The Huntington Library, Art Collections, and Botanical Gardens

Cause, Effect, or Coincidence?

Overview

Students will sharpen their critical thinking skills by understanding causality, correlation, and coincidence. In addition, this lab will familiarize students with the two conditions that support causality, necessity and sufficiency.

Introduction

Many times in our lives two events happen and we find ourselves asking, "What are the chances of that happening?" However, other then speculation how do we determine if two events are connected in some meaningful way? This is the job of a scientist. **A scientist's job is to try to determine causal relationships.** A cause is something that produces an effect. In order to determine causality, the scientist must prove that two events are not a coincidence or a simple correlation, but have a direct effect. First we should review coincidence and correlation.

Coincidence

A coincidence is a sequence of events that, although accidental, seem to have been planned, arranged, or correlated.

Examples

- Ice cream sales reflect the number of shark attacks on swimmers.
- As stock prices go up, skirt lengths get shorter.
- Elementary school children with more cavities have larger vocabularies.
- The more churches there are, the more violent crime there is (1988 census data).
- "On NPR a few weeks ago, I heard a most interesting statistic. A researcher at some university had arrived at the astounding conclusion that the more male children you have in a family, the more likely it is that one of them will grow up to become a homosexual. Why stop there? I would argue that this also increases the likelihood that one of them will grow up to be president, or a used car salesman, or be hit by a meteorite, or win the lottery, or contract syphilis, or rob a bank, or die while ice-fishing, or... " - Scott Hancock (submitted to members.cox.net/mathmistakes/reader.htm)

Correlation

A correlation means that two events happen together.

Example

- Sally dropped her hammer, and then felt a sharp pain in her foot.
- Observation: Two unusual events: dropping hammer, feeling pain.
- Conclusion: These events are correlated.

Just because two events are correlated doesn't mean that one causes the other. They could both be caused by something else, or they could just be a coincidence.

Correlation is a tool, and any tool, if misused, is capable of doing harm. Use a hammer the wrong way, and you will smash your thumb. Use correlation the wrong way, by jumping too quickly and simple-mindedly to inferences about cause and effect, and you will arrive at conclusions that are false and misleading.

For example, suppose you do a study on traffic safety and discover that a greater proportion of people driving to a vacation spot are injured in traffic accidents than people driving to work. Can you conclude that vacationers are worse drivers than commuters? Consider that on a vacation trip the driver covers many more miles than on a commute to work. Also consider that most commuters drive alone, while on a vacation trip they take their families. Recall that a cause is something that produces an effect. Demonstration of a cause and effect relationship requires much more than just a correlation. There must be demonstrable evidence that the alleged cause is actually responsible for the effect. Demonstration of cause can be very difficult. Sometimes a correlation, such as the one between cholesterol levels and heart disease, is so strong that we decide that high cholesterol must cause heart disease. Yet there is the possibility that the same risk factors that cause heart disease coincidentally cause blood cholesterol levels to rise. Careful application of the scientific method can sort out causation from coincidence.

Cause

A causation argument is that which ties two things together as cause and effect. How do you identify a causal relationship? In a controlled experiment, you change some variable (X) and look for changes in another variable (Y).

An entertaining demonstration of the misuse of causality appeared in an episode of The Simpsons (Season 7, "Much Apu about Nothing"):

Homer: *Not a bear in sight. The "Bear Patrol" must be working like a charm.*

Lisa: *That's specious reasoning, Dad.*

Homer: *Thank you, dear.*

Lisa: *By your logic I could claim that this rock keeps tigers away.*

Homer: *Oh, how does it work?*

Lisa: *It doesn't work.*

Homer: *Uh-huh.*

Lisa: *It's just a stupid rock. But I don't see any tigers around, do you?*

Homer: *Lisa, I want to buy your rock.*

The philosopher David Hume argued that any form of causality cannot be perceived (and therefore cannot be known or proven), and instead we can only perceive correlation. However, we can use scientific method to rule out false causes. Let's look at the following data set and consider that

- A cause is **necessary** when a particular variable is required to produce an effect, although there may be other variables involved.
- A cause is **sufficient** when a particular variable inevitably initiates or produces an effect. A cause can be necessary, sufficient, neither, or both.

