Chapter 11

Experiments and Observational Studies
11.1

Observational Studies
Music and Good Grades

The Study:
- Compared GPAs of music students and non-music students at Mission Viejo High School.

The Results
- Music Students: 3.59
- Non-Music Students: 2.91

Conclusions
- Should we make all students play an instrument?

Issues
- Could there be something else resulting in both?
Observational Studies

• Researchers don’t assign choices.
• Passively observe participants
• Good for discovering relationships related to rare outcomes
• Bad for establishing cause-and-effect relationships
• Tough to handle lurking variables
  • Do musicians have more supportive parents that help GPA?
  • Are smarter people more inclined to play an instrument?
Retrospective Studies

• Collect data on something that has already occurred
• Similar pros and cons as observational studies
• Additional issues can include:
  • Unreliable memories
  • Incomplete historical records
  • Often limited to a small part of the population
Prospective Studies

A **prospective study** is a study where we identify subjects in advance and collect data as events unfold.

- **Pros:**
  - Possible to isolate the variables.
  - With care, can establish cause and effect.
  - Can design the study to your specifications.

- **Cons:**
  - Can be expensive.
  - Rare occurrences require very large samples.
  - Can take too long:
    - Do breast-fed babies live longer than bottle-fed?
Why Did Fido and Fluffy Die?

Early 2007, many dogs and cats died of kidney failure. Should you conduct a retrospective or prospective study to find out why?

• Retrospective Study
  • The event happened in the past.
  • It may have been a rare event.
  • The retrospective study may provide clues on the cause.

• Once possible causes are identified, try a prospective study to verify the causes - if it doesn’t kill any pets.
11.2

Randomized, Comparative Experiments
Experiments

Is it possible to establish a cause and effect relationship?

- Take 100 young children. Randomly select 50 to be in a music program. The other 50 will not be allowed to play an instrument.

- An experiment requires random assignment of subjects to treatments.

- Only experiments can establish cause and effect.
How Experiments Work

- Identify the explanatory variable(s), called the factor(s).
- Identify the response variable.
- Select subjects or participants (if human) or experimental units (if not human).
- Decide on the levels to choose for each factor.
  - Music program or no music program
  - Sleep hours: 4, 6, or 8
- The combination of specific levels from all factors that a subject receives is called its treatment.
Assigning Participants to Treatments

- Don’t let them choose.
- Don’t assign based on what’s best for each.
- Randomly assign participants into groups. Each group receives a different treatment.
- Only through **random assignment** can a cause-and-effect relationship be established.
- What ethical dilemmas might this introduce?
Women’s Health Initiative

- A major NIH observational study of 93,000 women
- Goals:
  - Known and unknown risk factors for serious diseases
  - Compare the risk factors, presence of disease at start of study and new occurrences as study progressed across all study components.
  - Create a resource to identify biological indicators of the disease.
Women’s Health Initiative

• The goals could only be achieved with an experiment.

• Postmenopausal women randomly assigned to one of two groups:
  • Group 1: Hormone replacement therapy
  • Group 2: Inactive pill

• Result: Hormone replacement with estrogen increased risk of stroke.
Are Fido and Fluffy Still Dying?

- The deaths were traced to food contamination.
- The contaminated food was thrown out.
- To test if the new food is safe, what would be the treatments and response variables?
  - **Treatments:** Some pets receive food from the companies that now claim the problem has been dealt with. The other pets would be given food from known safe stocks.
  - **Response Variable:** Veterinarian’s assessment
11.3

The Four Principles of Experimental Design
Control and Randomization

1. Control
   - Make all conditions as similar as possible for all treatment groups.
   - Control allows us to isolate the one thing that is being studied. Helps avoid lurking variables

2. Randomize
   - Equalizes the effects of variation that we cannot control
   - Distributes the uncontrollable factors equally

Control what you can, randomize the rest.
Replicate and Block

Replicate
- Apply each treatment to a number of subjects.
- Repeat the entire experiment on an entirely different population of experimental units.

Block
- Group similar individuals together and randomize within each of these blocks.
- Blocking helps account for the variability due to the difference between blocks.
Is the Pet Food Safe?

Control

- Food and water portions, housing, exercise, sleep. Stick to one breed.

Randomize

- Assign dogs to two feed treatments randomly.

Replicate

- Redo the entire experiment with a different breed.
Diagrams

A diagram is a useful organizational tool.

- Flow charts help to understand the process
- Side-by-side boxplots help compare the treatment groups.
Chlorine and Nail Polish

Think →

- **Plan:** Does chlorine react differently for two nail polish colors?

- **Response:** Measure the percent of polish chipped away.

- **Treatments:** Red and Nude colored nail polish

- **Experimental Units:** Control variability by gluing 30 acrylic nails to chopsticks.
Chlorine and Nail Polish

Experimental Design

- Control: Same brand of nails and polish, same painting time and soaking time
- Replicate: 15 nails of each color
- Randomly Assign: Label nails 1-30. Use StatCrunch to select 15 of these to be red.

Make a Picture
Chlorine and Nail Polish

Show

- Use side-by-side boxplots to compare the treatments.
- Compare the means.

Tell

- If the difference in the means is larger than what I would expect due to sampling error, then I may conclude the difference is due to color.
Statistical Significance

A difference is called **statistically significant** if the difference is greater than what we would expect from random chance.

