Intervening is Imaging is Conditioning

Franz Huber, University of Toronto

Abstract

I show that, in acyclic causal models, post-intervention probabilities are imaging probabilities and both are conditional probabilities.

1 Acyclic causal models

Let us consider an acyclic causal model $\mathcal{M}$ of the sort that is central to causal modeling (Spirtes et al. 1993/2000, Pearl 2000/2009, Halpern 2016, Hitchcock 2018). Readers familiar with them can skip this section.

$\mathcal{M} = \langle \mathcal{S}, \mathcal{F} \rangle$ is a *causal model* if, and only if, $\mathcal{S}$ is a signature and $\mathcal{F} = \{F_1, \ldots, F_n\}$ represents a set of n structural equations, for a finite natural number n. $\mathcal{S} = \langle \mathcal{U}, \mathcal{V}, R \rangle$ is a *signature* if, and only if, $\mathcal{U}$ is a finite set of exogenous variables, $\mathcal{V} = \{V_1, \ldots, V_n\}$ is a set of n endogenous variables that is disjoint from $\mathcal{U}$, and $R : \mathcal{U} \cup \mathcal{V} \to \mathcal{R}$ assigns to each exogenous or endogenous variable X in $\mathcal{U} \cup \mathcal{V}$ its *range* (not co-domain) $R(X) \subseteq \mathcal{R}$. $\mathcal{F} = \{F_1, \ldots, F_n\}$ *represents a set of n structural equations* if, and only if, for each natural number i, $1 \leq i \leq n$: F_i is a function from the Cartesian product $\mathcal{W}_i = \times_{X \in \mathcal{U} \cup \mathcal{V} \setminus \{V_i\}} R(X)$ of the ranges of all exogenous and endogenous variables other than V_i into the range $R(V_i)$ of the endogenous variable V_i. The *set of possible worlds* of the causal model $\mathcal{M}$ is defined as the Cartesian product $\mathcal{W} = \times_{X \in \mathcal{U} \cup \mathcal{V}} R(X)$ of the ranges of all exogenous and endogenous variables.

A causal model $\mathcal{M}$ is *acyclic* if, and only if, it is not the case that there are m endogenous variables $V_{i1}, \ldots, V_{im}$ in $\mathcal{V}$, for some natural number m, $2 \leq m \leq n$, such that the value of $F_{i(j+1)}$ depends on $R\left(V_{ij}\right)$ for $j = 1, \ldots, m-1$, and the value of F_{i1} depends on $R(V_{im})$. Importantly, dependence is just ordinary functional dependence: F_i depends on $R\left(V_j\right)$ if, and only if, there are arguments $\vec{w}_i$ and $\vec{w}_i{}'$ in the domain $\mathcal{W}_i = \times_{X \in \mathcal{U} \cup \mathcal{V} \setminus \{V_i\}} R(X)$ of F_i that differ only in the value from $R\left(V_j\right)$ such that their values under F_i differ, $F_i(\vec{w}_i) \neq F_i\left(\vec{w}_i{}'\right)$.

Let $Pa(V_i)$ be the set of variables X in $\mathcal{U} \cup \mathcal{V}$ such that F_i depends on $R(X)$. The elements of $Pa(V_i)$ are the *parents* of the endogenous variable V_i, that is, the set of variables that are *directly causally relevant* to V_i. Let $An(V_i)$ be the ancestral, or transitive closure, of $Pa(V_i)$, which is defined recursively as follows: $Pa(V_i) \subseteq An(V_i)$; if $V \in An(V_i)$, then $Pa(V) \subseteq An(V_i)$; and, nothing else is in $An(V_i)$. The elements of $An(V_i)$ are the *ancestors* of the endogenous variable V_i. A variable Y is a *non-descendant* of a variable X if, and only if, X and Y are different and X is not an ancestor of Y.

