Dependence, Causal and Probabilistic

It is useful to start by distinguishing between *causal* dependence and *probabilistic* dependence.

- Two events are **causally independent** if neither of them is a cause of the other; otherwise, the effect is **causally dependent** on the cause.
- Two events are **probabilistically independent** if the assumption that one of them occurs does not affect the probability that the other one will occur; otherwise, each of them is **probabilistically dependent** on the other.

The umbrella example can be used to illustrate the difference between these two types of dependence:

- There being wet sidewalks is *probabilistically dependent* on umbrella use, because the assumption that people are using umbrellas increases the probability of wet sidewalks.
- There being wet sidewalks is *causally independent* of umbrella use, because umbrella use does not *cause* sidewalks to be wet and wet sidewalks do not cause umbrella use. (What we have instead is a *common cause*: rain causes both wet sidewalks and umbrella use.)

Something similar happens in the case of Newcomb's Problem:

- Whether or not the large box contains a million dollars is *probabilistically dependent* on your choice to one-box or two-box, because the assumption that you one-box increases the probability that the large box contains the money.
- Whether or not the large box contains a million dollars it is *causally independent* from your choice to one-box or two-box, because your action doesn't cause the box to have the money in it, and the amount of money in the box does not cause your action. (Here too we have a *common cause*: your psychological constitution causes both your decision and the predictor's prediction.)

Problem 1

2/2 points (ungraded)

Alice and Bob make independent decisions to go for a walk in the park on Tuesday afternoon, and neither of them brings an umbrella. Let A be the event of Alice's getting soaked in rain, and B be the event of Bob's getting soaked in rain. Are A and B probabilistically independent?

- ○ Yes.
- ◉ No. ✔

Are they causally independent?

- ◉ Yes. ✔
- ○ No.

Explanation

The occurrence of A would make it much more likely that B occurs, since it raises the probability that it rained and therefore that they both got soaked. So A and B are not probabilistically independent.

Since A and B made their decisions independently, we have been given no reason to think that A's occurrence would cause B to occur, or that B's occurrence would cause A to occur. (If either of these events occurs, it will presumably be caused by the relevant subject's decision go for a walk without an umbrella, and by the presence of rain.) So we can expect A and B to be causally independent.

Submit

ℹ Answers are displayed within the problem

Problem 2

2/2 points (ungraded)

Let S be the event of Jones's smoking, and C be the event of Jones's being diagnosed with lung cancer. Are S and C probabilistically independent?

- ○ Yes.
- ◉ No. ✔

Are they causally independent?

- ○ Yes.
- ◉ No. ✔

Explanation

Smoking is a cause of lung cancer. So C is not causally independent of A.
Because smoking is a cause of lung cancer, Jones's smoking increases the probability that he will be diagnosed with lung cancer. So C and A are not probabilistically independent.

Submit

Answers are displayed within the problem

Discussion

Topic: Week 5 / Dependence, Causal and Probabilistic

Hide Discussion

Add a Post

All Posts

Does smoking cause cancer?

discussion posted 4 days ago by **ericcoker1974**

I really like the terminology introduced here, as it plays right into one of my favorite pet peeves : people confuse causation with correlation all the time. It actually helps explain why I'm a 1-boxer. I'm not that attached to causation because of how I think about correlation and explanation.

The smoking question is case in point. It is generally accepted that smoking causes cancer. However, what we really have is a very good correlation between smoking and cancer, combined with a very good explanation of why it correlates.

So we give it the label of "cause." This is clearly better than seeing a bunch of people with umbrellas out and concluding they caused it to rain. You could come up with an explanation, but it would be a bad one.

The difference between causation and correlation, therefore, is a problem of epistemology. We can agree on the empirical evidence, which would be the correlation, but can we always agree on the explanation? In the case of smoking and cancer, we mostly do, smokers being the usual exception, but that doesn't make it 100% true. In fact, scientific thinking should forbid us from ever assuming that we are 100% correct about anything. We always have to think about things probabilistically. And the probability that smoking causes cancer is high. But you could also say one of the other causes is living past 40. It is the understanding of the intersectionality of all the correlations that helps us come up with better and better explanations, and allows us to approach causation.

I'm not saying anything about the ontology of causation here. Causation could exist. I'm just saying that what we normally call causual dependance (like smoking and cancer) is just highly probabilistic dependence. True causation is the limit of having infinite amount of data and a perfect explanation.

Despite the fact that some older people have smoked all their lives and don't have cancer and there are other people that treat their body like a temple and do have cancer, I don't smoke. Probabalistic dependance is good enough for me! But for many people, they want certainty. And this causes, oops, I mean correlates, to all sorts of poor decision making.

This post is visible to everyone.

Add a Response

1 response

Jimbof
4 days ago

You are making an important distinction in your post! Thanks for bringing this question into consideration.

In the case of cancer I read an article explaining the process causing cancer from smoking for stomachal and esophagichal cancer, and conclusions were validated by statistical correlations. I belive there are similar studies on Lung cancer. Unfortunately I lost the links to those articles, so you may just take my comments as a mere possibility, my point being I believe nowadays we have a complee theory explaining the process of cause-effect between smoking and cancer (at least for some kind of cancer).

In the other hand sometimes statistical correlations are considered to be enough as demonstration on cause - effect relationship. In the case of umbrellas and wet streets it's clear the has to be a common cause for both, but sometimes a common cause is out of the question.

For instance I remember doing a homework on the statistical correlation between CO2 human emissions and ocean'ss warming with data from late XIX century up to 2010. The result (coinciding with the results obtained by other students) was there is a correlation greater than 99.9% (I don't remember the exact value, sorry) between the two subjects on under study. Now there's difficult to think of a common cause for oceans being warmed and humans emitting CO2 (other than aliens influencing humans while simultaneously warming the oceans, or the such), so the conclusion was human CO2 emissions are causing ocean warming at some extend.

