Confounding

1000: Introduction

In May 2000, researchers in Pittsburgh found that children convicted of delinquency had much higher bone-lead levels than youngsters who had no juvenile convictions. In other words, there is an association between **JUVENILE CRIME** [Yes, No] and **BONE-LEAD** [High, Low]. At the same time, a study in Cincinnati found that, as blood lead levels increased in children, their reading and math scores decreased. So, there is also a negative association between **BLOOD LEAD** and **TEST SCORES**.

Given that we would like to reduce juvenile crime and increase test scores, we might naturally think that we should try to reduce the amount of lead in children's blood and bones. However, as we have discussed before, simply finding an association is not sufficient for drawing a causal conclusion. Two variables can be associated without any causal connection between them, in which case an intervention to reduce lead levels will have no effect at all. Can we determine ahead of time whether our intervention will be successful? What do we need to know/measure to estimate the effect (if any) of our intervention?

Consider another example. Suppose we collect data on 5,000 40-year-old Americans. We record the highest educational level that each achieved (**EDUCATION** [HS, College, Graduate Degree]), and we record each person's current income (**INCOME** [Low, Medium, High]). Suppose we find that **EDUCATION** and **INCOME** are associated. Is **EDUCATION** a cause of **INCOME**? We can even suppose that, for each individual, **EDUCATION** takes on its value before **INCOME**, so we can rule out the possibility that **INCOME** causes **EDUCATION**. But given just this association, we don't know whether **EDUCATION** is a cause of **INCOME**, or whether there is an unmeasured/unobserved variable (such as **HARD WORK**) that is a common cause of both **EDUCATION** and **INCOME**. In other words, the association alone doesn't tell us which of the following graphs is correct:

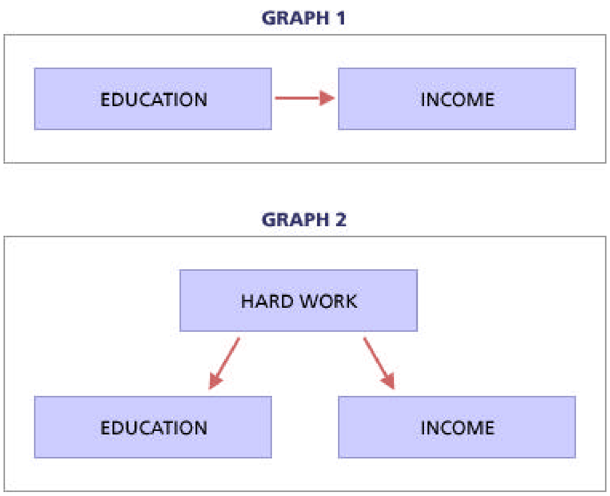

FIGURE 1000-1

In the module "Causal Prediction vs. Causal Discovery", we discussed the different ways in which associations underdetermine causal theories. This underdetermination is what makes the questions asked earlier hard. One specific way underdetermination can arise, and the way we will focus on in this module, is through the presence of unmeasured or unobserved common causes, which we call **confounders**.

This module is all about qualitative confounding, and the problem confounders present for causal discovery. We will explain exactly what confounders are, and how they pose a problem. After that, we briefly introduce a strategy for dealing with confounding: conditioning on them. There are limits to this strategy, however, and we will look at when and why conditioning on confounders is misleading.

2000: Confounding

2100: The General Problem

We are often trying to learn a particular causal structure because we want to know how to control something about our world. We want to get a particular variable to have a different value, and we need to know what interventions would lead to that change. However, we typically only have information about associations, and associations can be due to something other than one of the variables causing the other. In an earlier module, we learned that a common cause produces an association between its effects. In other words, it is possible for two variables to be associated, but for any intervention on one of the variables to have no effect on the other variable.