A *context* is a specification of the values of all exogenous variables. It can be represented by a vector $\vec{u}$ in the Cartesian product $R(\mathcal{U}) = \times_{U \in \mathcal{U}} R(U)$ of the ranges of all exogenous variables. A basic fact about causal models is that every acyclic causal model has a unique solution $w_{\vec{u}}$ for any context $\vec{u}$. Let $\mathcal{W}_0$ be the set of these "legal" possible worlds (Glymour et al. 2010). An acyclic causal model determines a unique directed acyclic graph whose nodes are the exogenous and endogenous variables in $\mathcal{U} \cup \mathcal{V}$ and whose arrows point into each endogenous variable V_i from all of the latter's parents in $Pa(V_i)$.

Acyclic causal models provide a semantics for some counterfactuals. The language includes atomic sentences of the form $V = v$ which say that endogenous variable V takes on a specific value v from its range $R(V)$, as well as the Boolean combinations that can be formed from these atomic sentences by finitely many applications of negation $\neg$, conjunction $\wedge$, and disjunction $\vee$. The variables must be endogenous. Sentences of the form $V \in S$, for a subset S of $R(V)$ with more (or less) than one element are not allowed. The antecedent of a counterfactual must by a finite conjunction $X_1 = x_1 \wedge \ldots \wedge X_k = x_k$ of one or more atomic sentences with distinct endogenous variables. The consequent must a Boolean combination ϕ of atomic sentences. Among others, this means that we cannot consider counterfactuals with a counterfactual in the antecedent or consequent.

An atomic sentence $V = v$ is true in $\mathcal{M}$ in $\vec{u}$ if, and only if, all solutions to the structural equations represented by $\mathcal{F}$ assign value v to the endogenous variable V if the exogenous variables in $\vec{\mathcal{U}}$ are set to $\vec{u}$. Since we are restricting the discussion to extended acyclic causal models which have a unique solution in any given context, this means that $V = v$ is true in $\mathcal{M}$ in $\vec{u}$ if, and only if, v is the value of V in the unique solution $w_{\vec{u}}$ to all equations in $\mathcal{M}$ in $\vec{u}$. The truth conditions for negations, conjunctions, and disjunctions are given in the usual way. The counterfactual $X_1 = x_1 \wedge \ldots \wedge X_k = x_k \mathrel{\Box\!\!\rightarrow} \phi$, or simply $\vec{X} = \vec{x} \mathrel{\Box\!\!\rightarrow} \phi$, is true in $\mathcal{M} = \langle \mathcal{S}, \mathcal{F} \rangle$ in $\vec{u}$, $\mathcal{M}, \vec{u} \models \vec{X} = \vec{x} \mathrel{\Box\!\!\rightarrow} \phi$ if, and only if, ϕ is true in $\mathcal{M}_{\vec{X}=\vec{x}} = \langle \mathcal{S}_{\vec{X}}, \mathcal{F}^{\vec{X}=\vec{x}} \rangle$ in $\vec{u}$.

The latter causal model results from $\mathcal{M}$ by removing the structural equation for X_i and by freezing the value of X_i at x_i, for each $i = 1, \ldots, k$. Formally, this means that $\mathcal{S}$ is reduced to $\mathcal{S}_{\vec{X}} = \langle \mathcal{U}, \mathcal{V} \setminus \{X_1, \ldots, X_k\}, \mathcal{R} \restriction_{\mathcal{U} \cup \mathcal{V} \setminus \{X_1, \ldots, X_k\}} \rangle$, where $\mathcal{R} \restriction_{\mathcal{U} \cup \mathcal{V} \setminus \{X_1, \ldots, X_k\}}$ is $\mathcal{R}$ with its domain restricted from $\mathcal{U} \cup \mathcal{V}$ to $\mathcal{U} \cup \mathcal{V} \setminus \{X_1, \ldots, X_k\}$; as well as that $\mathcal{F}$ is reduced to $\mathcal{F}^{\vec{X}=\vec{x}}$ which results from $\mathcal{F}$ by deleting, for each $i = 1, \ldots, k$, the function F_{X_i} representing the structural equation for X_i and by changing the remaining functions F_Y in $\mathcal{F} \setminus \{F_{X_1}, \ldots, F_{X_k}\}$ as follows: restrict the domain of each F_Y from $\times_{X \in \mathcal{U} \cup \mathcal{V} \setminus \{Y\}} R(X)$ to $\times_{X \in \mathcal{U} \cup \mathcal{V} \setminus \{Y, X_1, \ldots, X_k\}} R(X)$; and, replace F_Y by $F_Y^{\vec{X}=\vec{x}}$ which results from F_Y by setting $X_1, \ldots, X_k$ to $x_1, \ldots, x_k$, respectively.