Experiment	L: Light	D: Dirt	W: Water	G: Seed germinates?	S: Plant survives 1 week?
1	Yes	Yes	Yes	Yes	Yes
2	Yes	Yes	No	No	No
3	Yes	No	Yes	Yes	Yes
4	Yes	No	No	No	No
5	No	Yes	Yes	Yes	No
6	No	Yes	No	No	No
7	No	No	Yes	Yes	No
8	No	No	No	No	No

What conclusions can you draw from this data?

	L->G	D->G	W->G	L->S	D->S	W->S	G->S
Necessary?	No	No	Yes	Yes	No	Yes	Yes
Sufficient?	No	No	Yes	No	No	No	No

From this experiment, we draw the conclusions that:

1. Light is neither necessary nor sufficient for germination.
   Dirt is neither necessary nor sufficient for the plant to germinate.
   Water is necessary for the plant to germinate.
   Water is sufficient for the plant to germinate.
   Therefore, water is necessary and sufficient to cause germination.

2. Light is necessary for the plant to survive one week.
   Light is not sufficient for the plants to survive one week.
   Therefore, light is necessary, but not sufficient to allow the plant to survive one week.

3. Dirt is neither necessary nor sufficient for the plant to survive one week.

4. Water is necessary for the plant to survive one week.
   Water is not sufficient for the plants to survive one week.
   Therefore, water is necessary, but not sufficient to allow the plant to survive one week.

5. Germination is necessary for the plant to survive one week.

In the previous example, we saw several of the variables had causal relationships. Let us look at the four possibilities more closely.

The four possibilities of necessity and sufficiency:

1. **Necessary and sufficient causes:**
   X is required and will alone lead to Y
   Example: Trisomy of the human chromosome 21 causes Down's syndrome.

2. **Necessary but not sufficient causes:**
   X will, under the right circumstances, lead to Y
   Examples: Early exposure to language will lead to language development in children.
   Having the money to buy a car will lead to the purchasing of a new car.

3. **Sufficient but not necessary causes:**
   X will lead to Y, but so will other things.
   Example: A severe ice storm on a weekday will lead to school closings

4. **Contributory causes:**
   X is neither a necessary nor sufficient cause of Y, but changes the likelihood that Y occurs.
   Examples: Smoking and lung disease
   Viewing T.V. violence and aggressive behavior

The scientist's job is to try to determine causal relationships. This process often starts with an observation of some variable and an outcome:
"Wow, its seems like the seeds in my garden don't grow if I don't water"

You just discovered what seems like a nice correlation: if you water the seeds, they grow. You may even quantitate how often this happens: nine out of ten times when I added water to my garden, I got flowers. The strengths of using correlation in experiments lie in the ability to make observations in a "natural setting" (like your home garden), and the ability to tie a hypothesis directly to the observations: water appears to play an important role in germination. The weaknesses of using correlation can be the inability to establish the directionality of the two factors, and the possibility of unknown variables complicating the situation.

Directionality Problem: X --> Y or Y --> X
   Example: Viewing violent TV causes aggressive behavior, or "Do aggressive children like to watch violent TV?"
Third or other complicating variables problem: Z --> X and Z --> Y
   Example: Aggression in the home leads to both the viewing of violent TV and the aggressive behavior of children.

So, as a scientist, you next decide to design an experiment in your lab where you control several variables, like our experiment above with light, dirt, and water. The strengths of this method are that it can identify casual relations and that you can standardize all other conditions (e.g., the humidity of the plants or the type of soil). The weaknesses of this method are that experiments are often conducted in artificial settings, so that the result(s) may not generalize to the "real world". In addition, it may be impractical, or unethical, to perform the manipulations of the variable you would like to alter (for example, this crops up often when trying to study human genetics: it would be unethical to force people to have a kid so that you could see if disease X manifests).

Motivation

The students, as prominent scientists, have been chosen by the publishing office of Science Magazine to peer-review the following experiment and determine if the scientists who collected the data have enough evidence to evaluate the causes of seed germination and plant survival.

Objectives

Upon completion of the lab students will be able to

1. Define coincidence and give some examples
2. Define correlation and give some examples
3. Define causality
4. Describe and give examples of the conditions that need to be meet to determine causality

Materials

- Student handouts

Associated California Science Standards- Investigation and Experimentation

1b. Identify and communicate sources of unavoidable experimental error
1c. Identify possible reasons for inconsistent results, such as sources of error or uncontrolled conditions.
1d. Formulate explanations by using logic and evidence.
1f. Distinguish between hypothesis and theory as scientific terms.
1g. Recognize the usefulness and limitations of models and theories as scientific representations of reality.
1j. Recognize the issues of statistical variability and the need for controlled tests.
1k. Recognize the cumulative nature of scientific evidence.
1n. Know that when an observation does not agree with an accepted scientific theory, the observation is sometimes mistaken or fraudulent and that they theory is sometimes wrong.

