

Variables

by Sophia



WHAT'S COVERED

This tutorial will discuss variables within the field of statistics, and introduce the concept of confounding variables. The following elements will be the main focus of this tutorial:

1. Variables

1a. Variables of Interest

1b. Explanatory and Response Variables

2. Confounding Variables

1. Variables

In statistics, a **variable** is any attribute that we can measure about a population, used in a study. It is very important to carefully define the variables to be measured when creating a study.

Think of things that we could find out about people:

- Age
- Weight
- Gender
- Ethnicity
- Favorite Food
- Number of Pets
- Smoker or Non-Smoker
- ZIP Code
- Number of Siblings
- Political Affiliation
- Favorite Sport

All sorts of these things are variables. You might only want to know one of these things or some of these things.



TERM TO KNOW

Variable

Any attribute or number that can be measured about individuals in a study.

1a. Variables of Interest

For a political poll, for example, you wouldn't necessarily need to know if a candidate was a smoker or the number pets they have. However, you might want to know about their age, gender, state, political affiliations, zip code, ethnicity, and city.

Since those variables could potentially have some bearing on a political poll. They are the **variables of interest** for this study--literally, the variables you would be interested in measuring.

However, if you were conducting a weight loss study, the political affiliation will likely not be a variable to measure, but favorite food might seem important.



TERM TO KNOW

Variable of Interest

Any variable which we need to know about in the context of a study.

1b. Explanatory and Response Variables

Some studies try to determine a cause-and-effect relationship between two variables in that one variable causes the other. An increase in one corresponds to an increase or decrease in the other.

In those cases, we define the one that causes the other as the **explanatory variable**. In a study, you can have more than explanatory variable.

Then, variables that are the result are called **response variables**.

Examples of Explanatory and Response Variables

Explanatory: Number of hours you study Response: Grade on the exam	You might hypothesize that as you increase the number of hours that you study, your grade on the exam will increase as well. So the number of hours you study, therefore, helps to explain your grade.
Explanatory: Average monthly temperature Response: Ice cream sales	You might assume that as the temperatures get warmer, that ice cream sales would go up in kind.

Something that's a little bit less obvious is whether or not gender, which is a categorical variable, plays a role in which political party people will choose. Are males more likely to be Republican? Or are women more likely to be independent voters? We don't know. But that would be an interesting question to investigate.



TERMS TO KNOW

Explanatory Variable

A variable that we believe is predictive of something else. An increase in this variable will correspond to an increase or decrease in some other variable.

Response Variable

A variable that is affected by the explanatory variable.

2. Confounding Variables

The word **confounding** refers to when two variables get mixed up with one another and you can't tell the effect of one variable from the effect of the other variable. The **confounding variable** is the one not accounted for in a study. It is an unseen variable that has a significant effect on the response variable and is also related to the explanatory variable.

IN CONTEXT

Suppose that a researcher wants to know whether a high protein diet will help lab rats gain more weight than a low protein diet. The researcher has 26 lab rats and she selects 13 of the smallest rats to receive the low protein diet and 13 of the largest to receive the high protein diet. At the end of the study, she weighs the rats to determine their weight gain and finds that the rats on the high protein diet gained more weight.

Can you think of anything that she did wrong in this study?

The answer involves the occurrence of confounding. Remember, confounding is when two variables get mixed up and you can't tell the effect of one variable from the effect of the other variable.

So in this case, the effect of the diets--whether or not the high protein diet caused the rats to gain more weight--was confounded by the fact that the heaviest rats were put on the high protein diet. It's not clear if the high protein diets were effective at weight gain. Something else may have caused the weight gain since they were heavy already.

Therefore, these are the two variables of interest in the study. The high protein diet was supposed to be the explanatory variable. The weight gain was supposed to be the response variable. The

researcher was going to try to figure out a link between the two.

However, because of the way she assigned the rats, only a limited conclusion could be drawn. She wasn't able to draw the direct conclusion that she was hoping for--and that is confounding. Confounding should be limited in experiments when possible.



TRY IT

A high school math teacher, hoping to have his students do well on the final, offers an optional review session. He states, "No one who's ever attended the review session has ever scored less than a B".

What is the teacher trying to imply? Why isn't his implication correct?

You may have come up with that he's trying to imply that the review sessions will cause the students to do better. That may be true; however, there may be a few confounding variables. Maybe only his best and brightest students attend the optional review and these are students that may have done well on the final exam anyway. The effects, if any, are confounded by the intrinsic motivation of students to show up to the session.



TERMS TO KNOW

Confounding

Occurs when the effects of the treatments, if any, are indistinguishable from the potential effects of some other variable which was unaccounted for.

Confounding Variable

A variable which was not accounted for in a study, which limits the conclusions that the study can draw.



SUMMARY

Variables are what we choose to measure in a study. The variables of interest will depend on the questions that you're trying to answer. Not every variable must be measured--just the ones that are of interest. By looking at variables in context, you learned that if a cause and effect relationship is thought to exist, you can break the variables down even further into explanatory and response variables. Confounding occurs when there is a variable that is chosen as an explanatory variable in an experiment, but because another variable got in the way, it cannot be determined to explain a cause. You explored confounding variables in action to demonstrate how they can limit the conclusions that can be drawn from the supposed explanatory variable. In effect, the confounding variable inhibits a cause-and-effect conclusion. Often, it's one that you didn't think to measure, which is problematic.

Good luck!



TERMS TO KNOW

Confounding

Occurs when the effects of the treatments, if any, are indistinguishable from the potential effects of some other variable which was unaccounted for.

Confounding Variable

A variable which was not accounted for in a study, which limits the conclusions that the study can draw.

Explanatory Variable

A variable that we believe is predictive of something else. An increase in this variable will correspond to an increase or decrease in some other variable.

Response Variable

A variable that is affected by the explanatory variable.

Variable

Any attribute or number that can be measured about individuals in a study.

Variable of Interest

Any variable which we need to know about in the context of a study.