2 Probability

Next let us consider a regular probability measure Pr on the power-set of $\mathcal{W}$. This means that every non-empty proposition over $\mathcal{W}$ receives a positive probability, including the singletons containing a possible world which I will identify with each other. The conditional probability $\Pr(\cdot \mid \mathcal{W}_0)$ is the probability measure conditional on the assumption that $\mathcal{M}$ is true and no intervention takes place. Note that

$$\Pr(w_{\vec{u}} \mid \mathcal{W}_0) = \Pr\left([\![\vec{U} = \vec{u}]\!]_{\mathcal{W}}\right),$$

where $[\![\vec{U} = \vec{u}]\!]_{\mathcal{W}}$ is the proposition over $\mathcal{W}$ that is expressed by the sentence $\vec{U} = \vec{u}$. This means that $\Pr(\cdot \mid \mathcal{W}_0)$ allocates the entire probability mass of context $\vec{u}$ onto the single possible world $w_{\vec{u}}$; every other possible world that agrees with $w_{\vec{u}}$ on the values of the exogenous variables $\mathcal{U}$ receives probability zero.

If the set of exogenous variables $\mathcal{U}$ is probabilistically independent in the sense of $\Pr(\cdot \mid \mathcal{W}_0)$, Pearl (2000/2009: 30)'s causal Markov condition theorem applies: $\Pr(\cdot \mid \mathcal{W}_0)$ satisfies the causal Markov condition for the directed acyclic graph determined by $\mathcal{M}$ (each variable in $\mathcal{U} \cup \mathcal{V}$ is probabilistically independent of its non-descendants given its parents). In this case the pair $\langle \mathcal{M}, \Pr(\cdot \mid \mathcal{W}_0) \rangle$ is Markovian; it is semi-Markovian, if the set of exogenous variables $\mathcal{U}$ is not probabilistically independent in the sense of $\Pr(\cdot \mid \mathcal{W}_0)$. The significance of this theorem lies in connecting acyclic causal models to probability.

It is here that I am departing slightly from the approach usually taken. Usually (e.g., Pearl 2000/2009: ch. 3), one starts with a regular probability measure $\Pr_{\mathcal{U}}$ over the power-set of $R(\mathcal{U})$ and then extends $\Pr_{\mathcal{U}}$ to a unique regular probability measure $\Pr_{\mathcal{M}}$ over the power-set of $\mathcal{W}_0$. While

$$\Pr{}_{\mathcal{U}}\left([\![\vec{U}=\vec{u}]\!]_{R(\mathcal{U})}\right)=\Pr{}_{\mathcal{M}}\left([\![\vec{U}=\vec{u}]\!]_{\mathcal{W}_0}\right)=$$
$$=\Pr\left([\![\vec{U}=\vec{u}]\!]_{\mathcal{W}}\mid\mathcal{W}_0\right)=\Pr\left([\![\vec{U}=\vec{u}]\!]_{\mathcal{W}}\right)$$

for every context $\vec{u}$, as well as, for every possible world $w_{\vec{u}}$ that is legal in $\mathcal{M}$,

