4.2: Experiments

SAT Survey vs. SAT Experiment

Describe a survey and an experiment that can be used to determine the relationship between SAT scores and hours studied?

Section 4.2 Experiments

After this section, you should be able to…

- ✓ DISTINGUISH observational studies from experiments
- ✓ DESCRIBE the language of experiments
- ✓ APPLY the three principles of experimental design
- ✓ DESIGN comparative experiments utilizing completely randomized designs and randomized block designs, including matched pairs design

- **Survey:** Asking students about how many hours they studied for the SAT and their resulting scores.
- **Experiment: Randomly** selecting a group of same/similar IQ students and **assign** each student a different number of hours to studying for the SAT. The student is ONLY allowed to study the mandated amount of hours. Then, **compare** their result scores

Observational Study vs. Experiment

An **observational study** observes individuals and measures variables of interest but does not attempt to influence the responses.

An **experiment** deliberately imposes some treatment on individuals to measure their responses.

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

Confounding Variables

A **confounding variable** is one whose effects on the response variable cannot be distinguished from one or more of the explanatory variables in the study. Confounding = included.

Confounding = Confused

Confounding Variables

- Confounding refers to a problem that can arise in an experiment, when there is another variable that may effect the response and is in some way tied together with the factor under investigation, leaving us unable to tell which of the two variables (or perhaps some interaction) caused the observed response.

What's Lurking?!

1. As shoe size increases…so does reading ability.
2. An increase in ice cream consumption equals an increase in the number of drowning deaths for a given period.
3. As savings account balances increase….life expectance decreases.

Confounding Variables

- For example, we plant tomatoes in a garden that's half-shaded. We test a fertilizer by putting it on the plants in the sun and apply none to the shaded plants. Months later the fertilized plants bear more and better tomatoes.
- Why?
- **Well, maybe it's the fertilizer, maybe it's the sun, maybe we need both.** We're unable to conclude that the fertilizer works because any effect of fertilizer is confounded with any effect of the extra sunshine.

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. 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? Is there a lurking variable? Is there a confounding variable?

Lurking Variable

A **lurking variable** is a variable that is not among the explanatory or response variables in a study but that may influence the response variable. Lurking = **not included.**

Lurking = Missing Link

Three Principles of Experimental Design

1. **Control** for lurking variables that might affect the response: Use a comparative design and ensure that the only systematic difference between the groups is the treatment administered.
2. **Random assignment**: Use impersonal chance to assign experimental units to treatments. This helps create roughly equivalent groups of experimental units by balancing the effects of lurking variables that aren't controlled on the treatment groups.
3. **Replication**: Use enough experimental units in each group so that any differences in the effects of the treatments can be distinguished from chance differences between the groups.

The Randomized Comparative Experiment

- 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.

Factor v. Treatment

For example, three different groups of runners are subjected to different training methods.

Experimental units- runners

Factor- Training methods

Treatments- Specific type of workout: Speed, strength training and distance workouts

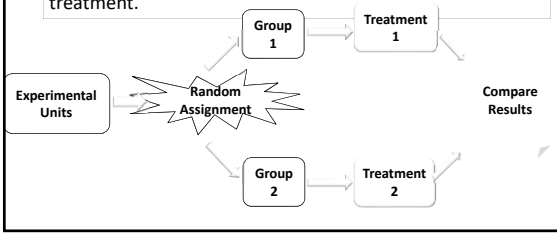
Factor = General Group

Treatment = Specific Implementation

The Randomized Comparative Experiment

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.



A cookie manufacturer is trying to determine how long cookies stay fresh on store shelves, and the extent to which the type of packaging and the store's temperature influences how long the cookies stay fresh. He designs a completely randomized experiment involving low (64 Fº) and high (75 Fº) temperatures and two types of packaging—plastic and waxed cardboard.

List the **experimental units**, **factors**, and **treatments** in this experiment.

The Language of Experiments

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

Factors: General name for explanatory variables in an experiment (multi-vitamin regime).

Treatment: a specific condition (given vitamin A vs. vitamin B; time frame vitamin taken) applied to the individuals in an experiment.

Experimental units: packages of cookies.

Factors: Temperature and packaging.

Treatments: Low temp and plastic, high temp and plastic, low temp and waxed cardboard, high temp and waxed cardboard.

Specific Types of Experimental Design

- Double-Blind
- Single-Blind
- Matched Pairs
- Block Design

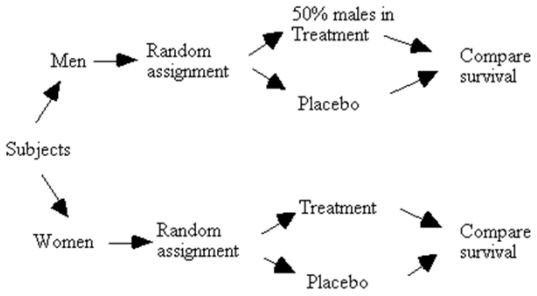
Matched Pair Design

- **Example:**
  - Tire wear and tear.
  - Put one set of tires on the left side of the car and a different set on the right side of the car.
  - This would help control the lurking variable of different driving styles (between teenage boys vs. teachers) and mileage driven.

Double-Blind

- In a double-blind experiment, neither the subjects nor the experimenters know which treatment a subject received.

Block Design



Matched Pair Design

- In a matched pair design, subjects are paired by matching common important attributes.
- Some times the results are a pre-test and post-test with the unit being "matched" to itself.

Block Design

- A block is a group of experimental units or subjects that are known before the experiment to be similar in some way that is expected to affect the response to the treatments.
- In a block design, the random assignment of units to treatments is carried out separately within each block.
- Helps control for lurking variables.

Block Design

- Experiments are often blocked by
  - Age
  - Gender
  - Race
  - Achievement Level (Regular, Honors, AP, IQ level, etc.)

Inference for Experiments

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.