- Flip a coin 100 times:
  - 52 tails is not statistically significant since it would not be surprising to observe this outcome.
  - 93 tails is statistically significant since it would be surprising to observe this outcome.
Statistical Significance

Possible boxplots from the nail polish experiment.

Statistically significant since the medians of each are outside the typical values of the other.

Not statistically significant since the medians of each are within the typical values of the other.
Random Samples and Random Treatments

- Surveys use a random group of participants.

- Experiments find a homogeneous group, separate them into random subgroups for treatment.

- Experiments do not use a random sample from the population.

- Beware of stating that the participants from the experiment represent the larger population.
Control Treatments
Control Groups and Control Treatments

Control

- Does eating ten carrots a day help you lose weight?
- Find 200 participants and randomly select 100 of them to eat ten carrots a day.
- The other 100 are the control group.
- Not eating ten carrots a day is the control treatment.
Blinding

What brand of cola is the best?

- If you give participants cans of cola and ask how much they like it, the label can be an influence.
- Instead give each an unlabeled cup of soda.
- **Single-blindling** involves the participants not knowing whether they are in the control or treatment group.
- If the person handing out the cups hands out her favorite soda she may bias the results.
- **Double-blindling** means neither the participant nor the person handing out the soda knows the label.
Who Can Affect the Experiment

There are two main classes of individuals who can affect the experiment.

- Those who can influence the results.
  - Subjects
  - Treatment administrators
  - Technicians

- Those who evaluate the results.
  - Judges
  - Treating Physicians
Blinding Examples

Pet Food: Should blinding be used when comparing the questionable pet food with the known safe food?
  • Yes, the veterinarian should not know which is which.

Nail-polish chipping and color: Should the nail polish experiment have been conducted blindly?
  • Yes, black and white photos of the nails could have been taken and an independent judge could have looked at these photos.
Placebos

- A *placebo* is a “fake” treatment that looks like the treatment being tested.

- Just telling a patient that they are being treated can aid recovery.

- This is called the *placebo effect*.

- Use a placebo for effective blinding.
Placebo Effect Really is True

The Study

- One group received Ginkgo and the other received a placebo.
- 13 memory tests were given.
- Ginkgo better for 6 tests.
- Placebo better for 7 tests.
Summary of Experimental Techniques

- Randomized
- Comparative
- Double-blind
- Placebo-controlled
11.5 Blocking
Nail Polish Revisited

Comparing Nail Polish on Real Hands

• To decrease variation, randomly assign half the right hand fingers red and other half nude.
• Do the same for the left hand side.
• This ensures that a color is not overrepresented by the right hand.
• This is similar to stratified sampling.
• Each hand serves as a separate experiment.
Blocking

• Experimental units can be separated into groups that are not the treatment, we call these groups blocks.

• Blocking involves randomly assigning the treatments within each block.

• Blocking helps isolate the variability due to the differences between blocks.

• Blocking helps clarify the difference between the treatments.

• The design is called a randomized block design.
Chart of a Randomized Block Design

Block

Block A
2 fingers
right hand

Block B
2 fingers
left hand

Random Assignment

Group 1
Treatment 1
red

Group 2
Treatment 2
nude

Group 3
Treatment 1
red

Group 4
Treatment 2
nude

8 fingers
(no thumb):
4 on the
right hand
and 4
on the
left hand

Compare
polish
chipped

Compare
polish
chipped
Fido and Fluffy as Blocks

Blocking to Determine Whether the Pet Food is Safe

- Block by animal type.
  - Half of the dogs get the safe food and half get the suspect food.
  - Half of the cats get the safe food and half of the cats get the suspect food.

- Measure the responses separately and look at the results afterward.
11.6

Confounding
Animated Teaching vs. Subdued Teaching

Professor Ceci taught the same course in the fall and the spring.

- Fall: Subdued manner, everything else the same
- Spring: High enthusiasm, animated gestures

Results: How much did you learn? (1-5)

- Fall: 2.93
- Spring: 4.05

Conclusions

- Animated teaching better than subdued teaching???
- Weather: Fall ends gloomy, spring ends pleasant.
Confounding Factors

Two factors are **confounded** if the levels of one are associated with the levels of the other.

- Weather and Professor Cecil’s style were confounded.
- Try to avoid confounding factors, but it is difficult and sometimes impossible.
- Avoiding confounding factors can introduce new ones.
- Compare morning and afternoon fall courses.
Two Factors

A bank sent out 50,000 low-rate-no-fee offers and 50,000 high rate with fee offers.

• Many more responses for low rate, no fee

• Was it the low rate or no fee that customers liked?

• Should have sent out four types of offers:
  – Low rate, no fee
  – Low rate, with fee
  – High rate, no fee
  – High rate, with fee
Lurking and Confounding

Lurking Variable
- Associated with both $x$ and $y$
- Makes it appear that $x$ causes $y$

Confounding Variable
- Associated in a noncausal way with a factor
- Affects the response
- Can’t tell if the cause was the factor or confounding variable
What Can Go Wrong?

- Don’t give up just because you cannot run the experiment.
- Sometimes we have to resort to an observational study. Do airbags lower the risk of dying?

Beware of Confounding
- If possible, use randomization.
- If not possible, report out likely confounding factors.
What Can Go Wrong? (Continued)

- Bad things can happen even to good experiments.
  - Record additional information that may help in the analysis.

- Don’t spend your entire budget on the first run.
  - Try a small pilot run first.
  - Then refine the factor levels and fix other problems.
  - Finally, perform the full-scale study.