting of an experiment may not realistically duplicate the conditions we really want to study;
e.g., extension of animal studies to humans; extension of college student studies to larger populations; etc....

Blocked Designs

In a **block,** or **stratified,** design, subjects are divided into groups, or blocks, prior to experiments, to test hypotheses about differences between the groups.



Control what you can, block on what you can't control, and randomize to create comparable groups.

Matched Pairs

Matched pairs: Choose pairs of subjects that are closely matched—e.g., same sex, height, weight, age, and race. Within each pair, randomly assign who will receive which treatment; e.g., lead poisoning study. See this link on my website:

http://people.uncw.edu/frierson/S215%5Clead_levels.jmp

It is also possible to just use a single person, and give the two treatments to this person over time in random order. In this case, the "matched pair" is just the same person at different points in time.

Identical twins are the ultimate matched pairs!

HW: Watch the StatTutor video on Randomized Block Designs and Matched Pairs and answer the questions asked during that video.

Homework

1. Get signed up for the Stats Portal.
2. Finish reading Chapter 1 and take the online quiz (under "Assignments"), deadline is August 27th at 9:00am – you get 2 chances, I'll average the two scores.
3. Begin reading Chapter 2, read section 2.1 and watch the video (under "Assignments")
4. Go over the 2.1 Summary on page 50 – work on # 2.23, 2.27, 2.30, 2.31, 2.37, 2.40