

--- JON BRAUER ---

Is it possible to have causation without correlation?

--- CHAT GPT ---

Yes, it is possible to have causation without correlation. While causation and correlation are often related, they are not the same thing.

Correlation refers to a statistical relationship between two variables, where a change in one variable tends to be associated with a change in the other variable. However, correlation does not necessarily imply causation, because there could be other factors at play that influence both variables. For example, there is a correlation between the number of ice cream cones sold and the number of drownings each year, but this does not mean that ice cream causes drownings.

Causation, on the other hand, refers to a relationship where one event (the cause) brings about another event (the effect). A causal relationship can exist even if there is no correlation between the two variables, because the relationship is not dependent on a statistical pattern. For example, smoking causes lung cancer, but not everyone who smokes gets lung cancer, and not everyone with lung cancer smokes.

In summary, while correlation and causation are often linked, they are not the same thing, and it is possible to have causation without correlation.

--- JON BRAUER ---

In your smoking example, if smoking causes lung cancer, wouldn't there also be a statistical relationship (i.e., a correlation) between smoking and lung cancer? The statement that "not everyone who smokes gets lung cancer, and not everyone with lung cancer smokes" does not rule out the possibility (or the likelihood) of a statistical relationship - it only rules out the possibility of a perfect correlation.

--- CHAT GPT ---

You are correct that in the case of smoking causing lung cancer, there is typically a statistical relationship (i.e., a correlation) between smoking and lung cancer. In fact, smoking is one of the strongest risk factors for developing lung cancer, and studies have consistently shown a strong correlation between smoking and lung cancer.

My previous statement about not everyone who smokes getting lung cancer and not everyone with lung cancer smoking was intended to highlight the fact that while there is a strong correlation between smoking and lung cancer, it is not a perfect correlation. In other words, there are some people who smoke and do not develop lung cancer, and there are some people who develop lung cancer even though they have never smoked.

So while a correlation between smoking and lung cancer does exist, it is not a one-to-one relationship, and it is possible for some smokers to avoid lung cancer and for some non-smokers to develop lung cancer. However, this does not negate the fact that smoking is a significant risk factor for developing lung cancer, and the statistical relationship between the two remains strong.

--- JON BRAUER ---

Yes, but the question was whether it is possible to have causation without correlation. You gave an example of causation with correlation.

--- CHAT GPT ---

You are correct, my apologies for any confusion. The smoking example I gave was indeed an example of causation with correlation.

To provide an example of causation without correlation, imagine a scenario where a scientist discovers that a specific gene mutation causes a rare disease that affects only a small portion of the population. In this case, there may not be a statistical relationship or correlation between the gene mutation and the disease in the general population, because the disease is so rare and the gene mutation is only found in a small portion of the population. However, there is a clear causal relationship between the gene mutation and the disease.

So, to answer the original question more accurately, it is possible to have causation without a statistically significant correlation between the two variables in the general population, as in the rare disease example described above.

--- JON BRAUER ---

In that example, again, there is a correlation between the gene mutation and the disease. I did not ask about whether there could be causation without a statistically significant correlation - statistical significance depends on power of the test. So, in this causal example, if you had a sufficiently large sample, the correlation would also be detected, making it another example of causation with correlation.

--- CHAT GPT ---

You are correct, and I apologize for the confusion. In the scenario I described, there is indeed a correlation between the gene mutation and the rare disease, even though it may not be statistically significant in the general population due to the low prevalence of the disease.

It is true that statistical significance depends on the power of the test and the sample size, and it is possible that a correlation between two variables could become statistically significant with a larger sample size. Therefore, in this case, with a larger sample size, the correlation between the gene mutation and the rare disease would likely be detected, and it would be an example of causation with correlation.

Thank you for pointing out this mistake in my previous responses, and I hope this clarification is helpful.

--- JON BRAUER ---

So, can you think of another example of causation without correlation that does not involve an insufficiently powered test for detecting correlation?

--- CHAT GPT ---

Yes, I can provide another example of causation without correlation that does not involve an insufficiently powered test for detecting correlation.

Consider the example of a person who takes a medication that causes a significant reduction in their symptoms of a medical condition. In this case, there is a clear causal relationship between taking the medication and the improvement in symptoms, but there may not be a statistical correlation between the two variables in the general population.