For example, suppose the following causal graph accurately depicts the causal relations among **EDUCATION**, **INCOME**, and **DOMESTIC VIOLENCE**:

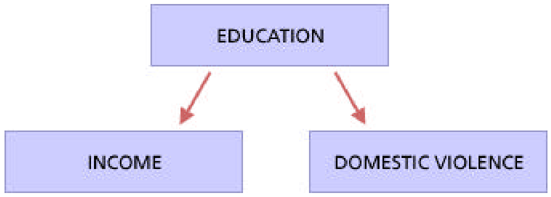

FIGURE 2100-1

This graph predicts that **INCOME** and **DOMESTIC VIOLENCE** are associated, but this association is (totally) confounded by **EDUCATION**. If a scientist observed the association and concluded that poverty causes domestic violence, her conclusion would be incorrect because of the confounding common cause **EDUCATION**.

What kind of factors can lead to a misleading association? Remember that an association is produced by a causal connection. **X** and **Y** are causally connected just when:

1. There is a causal path from **X** to **Y**, or **Y** to **X**; or
2. **X** and **Y** have a common cause.

If there is a causal path from **X** to **Y** or from **Y** to **X**, then that means that one variable is a cause (possibly indirect) of the other, so a causal path can't confound the association between two variables. However, a common cause can.

Definition: Confounding

The association between two variables is confounded whenever there is an unmeasured common cause. The common cause is called the confounder.

In this module, we will specifically be concerned with the problems confounding causes when we are trying to determine whether one variable is a cause at all of the other variable. Confounding also creates problems when we are trying to estimate the strength of the effect of an intervention; we will refer to those problems as "quantitative confounding," and we will leave those problems for another module. Here we are just concerned with "qualitative confounding," in which the presence/absence of a causal path is all we are asking about. We will find that there are plenty of issues, even with this restricted view.

In order to get a handle on how confounding can be a problem, here are two examples in which there might be a confounder that explains the association between two variables. Both of these cases are probably more complicated in real life, but we simplify them here for the purposes of demonstration. We'll consider more complicated examples later in the module.

Example 1

We know that playing violent video games is associated with committing violent crimes. We might explain this association by claiming that playing violent video games is a direct cause of committing violent crimes. This would be one way of explaining the association between the two variables. However, the association between these variables could also be due to a confounder (an unmeasured common cause), such as natural aggressiveness (or some trait like this). In other words, the graph might be:

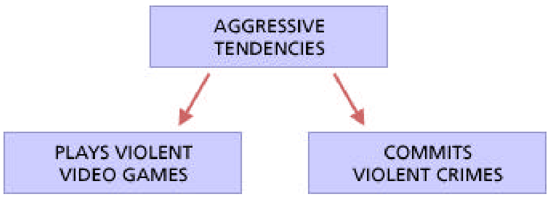

FIGURE 2100-2

In this graph (with a confounder), playing violent video games is not a cause of committing violent crimes. Instead, the association is entirely due to the confounding variable of **AGGRESSIVE TENDENCIES**. Therefore, changing whether someone plays violent video games would have no effect on whether the person commits violent crimes.

Example 2

As another example, let's assume that whether someone has a child before the age of 20 is associated with that person's career success. This might lead us to think that having a child before the age of 20 causes career success; that would explain the association between the variables. However, it's possible that this association is entirely due to the person's socioeconomic status as a child. So, the graph might be:

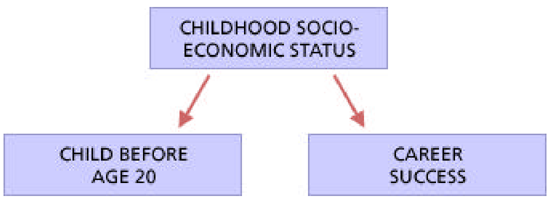

FIGURE 2100-3

Although **CHILD BEFORE AGE 20** and **CAREER SUCCESS** are associated, neither one causes the other in this graph. The association could be entirely due to the person's earlier socioeconomic status.