The exception was a student concluding ocean warming was causing humans to increase the use of air conditioning worldwide, therefore increasing the comsuption of energy, therefore increasing CO2 emissions.

We all had a blast, but the teacher took it seriously and showed how that increase in energy comsumption doesn't match the increase on CO2 emission for a huge amount.

My point is, I agree with you sometimes statistical correlation is enough for a causal explanation.

Generally the relationship between statistical correlation and causation are defined by the significance level (alpha) or the probability of rejecting the null hypothesis when it is actually true. In my field we usually assign this alpha = .05, or a 5% chance of rejecting the null. The underlying assumption here is 5% represents a fairly low probability. Following this you calculate the P value, or the probability of getting a a more extreme mean value from testing assuming the null hypothesis is true. When the P value is less than alpha we declare there's a significant difference between the population mean and the testing mean, or that the result of the test is significant. True, use of the term "significant" doesn't indicate 100% causality but it does indicate I'm willing to accept there's a serious relationship between the dependent and independent variables in my test.

The vast majority of smoking/cancer studies I've seen show significance for the hypothesis "smoking causes lung cancer." Note, these studies don't evaluate the statement "only smoking causes lung cancer" or that "all lung cancer is caused by smoking". In general though, like the box choice problem, it raises the question for the individual - is smoking worth it?

posted 3 days ago by **mel_mann**

I totally get the concept of significant difference ... but there is no such thing as 100% confidence. This is the difference between the causal and probabilistic thinker (or, 2-boxers and 1-boxers)

Now, it makes a lot of sense to live your life based treating things that have an alpha of 0.05 as true. It is a fantastic simplifying assumption and you won't go far wrong... however life is complicated, and we may have to make decisions, in real time, about hundreds or thousands or even millions of things that may have a correlations alpha of 0.05. Let's say there are 100. That means you are likely wrong about the causal chain of 5 of those things. And since all these correlations interact, it isn't even that simple. Some errors could be correlated... like polling results by state before an election... which is why the Donald won even though many predicted a 99% chance of Hilary winning... they didn't account for a systematic error among polling between the states, except for the stats guru Nate Silver (of 538) who gave her a 2/3 chance of winning, stating that all it would take was a 2 point systematic polling error, which had been seen historically and turned out to be the case is 2016.

So, how do we decide of the 100 things that we are 95% confident are true or not? We can always get more data, but that takes time. Some people guess... but I prefer to find explanations for why some things would be true and some might not be. This is how science works. The best explanations that are backed by the lowest alpha correlations are the ones we call theories. And when we have a good theory, like smoking leads to cancer, we then use the stronger language that smoking causes cancer, despite the fact that we continue to use the term theory, not truth. But we can never be sure of causation, just like we can never be sure we have found the truth.

You can also think about the beta risk, right? In the case of smoking, sure, there is a small chance the scientists are wrong, but what, exactly, are you missing out on? An expensive habit that, unless you hang out with other smokers, people find disgusting? That goes into your decision making as well. It goes the other way as well. Sometimes we can be really sure of something, but the cost of being wrong is so high and the benefit of being correct so low that we should not act as if it were true. Climate change deniers should be thinking this way, but that would require them to be rational. Ah well.

2-boxers claim to have a perfect understanding of how causation works. I'm claiming it is always rational, no matter what you are talking about, to assume that you could be wrong, including intuitively obvious things like the future can't effect the past and I have free will. And that leads to 1-boxing, and the reason that 1-boxers will always be richer than 2-boxers... which is fine if that if being rich is your goal. However, maybe your goal is to be really sure of yourself.... which is something evolution programmed into us along with a sense of free will.

posted 2 days ago by **ericcoker1974**

We can have p-values so low that they are almost zero. But even with near 100% confidence, we have to ask: is causation really an empirical question or is it also a way of looking at the world? I would say it is the latter. We can all observe that if you behead a chicken, the chicken will die. But does this mean that your action causes the death of the chicken? Ghazali would disagree and assign causal agency only to God. https://www.bu.edu/wcp/Papers/Medi/MediAdam.htm

posted a day ago by **Kallikles**

I'm just reading Rebecca Goldstein's Betraying Spinoza and am getting a flavor for his philosophy. So if Ghazali means Spinoza's God, which he defines as the cause of everything but is not an anthropomorphized, scary, insecure, lonely man in the sky but instead all possible logical entailment, then sure, God is the cause of all things.
So when I kill a chicken, what caused me to do so? Free will? An evolutionary drive we call hunger? There are always causes for causes ... like a reverse omega sequence.
Granted... that is a different argument than the one I was making before. You are correct in pointing out that there is some special sort of knowledge you get about causation when you witness an event. But if there are 2 people in a room and a dead chicken and a 3rd person walks into a room, it is impossible for that 3rd person to know for certain who killed it, but if one person is covered in blood and holding a knife, you can be highly confident. And scientists are that 3rd person... they aren't intersted so much in things they cause themselves.

I'm not arguing against treating correlation as cause, I'm just trying to point out that there is a danger in doing so when we do it a very large number of times, as we do during the course of our life, not just as individuals, but as a society. For me, the Newcomb problem points to the conclusion that cause is a more slippery concept than we normally think. If it wasn't, then we would all be 1-boxers or 2-boxers, and this would be a very boring lecture.

posted a day ago by **ericcoker1974**

Add a comment

Showing all responses

Add a response:

Preview

Submit