$$\Pr{}_{\mathcal{M}}(w_{\vec{u}})=\Pr(w_{\vec{u}}\mid\mathcal{W}_0)\,,$$

the sentence $\vec{U}=\vec{u}$ picks out different propositions over $R(\mathcal{U})$, $\mathcal{W}_0$, and $\mathcal{W}$. In addition, the probability measures $\Pr_{\mathcal{U}}$ and $\Pr_{\mathcal{M}}$ do not assign any probability to propositions comprised by possible worlds that are illegal in $\mathcal{M}$, while these propositions receive probability zero from $\Pr(\cdot\mid\mathcal{W}_0)$ and positive probability from Pr. It is this slight departure that enables me to prove my claims.

The post-intervention probability $\Pr(\cdot\mid\mathcal{W}_0)_{do\left(\vec{X}=\vec{x}\right)}$ after intervening on the endogenous variables $\vec{X}$ and setting their values to $\vec{x}$ is usually defined to be the unique regular probability measure $\Pr_{\mathcal{M}_{\vec{X}=\vec{x}}}$ over the power-set of the set of legal possible worlds $\mathcal{W}_0^{\vec{X}=\vec{x}}$ of the acyclic causal model $\mathcal{M}_{\vec{X}=\vec{x}}$ that extends $\Pr_{\mathcal{U}}$ (in the same manner as $\Pr_{\mathcal{M}}$ extends $\Pr_{\mathcal{U}}$ to a unique probability measure over the power-set of $\mathcal{W}_0$). It can also be calculated from the pre-intervention probability $\Pr(\cdot\mid\mathcal{W}_0)$ as follows (see Spirtes et al. 1993/2000: 51's manipulation theorem): for any possible world w in $\mathcal{W}$,

$$\Pr(w\mid\mathcal{W}_0)_{do\left(\vec{X}=\vec{x}\right)}=\Pr{}^*\left([\![\vec{X}=X(\vec{w})]\!]_{\mathcal{W}}\right)\times$$
$$\times\prod_{Y\in\mathcal{U}\cup\mathcal{V}\setminus\{X_1,\ldots,X_k\}}\Pr\left([\![Y=Y(w)]\!]_{\mathcal{W}}\mid[\![Pa\vec{(Y)}=Pa(\vec{Y})(w)]\!]_{\mathcal{W}}\cap\mathcal{W}_0\right),$$

where $\vec{X}(w)$ are the values of the variables $\vec{X}$ in w and the intervention-function $\Pr^*$ takes on value 1 for $\vec{X}(w)=\vec{x}$ and value 0 for $\vec{X}(w)\neq\vec{x}$. The post-intervention probability $\Pr(\cdot\mid\mathcal{W}_0)_{do\left(\vec{X}=\vec{x}\right)}$ satisfies the causal Markov condition for the directed acyclic graph determined by the acyclic causal model $\mathcal{M}_{\vec{X}=\vec{x}}$.

To establish my claims, let me amend a concept of Lewis (1973: 133)': *the interventionist theory of* $\vec{X}=\vec{x}$ *in context* $\vec{u}$, $IT\left(\vec{X}=\vec{x},\vec{u}\right)$, is the set of sentences that would be true in $\mathcal{M}$ in $\vec{u}$ if $\vec{X}=\vec{x}$ were made true in $\mathcal{M}$ in $\vec{u}$ by an intervention that sets the values of $\vec{X}$ to $\vec{x}$,

$$\left\{\phi:\mathcal{M},\vec{u}\models\vec{X}=\vec{x}\mathrel{\Box\!\!\rightarrow}\phi\right\}.$$