2200: Case Studies

2210: Lead Exposure (again)

In the Introduction, we described a study which examined the association between various measures of exposure to lead (which we will just call **LEAD EXPOSURE**) and other variables. Specifically, the studies found that:

LEAD EXPOSURE is associated with **JUVENILE CRIME**

We can make the rather safe assumption that **JUVENILE CRIME** is not a cause of **LEAD EXPOSURE**, since the exposure must have come at a much younger age. But it is quite natural to think that the exposure to lead might be a cause of juvenile crime.

Consider a study that shows that **LEAD EXPOSURE** is associated with **JUVENILE CRIME**. Here is an excerpt from a Associated Press article about the study:

"Lead is a toxic metal that can interfere with development of the central nervous system and can be detected in blood and bones. Severe lead poisoning, which can cause seizures and even death, can sometimes be treated with medication, but Lanphear said the more subtle declines in mental functioning linked to lead are persistent and may be permanent.

"In a study of 417 youngsters in Allegheny County, Pa., significantly higher bone-lead levels were found in those convicted of delinquency than in a comparison group, reported Dr. Herbert Needleman of the University of Pittsburgh.

"Among boys, convicted juveniles were nearly twice as likely to have high bone-lead levels as the youngsters in the comparison group. The risk for girls was even higher, partly because only a small number of female delinquents - 21 - was studied, Needleman said.

"The findings suggest a possible link between early lead exposure and 11 percent to 37 percent of arrested delinquents, said Needleman, whose previous research linked aggressive and anti-social behavior to lead.

"Lead exposure may be one of the most preventable causes of criminal behavior, he said."

In short, the data in the study showed an association between lead exposure and juvenile crime.

In answering the questions below about this case, you need to know that the two primary modes of exposure to lead are 1) lead-based paint that children ingest when they live in houses with such paint on the walls and woodwork and 2) gasoline with lead additives, which was sold in the US until about a decade ago.

< A link to exercises in the interactive version of this module. >

2220: Drop in Crime Rates

Throughout most of the 1990s, crime rates have decreased in the United States. In 1999, two researchers (Steven D. Levitt of the University of Chicago and John J. Donohue III of Yale University) gave a rather startling potential explanation for this decrease: that the legalization of abortion in the early 1970s had led to fewer unwanted children, who are more likely to grow up and commit crimes than other children. The point of this example is not to have a debate about the morality of abortion. Rather, we want to look at the kind of evidence that Levitt and Donohue put forth for their conclusion, and we want to ask: is confounding potentially a problem for these researchers?

Levitt and Donohue found the following associations:

1. States that legalized abortion before Roe vs. Wade saw their crime rates drop earlier than the country as a whole; and
2. States with higher abortion rates have had steeper drops in crime

We can summarize these associations as: **ABORTION RATE** is negatively associated with **CRIME RATE 18 YEARS LATER**. Clearly the later crime rate does not cause the earlier abortion rate, and so we can rule out one group of graphs (those with **CRIME RATE 18 YEARS LATER → ABORTION RATE**). Levitt and Donohue's explanation for the association is:

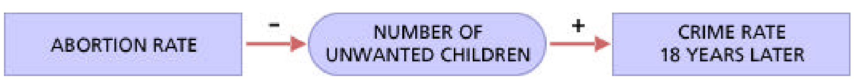

FIGURE 2220-1

The questions below deal with possibile counfounders.
< A link to exercises in the interactive version of this module. >

3000: Controlling for Confounders

3100: The General Theory

The previous sections introduced the concept of a confounder, and showed that associations can result between two variables when neither one causes the other. In all of those examples, we found alternate explanations for the observed associations, and so we realized that it is possible that neither variable caused the other. However, in all of those examples, it was also possible that one variable actually was a cause of the other.

Consider the earlier example of the association between lead exposure and both test scores and juvenile crime. As part of that example, we found a number of other variables that might be confounders. However, it's also possible that the scientists' conclusion is correct: lead exposure might be a cause of lower test scores and higher juvenile crime. Looking at just one effect, we can ask: can we determine if **LEAD EXPOSURE** actually is a cause of **TEST SCORES**?