This lack of correlation could be due to a number of factors, such as individual differences in the severity or type of the medical condition, or differences in the dosage or timing of the medication. Additionally, the relationship between the medication and the symptoms could be confounded by other factors, such as the person's diet, exercise habits, or stress levels.

So, in this scenario, while there may not be a statistically significant correlation between taking the medication and the improvement in symptoms in the general population, there is a clear causal relationship between the two variables in the individual person who experienced the improvement in symptoms.

--- JON BRAUER ---

This example again leverages a lack of statistical significance, which I am unmoved by such arguments. For instance, here, there is causation and likely correlation - but the correlation is not detectable due to improper modeling due to issues like heterogeneity in the strength of the effect, confounding, poor measurement, or inadequately specifying dosage and time. Can you think of any other examples that are different from these?

--- CHAT GPT ---

Thank you for the feedback and clarification. I will provide another example of causation without correlation that is different from the previous examples.

Consider the example of a person who develops a fear of flying after experiencing a traumatic event on a plane, such as severe turbulence or an emergency landing. In this case, the traumatic event on the plane is the direct cause of the person's fear of flying.

However, there may not be a statistical correlation between traumatic events on planes and the development of a fear of flying in the general population. This is because not everyone who experiences a traumatic event on a plane will develop a fear of flying, and some people may develop a fear of flying for other reasons unrelated to traumatic events.

Therefore, in this scenario, there is a clear causal relationship between the traumatic event on the plane and the person's fear of flying, but there may not be a statistically significant correlation between the two variables in the general population.

--- JON BRAUER ---

Interesting. If it was a singular event, then would there not be a perfect correlation between flight exposure and fear of flying since there are two data points (pre-exposure level of fear & post-exposure levels of fear)? If it is not a singular event and there is some proportion of the population that develops fear of flying after exposure to flying, then again we are back to the rare genetic cause of disease example - with sufficient data, we would be able to detect this causally driven correlation in the population. Any other ideas?

--- CHAT GPT ---

You make a good point, and I apologize for any confusion. The example I provided may not be a good example of causation without correlation.

Another example of causation without correlation could be the case of a person who develops an allergic reaction to a specific food after consuming it for the first time. In this case, the food is the direct cause of the allergic reaction, but there may not be a statistically significant correlation between consuming the food and developing an allergic reaction in the general population.

This lack of correlation could be due to individual differences in genetics or immune system function, or differences in the amount or preparation of the food consumed. Additionally, the relationship between the food and the allergic reaction could be confounded by other factors, such as the person's age or other underlying health conditions.

Therefore, in this scenario, while there may not be a statistically significant correlation between consuming the food and developing an allergic reaction in the general population, there is a clear causal relationship between the two variables in the individual person who experienced the allergic reaction.

--- JON BRAUER ---

Again, this example seems to have a very similar logical structure and empirical implications of your previous examples. I am going to write a blog entry on causation without correlation that covers a different type of example. I will share the idea with you in case it helps you answer questions like this. Another reason there might be causation without correlation is that an exposure affects multiple mechanisms that have countervailing effects on an outcome. For instance, the exposure might be causally related to the mediators and correlated with the mediators, but the mediators are causally related to the outcome and correlated with the outcome but in opposite directions, thereby resulting in a null (negated or cancelled out) correlation in the population. There would still be causation, though, since one could imagine intervening on one mediating pathway to hold it constant, then the exposure would cause change the outcome - and be correlated with the outcome - through the other mediating pathway. Have you encountered any examples like this before?

--- CHAT GPT ---

Thank you for sharing your idea. That is a very interesting perspective on causation without correlation, and it makes sense that an exposure affecting multiple mechanisms could result in countervailing effects on an outcome, leading to a null correlation in the population.

I have not encountered this specific example before, but I can see how it could arise in situations where there are complex and interrelated causal pathways between an exposure and an outcome. For example, in a study investigating the effects of a particular medication on a disease outcome, the medication may affect multiple biological pathways that have opposing effects on the disease outcome. As a result, the overall correlation between medication use and disease outcome may be null or weak, even though the medication is causing changes in the underlying biological pathways and ultimately influencing the disease outcome.

I hope this example helps to provide a different perspective on causation without correlation, and thank you for bringing up this interesting topic.