$IT\left(\vec{X}=\vec{x},\vec{u}\right)$ is a set of Boolean combinations of atomic sentences. Unlike the counterfactuals in section 1, atomic sentences and Boolean combinations thereof are assigned truth values in $\mathcal{M}$ not just relative to contexts $\vec{u}$, but relative also to possible worlds w. In fact, I presupposed this when considering the propositions expressed by $\vec{U}=\vec{u}$ over various sets of possible worlds (an atomic sentence $V=v$ is true in $\mathcal{M}$ in w if, and only if, v is the value of V in w; the truth conditions for negations, conjunctions, and disjunctions continue are given in the usual way). Therefore, we can say that $IT\left(\vec{X}=\vec{x},\vec{u}\right)$ is true in $\mathcal{M}$ in precisely one possible world, viz. the unique solution $w_{\vec{u}}^{\vec{X}=\vec{x}}$ to all equations in $\mathcal{M}_{\vec{X}=\vec{x}}$ in $\vec{u}$. In the framework of Stalnaker (1968), whose central ingredient is a selection function f, the corresponding set of sentences picks out the unique possible world $f\left(\vec{X}=\vec{x},w_{\vec{u}}\right)$ that is selected by f in the possible world $w_{\vec{u}}$ for the antecedent $\vec{X}=\vec{x}$ as the unique possible world among all possible worlds in which $\vec{X}=\vec{x}$ is true that is closest or most similar to $w_{\vec{u}}$. In the slightly less demanding framework of Lewis (1973) the corresponding set of sentences may pick out the empty set if one does not make Lewis (1973: 19)'s "limit assumption" (Herzberger 1979).

When one brings about $\vec{X}=\vec{x}$ by an intervention that sets the values of $\vec{X}$ to $\vec{x}$ and one assumes that $\mathcal{M}$ is true, the information one receives is the proposition expressed by the disjunction or intersection of all sets $IT\left(\vec{X}=\vec{x},\vec{u}\right)$, for every context $\vec{u}$, i.e.,

$$IT\left(\vec{X}=\vec{x}\right)=\bigcap_{\vec{u}\in R(\mathcal{U})}\left\{\phi:\mathcal{M},\vec{u}\models\vec{X}=\vec{x}\mathrel{\Box\!\!\rightarrow}\phi\right\}.$$

$IT\left(\vec{X}=\vec{x}\right)$ is true in all and only the possible worlds in $\mathcal{W}_0^{\vec{X}=\vec{x}}$, which is the set of legal possible worlds of the acyclic causal model $\mathcal{M}_{\vec{X}=\vec{x}}$. $IT\left(\vec{X}=\vec{x}\right)$ says that, assuming that $\mathcal{M}$ is true, what would be the case if $\vec{X}=\vec{x}$ were made true by an intervention that sets the values of $\vec{X}$ to $\vec{x}$ *is* the case. $IT\left(\vec{X}=\vec{x}\right)$ implies $\vec{X}=\vec{x}$ (but, in general, is not implied by it). This is so also in the frameworks of Stalnaker (1968) and Lewis (1973), as well as any other that validates $\vec{X}=\vec{x}\mathrel{\Box\!\!\rightarrow}\vec{X}=\vec{x}$ (such as Huber 2021's typicality framework).

Note that, for every context $\vec{u}$,

$$\Pr\left(w_{\vec{u}}^{\vec{X}=\vec{x}}\mid\mathcal{W}_0\right)_{do\left(\vec{X}=\vec{x}\right)}=\Pr\left(\llbracket\vec{U}=\vec{u}\rrbracket_{\mathcal{W}}\mid\mathcal{W}_0\right)_{do\left(\vec{X}=\vec{x}\right)},$$

as well as

$$\Pr\left([\![\vec{U}=\vec{u}]\!]_{\mathcal{W}} \mid \mathcal{W}_0\right)_{do\left(\vec{X}=\vec{x}\right)} = \Pr\left([\![\vec{U}=\vec{u}]\!]_{\mathcal{W}} \mid \mathcal{W}_0\right)_{do\left(\vec{Y}=\vec{y}\right)} = \Pr\left([\![\vec{U}=\vec{u}]\!]_{\mathcal{W}}\right)$$