It turns out that we **can** figure out whether one variable is a cause of the other, but only if, for each confounder, we measure (and then condition on) some variable that eliminates the association produced by that confounder. (Later in this section, we'll talk about what happens if we don't account for all of the confounders.) Let's imagine that the actual graph for the **LEAD EXPOSURE**/**TEST SCORES** case is:

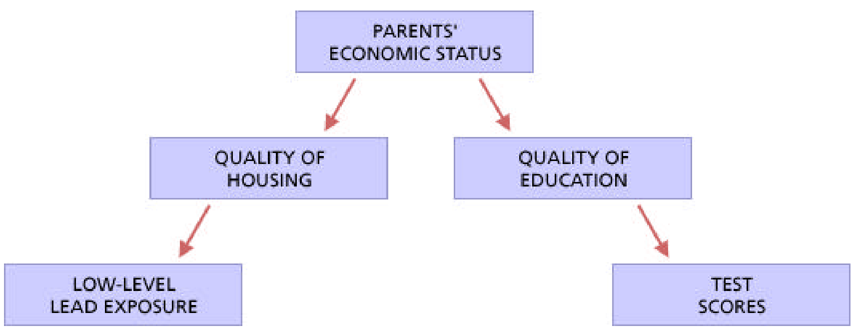

FIGURE 3100-1

This graph says that **PARENTS' ECONOMIC STATUS** is a confounder of the association between **LEAD EXPOSURE** and **TEST SCORES**, since the association between them is entirely due to the common cause. Assuming that this graph is correct, intervening to change the level of lead exposure will have no effect on children's test scores. But how can we predict that fact from only associational data? The idea is that we need to look at conditional associations also. Notice that this graph predicts:

1. **LEAD EXPOSURE** _X_ **TEST SCORES**
2. **LEAD EXPOSURE** _||_ **TEST SCORES** | **PARENTS' ECONOMIC STATUS**

The conditional independence holds because conditioning on the common cause (**PARENTS' ECONOMIC STATUS**) blocks the causal connection. If **LEAD EXPOSURE** **is** a cause of **TEST SCORES**, though, then the two variables will still be associated, even if we condition on **PARENTS' ECONOMIC STATUS**, since there will still be a causal connection that produces association: the causal path from **LEAD EXPOSURE** to **TEST SCORES**. So, if we measure the value of **PARENTS' ECONOMIC STATUS** for every individual in our population and it is the only confounder, then we can determine whether **LEAD EXPOSURE** is a cause of **TEST SCORES** by looking at the conditional association/independence.

In general, if the association between two variables is entirely due to confounders, then if we condition on all of the confounders, the two variables will be independent, since we will be blocking all of the causal connections between the two variables. If, on the other hand, one of the variables is a (direct or indirect) cause of the other variable, then if we condition on all of the confounding variables, the two variables will still be associated, since they will still be causally connected. In this second case, conditioning on the confounding variables eliminates some of the causal connections that produce associations, but not all of them.

Moreover, notice that we don't even need to condition on the confounders themselves. Rather, we just need to block the common cause's causal connection from producing an association. We can do this by conditioning on any variable, not just the common cause, on that causal connection (actually, there's an exception, but we'll talk about that later). So, in the lead and test scores example, we see that the following conditional independencies also hold:

1. **LEAD EXPOSURE** _||_ **TEST SCORES** | **QUALITY OF EDUCATION**
2. **LEAD EXPOSURE** _||_ **TEST SCORES** | **QUALITY OF HOUSING**

These conditional independencies also tell us that **LEAD EXPOSURE** is not a cause of **TEST SCORES** (assuming that we know that the other variables are part of the confounder's causal connection). So, there are multiple variables we can measure to determine whether **LEAD EXPOSURE** is a cause of **TEST SCORES**.

