Causal Analysis in Theory and Practice

On Mediation, counterfactuals and manipulations

Filed under: Discussion, Opinion — moderator @ 9:00 pm , 05/03/2010

Judea Pearl's reply:

1.
As to the which counterfactuals are "well defined", my position is that counterfactuals attain their "definition" from the laws of physics and, therefore, they are "well defined" before one even contemplates any specific intervention. Newton concluded that tides are DUE to lunar attraction without thinking about manipulating the moon's position; he merely envisioned how water would react to gravitaional force in general.

In fact, counterfactuals (e.g., f=ma) earn their usefulness precisely because they are not tide to specific manipulation, but can serve a vast variety of future inteventions, whose details we do not know in advance; it is the duty of the intervenor to make precise how each anticipated manipulation fits into our store of counterfactual knowledge, also known as "scientific theories".

2.
Regarding identifiability of mediation, I have two comments to make; ' one related to your Minimal Causal Models (MCM) and one related to the role of structural equations models (SEM) as the logical basis of counterfactual analysis.

In my previous posting in this discussion I have falsely assumed that MCM is identical to the one I called "Causal Bayesian Network" (CBM) in Causality Chapter 1, which is the midlevel in the three-tier hierarchy: probability-intervention-counterfactuals?

A Causal Bayesian Network is defined by three conditions (page 24), each invoking probabilities under interventions, and none relies on counterfactuals. Given that the three conditions imply the truncated product decomposition (G-formula), I have concluded that it is the same mathematical object as MCM, and I was glad to see a counterfactual characterization of this object,

It turns out, however, (and this was brought to my attention by Thomas, yesterday) that there is a 4th layer in the hierarchy, containing those couterfactuals that can be tested empirically (i.e., reducible to do(x) expressions) yet not preresented directly in CBN. A well known example is the binary ETT (Effect of Treatment' on the Treated = P(Y(0)|X=1), see Causality p. 396-7) The idea here is that assumptions from the 1st layer of the hierarchy (i.e., binary variables is a probabilistic restriction) can combine with assumptions from the 2nd lever (i.e., interventional) to lift a third-level quantity like ETT to the second level. Such "liftable" counterfactuals may deserve to be recognized as a 4th, intermediate level, between CBN and SEM.

Note-1, Shpitzer and Pearl (2007) (in "what counterfactuals can be tested" also characterized a set of "liftable" counterfactuals, but they allowed assumptions from the 3rd level, and prohibited distribution-specific assumptions (e.g., binary, normal))

Note-2, Another "liftable counterfactual" is the Probability of Necessity (PN= P(Y(0)=0|X=1, Y=1)) which can be identified given a specific combination of experimental and observational data, see Causality p. 302-3.

3.
Regarding the role of SEM in causal modeling, I do not agree with your critique; let me explain why SEM or NPSEM remains the logical basis for counterfactual analysis. SEM represents the scientific view of nature: a collection of invariant functions (or laws) that connect variables together. To deny SEM is to deny physics and physical thinking, according to which there ARE invariant laws in nature, regardless of whether or not we can test them. In many cases the assumption that such laws exist is all that is needed, not their details.

Indeed, the laws of physics, because they are counterfactual, cannot be given empirical tests. For example, we cannot test the statement: "If the weight on this spring was different, its length would be different as well." We can never be sure that the spring is the same spring, or that the time of day did not change its elasticity. Yet, thus far, this untested Laplacian assumption of invariance has not caused us a major disappointment. Only subatomic processes, governed by quantum mechanics managed to defy the counterfactual interpretation of physical laws (Causality, page 26).

The real question is whether reliance on SEM is more dangerous in epidemiology than our reliance on the untestable invariance of physical laws in every day life, a reliance that gave rise to the scientific revolution and, in my opinion, is not more dangerous than saying that a patient will remain the same patient even if we close our eyes for a split second.

Consider two variables, X and Y. If we assume that the laws of nature are deterministic and invariant we can write:

Y=f(x,u)

where u stands for all factors not affected by X, and then we can write:

Y(0) = f(0,u)
Y(1) = f(1,u).

Assuming that U is independent of X gives X _||_ {Y(0),Y(1)}, which asserts independence accross worlds. We see that, the untestable assumption of independence across worlds carries the same information as the standard unconfoundedness assumption that X is independent of the factor U that nature consults before determining the value of Y.

Judgmentally, this sort of assumption is routinely made whenever we deal with observational studies. True, to merely predict causal effects we can get by with weaker assumptions, expressible in do(x) notation, and verifiable (in principle) from manipulative experiments. But, realistically, human store scientific experience in the form of invariant laws of nature, not in the form of experimental probabilities, and it is this store of knowledge that we tap when we judge the plausibility of causal assumptions.

This sort of assumption is the basis for postulating unit level counterfactuals, sometimes written Y_x(u) . Indeed, what gives us the audacity to assume that Y(0) has a unique value for every unit regardless of the treatment actually received by that unit? I contend that he who resists writing the SEM equation Y = f(x,u) should also refrain from using unit-based counterfactuals as a basis for analysis. Purging NPSEM from epidemiology is tantamount to purging most counterfactuals from causal analysis, with the exception of the thin layer of "liftable" counterfactuals.

I take it, therefore, that you do not object to SEM as an analytical and conceptual tool, only to assumptions of form X_||_U, which cannot be tested by RCT.

Next let us attend to practical issues. You say "if someone posit the DAG X--->Z--->Y when in fact there is confounding between Z and Y, then...the assumption that the DAG is an NPSEM will have led us to an inconsistent estimate of a counterfactual contrast." This is a contradictory premise. "if someone posit the DAG X--->Z--->Y when in fact there is confounding between Z and Y," then this someone cannot simultaneously assume that that DAG is an NPSEM, because if it were, the confounding between Z and Y would be representable as a latent variables affecting both. This is what NPSEM is entrusted to represent: functional, not probabilistic dependencies.

Perhaps you are concerned with the possibility that someone will posit the DAG X--->Z--->Y being unaware of of confounding that exists between Z and Y. This is a mispecification omission, not a deficiency of NPSEM as a representation scheme. It is a failure of the modeler to take advantage of the powerful machinery of NPSEM and express the existence of confounding between Z and Y.
.
In summary, I submit that SEM retains its status as the logical basis for counterfactuals. I would like, however, to go further and argue that even untestable assumptions such as X_||_U are useful. They are useful, because scientists are useful. In other words, whenever our analysis depends on scientific knowledge obtained from human beings, beware, that knowledge is already molded, filtered, contaminated, perhaps distorted by assumptions of the type X_||_U, exactly the type that you warn us to avoid.

I believe the ability to defend those assumptions on the basis of scientific theories more than makes up for our inability to verify them experimentally in each specific case. I therefore] stand behind the Mediation Formula:

Indirect Effect = sum_z E(Y(x=0,Z=z)[P(z|X=1)-P(z|x=0)] (see: http://ftp.cs.ucla.edu/pub/stat_ser/r363.pdf), as long as I am willing to defend its underlying assumptions, here cast in a solid and transparent scientific language (i.e., NPSEM):

$X=h(U_x)$
$Z=f(X, U_z)$
$Y=g(X, Z, U_y)$
and $\{U_x, U_z, U_y\}$ mutually independent.

Best regards,
Judea Pearl