for any two interventions on endogenous variables $\vec{X}$ and $\vec{Y}$. This means that the post-intervention probability $\Pr(\cdot \mid \mathcal{W}_0)_{do\left(\vec{X}=\vec{x}\right)}$ re-allocates the probability mass $\Pr\left([\![\vec{U}=\vec{u}]\!]_{\mathcal{W}}\right)$ away from the possible world $w_{\vec{u}}$ that is legal in $\mathcal{M}$ to the possible world $w_{\vec{u}}^{\vec{X}=\vec{x}}$ that is legal in $\mathcal{M}_{\vec{X}=\vec{x}}$. This in turn means that the post-intervention probability $\Pr(\cdot \mid \mathcal{W}_0)_{do\left(\vec{X}=\vec{x}\right)}$ is what Lewis (1976: 310) calls *the image of* the pre-intervention probability $\Pr(\cdot \mid \mathcal{W}_0)$ (not Pr) *on* $[\![\vec{X}=\vec{x}]\!]_{\mathcal{W}}$ (modulo the fact that Lewis 1976 works with sentences rather than propositions). This imaging probability is the pre-intervention probability of counterfactuals which validate conditional excluded middle (as does the counterfactual one from section 1, as well as Stalnaker 1968's) with antecedent $\vec{X}=\vec{x}$,

$$\Pr\left([\![\vec{X}=\vec{x} \mathrel{\Box\!\!\rightarrow} \cdot]\!]_{\mathcal{W}} \mid \mathcal{W}_0\right).$$ [1]

My claims follow by noting that both of them are identical to the conditional probability $\Pr\left(\cdot \mid [\![IT\left(\vec{X}=\vec{x}\right)]\!]_{\mathcal{W}}\right)$, i.e., $\Pr\left(\cdot \mid \mathcal{W}_0^{\vec{X}=\vec{x}}\right)$.

These results remain true if the intervention on the endogenous variables $\vec{X}$ does not set their values to $\vec{x}$ but imposes a probability distribution on them so that the intervention-function $\Pr^*\left([\![\vec{X}=\vec{X}(w)]\!]_{\mathcal{W}}\right)$ takes on not just the values 1 and 0, but values between 1 and 0 that sum to 1. In this case we are conditioning in a generalized way (Jeffrey 1965/1983: ch. 11):

$$\Pr(\cdot \mid \mathcal{W}_0)_{do\left(\vec{X}=\vec{x}\right)} = \sum_{\vec{x}\in R\left(\vec{X}\right)} \Pr\left(\cdot \mid [\![IT\left(\vec{X}=\vec{x}\right)]\!]_{\mathcal{W}}\right) \times \Pr^*\left([\![\vec{X}=\vec{x}]\!]_{\mathcal{W}}\right).$$

[1] Pearl (2017) also notes the close relationship between intervening and imaging, though arrives at this result in a different way – and without relating either to conditioning. Pearl (2017)'s aim is to enrich the set of sentences for which the *do*-operator is defined. In the present context, this amounts to enriching the antecedents of counterfactuals from simple sentences of the form $X_1 = x_1 \wedge \ldots \wedge X_k = x_k$ to more complex sentences. Specifically, Pearl (2017) wants to allow for interventions on disjunctions (to calculate the expected utilities of disjunctive actions, among other things). This is *exactly* what causality models (Huber ms) allow for, which comprise the structure of acyclic causal models, but go beyond this structure. Pearl (2017)'s assessment that interventions on disjunctions require more structure than is present in acyclic causal models is water on the mills of the proponent of acyclic causality models. I thank Sander Beckers for pointing me to Pearl (2017).

In acyclic causal models, intervening is imaging is conditioning.