This ability to control for a single confounder using multiple variables can be particularly important in some situations. Consider the earlier example of **PLAYS VIOLENT VIDEO GAMES** and **COMMITS VIOLENT CRIMES**, and suppose the association is entirely due to a confounder: **AGGRESSIVE TENDENCIES**. However, it is unclear how we could measure a variable such as **AGGRESSIVE TENDENCIES**. Perhaps we could put someone in many different situations, and see how they respond. But it's doubtful that this idea would be a particularly accurate measure of the variable. So although the association is entirely due to the confounder, we can't discover that, since we can't condition on the confounder itself.

However, suppose that **AGGRESSIVE TENDENCIES** is a cause of **COMMITS VIOLENT CRIMES** through the mediator **PUTS ONESELF IN DANGEROUS SITUATIONS**. If we adopt the convention that variables in ovals are unmeasured, we can represent this situation as:

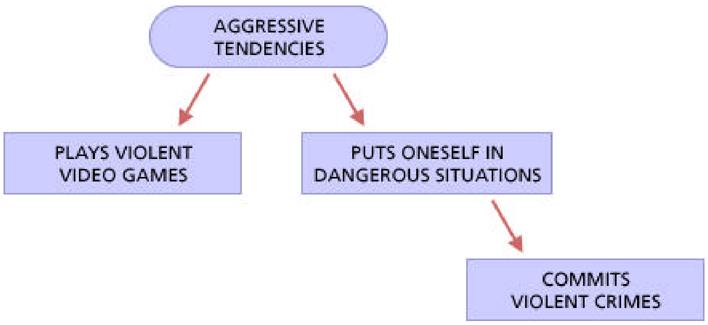

FIGURE 3100-2

Although it would not necessarily be practical to measure this new variable, it would certainly be easier and more ethical. And if the graph is as pictured above, then we would be able to control for the unmeasured confounder by conditioning on the measured mediator (on the confounder's causal connection).

So, we have found a way to use only associational data to determine whether the association between two variables is entirely due to confounders (though we must measure all of the confounders). The general strategy is:

To determine whether one variable is a (direct or indirect) cause of another, control for all of the confounding variables (by conditioning on at least one variable on each causal connection produced by a confounder). If the original variables are associated after controlling for the confounders, then one variable is a cause of the other. If they are independent after controlling for confounders, then neither is a cause of the other.

3200: Controlling in Practice

The process of looking at conditional associations/independencies to determine whether one variable is a cause of another is called **controlling for the confounders**. Don't be confused by the name into thinking that controlling for confounders is an intervention, though. We aren't intervening on the confounders and then checking the associations (since the confounder is an exogenous variable, and so the intervention wouldn't change the causal graph at all). Rather, we're just looking for the same conditional independencies that we covered in the "Conditional Independence" module. In other words, we're asking whether the two variables are still associated among those people who all (naturally) have the same values for the confounders.

Let's look at an abstract example to see just how this process works. Consider the simplest case of confounding: a single common cause of two variables (so the variables are unconditionally associated). Furthermore, let's suppose for simplicity that **Z** is the only confounder, and that **X** occurs before **Y**. Now we want to ask: is **X** a cause of **Y**?

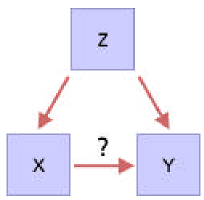

FIGURE 3200-1

Using the strategy from the previous section, we need to ask whether **X** _||_ **Y** | **Z**. If they are independent, then their (unconditional) association must be entirely due to the confounder **Z**, and so **X** is not a cause of **Y**. If they are conditionally associated, then since we assumed that **Z** is the only confounder, the only causal connection left to explain the conditional association is that **X** is a cause of **Y**. Now, since we are asking about **variable** independence, we need to check whether every value of **X** is independent of every value of **Y**, conditional on every value of **Z**. If the variables are all binary, we just need to check:

1. Fr(**X** = Yes | **Z** = Yes) = Fr(**X** = Yes | **Y** = Yes, **Z** = Yes); and
2. Fr(**X** = Yes | **Z** = No) = Fr(**X** = Yes | **Y** = Yes, **Z** = No).