Procedure

1. Hand out student activity sheet.
2. Review coincidence, correlation, and causality.
3. When discussing causality, demonstrate that two conditions must be met for there to be a causal relationship. The first event must be necessary to lead to the second event and the first event must be sufficient to cause the second event.
4. Have students finish handout.

Evaluation

The following questions are listed under the Analysis section of the student handout and may be used as part of a report, class discussion or assessment.

1. Define coincidence and give some examples.
2. Define correlation and give some examples.
3. Define causality.
4. Describe and give examples of the conditions that need to be met to determine causality.
5. Describe the role experimentation plays in determining causality and why it is so important.
6. What are the drawbacks of controlled experiments?
7. How is analyzing the results of an experiment different from seeing a correlation? (Please use the following words in your answer: **hypothesis** and **theory**).

Extension Activities

1. Have the students read a newspaper article of a scientific finding. Have the students determine if they have enough information from that article to determine if the scientist performing the study showed causality of two events.
2. Having done the first extension activity, have the students practice reading a scientific paper (preferably on the same scientific finding that the newspaper article was based on) and have them identify the hypothesis and if the scientists were able to determine causality between two events.
3. Have the students design an experiment to test a hypothesis and determine what they would have to show scientifically to prove causality of the variables to the outcome.

For further information:

http://usabig.com/autonomist/fallacies.html#falsff
http://www.fallacyfiles.org/examples.html
http://www.d.umn.edu/~tbacig/mindmath/mathles5.html
http://www.csicop.org/si/9809/coincidence.html
http://www.serebella.com/encyclopedia/article-correlation_implies_causation_(logical_fallacy).html

Student Handout: Cause, Effect, or Coincidence?

Name:___________________________________

A scientist's job is to try to determine causal relationships. This process often starts with an observation of some variable and an outcome: "Wow, it seems like the seeds in my garden don't grow if I don't water them." This is the discovery of what seems like a nice correlation: if you water the seeds, they grow. One may even quantify how often this happens, i.e., nine out of ten times when I added water to my garden, I got flowers. How do you go from noticing a correlation between two events to a causation argument that ties two things together as cause and effect? In a controlled experiment, you change some variable (X) and look for changes in another variable (Y).

You, as a prominent scientist, have been chosen by the publishing office of Science Magazine to peer-review the following experiment and determine if the scientists who collected the data have enough evidence to evaluate the causes of seed germination and plant survival. Recall that to prove causality, two events must be linked together as both necessary and sufficient.

A cause is **necessary** when a particular variable is required to produce an effect, although there may be other variables involved.

A cause is **sufficient** when a particular variable inevitably initiates or produces an effect. A cause can be necessary, sufficient, neither, or both.

Here is the scientist's tabulated data:

Experiment	L: Light	D: Dirt	W: Water	G: Seed germinates?	S: Plant survives 1 week?
1	Yes	Yes	Yes	Yes	Yes
2	Yes	Yes	No	No	No
3	Yes	No	Yes	Yes	Yes
4	Yes	No	No	No	No
5	No	Yes	Yes	Yes	No
6	No	Yes	No	No	No
7	No	No	Yes	Yes	No
8	No	No	No	No	No

What conclusions about the necessity and sufficiency of these elements can you draw from this data?

	L->G	D->G	W->G	L->S	D->S	W->S	G->S
Necessary?							
Sufficient?							

From this experiment, we can draw the following conclusions (write out your conclusions in complete sentences):

Analysis

On a separate sheet of paper, please complete the following:

1. Define coincidence and give some examples.
2. Define correlation and give some examples.
3. Define causality.
4. Describe and give examples of the conditions that need to be meet to determine causality.
5. Describe the role experimentation plays in determining causality and why it is so important.
6. What are the drawbacks of controlled experiments?
7. How is analyzing the results of an experiment different from seeing a correlation? (Please use the following words in your answer: **hypothesis** and **theory**).