If both of these equalities are satisfied, then **X** is not a cause of **Y**. If either one fails to be satisfied, then **X** is a cause of **Y** (assuming **Z** is the only confounder). In this particular example, we only need to check two equalities. Depending on how many confounders we have, though, we might need to check many more equalities (since we have to check for variable independence).

A recent study showed an association between being directly exposed to violence (seeing violent acts on person) and commission of violence crime. The following questions ask you to consider possible confounding variables in this case and how to control for them.

< A link to exercises in the interactive version of this module. >

Of course, before we can measure them, we need to know which variables are confounders. We've already seen that our procedure of "Control for confounders; 'Independent = No cause' and 'Associated = Cause'" won't work if we miss a confounder, since that confounder's causal connection will guarantee that the variables are associated. But what if we control for a non-confounder? That's the subject of the next sections.

4000: Limitations of Controlling for Confounders

4100: Controlling for Mediators

In the previous sections, we developed a general strategy for determining whether one variable is a cause of another. But this strategy assumed that we knew which variables were confounders, and we often don't know whether a variable is a confounder or not. Typically we just measure a set of variables without knowing all of the causal relations among the variables. We might hope that we could just change the strategy slightly: rather than controlling only for confounders, we could just control for all of the variables we measure. If this strategy worked, we would no longer need to know ahead of time which variables are confounders (though we would still need to make sure we controlled for all of them). Unfortunately, this proposed strategy won't work, and this section and the next one are about when and why it fails.

Consider a system with three variables: **LEAF COVER** [Thick, Thin], **PLANT HEALTH** [Healthy, Sick], and **ROOT SYSTEM** [Large, Small]. Let's assume that **PLANT HEALTH** and **ROOT SYSTEM** are associated. We would like to know whether **ROOT SYSTEM** is a cause of **PLANT HEALTH**, but we might be worried that **LEAF COVER** is a confounding variable. In other words, we might be think that the causal relations are:

FIGURE 4100-1

From the last section we know that, if **LEAF COVER** actually is a confounder of the two variables, then controlling for **LEAF COVER** will correctly tell us whether **ROOT SYSTEM** does not cause **PLANT HEALTH**. However, if we don't know whether **LEAF COVER** actually is a confounder, then even if **ROOT SYSTEM** _||_ **PLANT HEALTH** | **LEAF COVER**, we still don't know that **ROOT SYSTEM** is not a cause of **PLANT HEALTH**. Why? Because the actual causal graph might be:

FIGURE 4100-2

If this is the correct causal graph, then **ROOT SYSTEM** and **PLANT HEALTH** will be independent conditional on **LEAF COVER**, even though **ROOT SYSTEM** is a (indirect) cause of **PLANT HEALTH**. This is a case in which it is very important to know whether **LEAF COVER** is a confounder, so that we can correctly interpret the conditional independence. In general, if we control for a mediator (when we think it's a confounder), then we can get the wrong answer about whether one variable is a cause of the other.

For a more abstract example, consider the following graph:

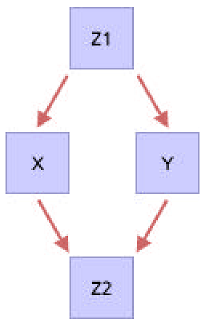

FIGURE 4100-3

We know that **X** _||_ **Y** | **Z1**, **Z2**, and so we might conclude that **X** is not a cause of **Y**, even though it is actually an indirect cause of **Y**. So we can see that controlling for every variable can give us the wrong answer if at least one of the variables is a mediator on a causal path from **X** to **Y** (or **Y** to **X**).

< A link to exercises in the interactive version of this module. >

4200: Controlling for Common Effects

We just saw that the strategy "control for everything" will fail if one (or more) of the variables is a mediator between the two variables we're considering, because we'll incorrectly think that neither variable is a cause of the other. The strategy also fails if one (or more) of the variables is a common effect of the two variables, because then we can incorrectly conclude that one variable **is** a cause of the other.

Consider the variables: **EXERCISE**, **BODY WEIGHT**, **FOOD EATEN**, and **METABOLISM**. We will assume that **EXERCISE** is associated with **BODY WEIGHT**, but we want to know whether **EXERCISE** is a **cause** of **BODY WEIGHT**. **FOOD EATEN** and **METABOLISM** could be confounders, and so we might think that we should control for them. However, the following graph might represent the actual causal relations:

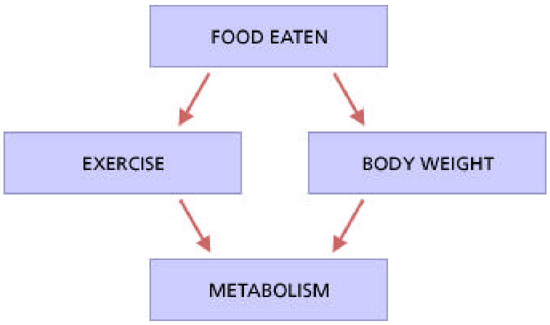

FIGURE 4200-1

If we condition on **FOOD EATEN** and **METABOLISM**, then **EXERCISE** and **BODY WEIGHT** will still be associated, since we're conditioning on a common effect (**METABOLISM**). So, using the "control for everything" strategy introduced in the previous section, we would (incorrectly) conclude that **EXERCISE** is a cause of **BODY WEIGHT**. This is a case where "controlling for everything" would give us the wrong answer because we would hypothesize a causal relation where none exists.

For an abstract example that integrates the lessons of the this section and the previous one, consider this graph:

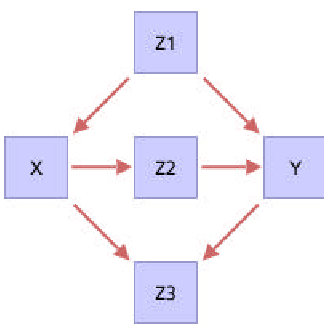

FIGURE 4200-2

If we want to know whether **X** causes **Y**, we should only control for **Z1**. If we also control for **Z2** (a mediator), we'll think that **X** is not a cause of **Y**, even though it is. If we control for **Z3** (a common effect), we'll correctly conclude that **X** is a cause of **Y**, but for the wrong reasons. So, it's important to control for all and only the confounding variables. Unfortunately, we often don't know ahead of time which variables are confounders.

< A link to exercises in the interactive version of this module. >

5000: Summary

There are many reasons why we want our causal hypotheses to be correct. Associational information is an important window onto causal information, but we need to be careful about how we use it. If one variable causes another, then the two will be associated. However, that same association could also arise because of a confounder: an unmeasured/unobserved common cause. Two variables can be associated (which would suggest causation) even though neither one is a cause of the other.

We can account for the problem of confounding by controlling for the confounding variables. To control for the confounders, we just need to condition on them. We look at groups of individuals who all have the same value for the confounders; we don't have to intervene on the confounders. In fact, we don't actually have to condition on the confounders themselves. Rather, we just need to ensure that the confounder's causal connection is blocked.

However, we have to be careful about which variables we control for. If we control for a mediator, then we might mistakenly conclude that neither variable is a cause of the other. On the other hand, if we control for a common effect, then we can decide that one variable **is** a cause of the other, even if neither variable actually does cause the other. These potential problems can be quite significant if we don't have substantial theoretical knowledge that allows us to rule out certain graphs as impossible.

Throughout this module, we've only considered the case in which confounding potentially accounts for **all** of the association between two variables. In other words, we've assumed that we just want to know whether one variable is a cause (regardless of strength) of another variable. Sometimes, though, we want to know the strength with which one variable causes another, and confounding arises there as well. We've left that topic for another module, but it's worth pointing out that all of the same strategies and problems arise in that situation as well. Confounding can be overcome, but only with substantial knowledge.