It is worth pointing out the role played by the distinction between exogenous and endogenous variables, as well as the causal assumptions contained therein, in arriving at these results. Given an ayclic causal model $\mathcal{M}$, there is no uncertainty about what is the case, as well as what would be the case if one were to intervene on some endogenous variables, once the value of every exogenous variable is fixed. In other words, given $\mathcal{M}$, the exogenous variables are causally sufficient for the endogenous variables. All uncertainty is restricted to the exogenous variables. Hence, the sets of possible worlds that are legal in $\mathcal{M}$ or any of its sub-models $\mathcal{M}_{\vec{X}=\vec{x}}$ are proper subsets of the set of all possible worlds (as long as there is at least one endogenous variable that has more than one value). This is why $\vec{X} = \vec{x}$ is, in general, less informative than $IT\left(\vec{X} = \vec{x}\right)$. However, the difference between observing $\vec{X} = \vec{x}$ and bringing it about that $\vec{X} = \vec{x}$ – or rather: between receiving the information that $\vec{X} = \vec{x}$ is true and that $\vec{X} = \vec{x}$ is made true – vanishes if no causal assumptions are made.

To see this, let $\rightarrow$ be an arbitrary conditional that is defined for at least the same antecedents and consequents as $\Box\!\!\rightarrow$ from section 1, at least as weak as an object-language counterpart of logical implication $\Rightarrow$, and at least as strong as the material conditional $\supset$. This includes $\Box\!\!\rightarrow$ as defined in section 1, but also as defined in the frameworks of Stalnaker (1968) and Lewis (1973) and many others. In addition, assume that $\rightarrow$- conditionals $\vec{X} = \vec{x} \rightarrow \phi$ have truth values relative to all possible worlds w in $\mathcal{W}$, not merely all contexts $\vec{u}$ in $R(\mathcal{U})$.

For any possible world w in $\mathcal{W}$, call

$$T^{\rightarrow}\left(\vec{X} = \vec{x}, w\right) = \left\{\phi : w \models \vec{X} = \vec{x} \rightarrow \phi\right\}$$

the $\rightarrow$-conditional theory of $\vec{X} = \vec{x}$ in w. For fixed $\vec{X} = \vec{x}$ and w, $T^{\rightarrow}\left(\vec{X} = \vec{x}, w\right)$ is the stronger, the weaker the conditional $\rightarrow$.

By assumption, for every antecedent $\vec{X} = \vec{x}$ and possible world w,

$$T^{\Rightarrow}\left(\vec{X} = \vec{x}, w\right) \subseteq T^{\rightarrow}\left(\vec{X} = \vec{x}, w\right) \subseteq T^{\supset}\left(\vec{X} = \vec{x}, w\right).$$

Hence, for every antecedent $\vec{X} = \vec{x}$,

$$T^{\Rightarrow}\left(\vec{X} = \vec{x}\right) \subseteq T^{\rightarrow}\left(\vec{X} = \vec{x}\right) \subseteq T^{\supset}\left(\vec{X} = \vec{x}\right),$$

where we define as follows:

$$T^{\rightarrow}\left(\vec{X} = \vec{x}\right) = \bigcap_{w \in \mathcal{W}} T^{\rightarrow}\left(\vec{X} = \vec{x}, w\right).$$

Note that we are now quantifying over all possible worlds w rather than merely all contexts $\vec{u}$, as we did in the definition of $IT\left(\vec{X}=\vec{x}\right)$. Because of this it follows that

$$T^{\Rightarrow}\left(\vec{X}=\vec{x}\right)=T^{\supset}\left(\vec{X}=\vec{x}\right).$$

Consequently, for every antecedent $\vec{X}=\vec{x}$ and conditional $\rightarrow$,

$$[\![T^{\rightarrow}\left(\vec{X}=\vec{x}\right)]\!]_{\mathcal{W}}=[\![\vec{X}=\vec{x}]\!]_{\mathcal{W}}.$$

Without any causal assumptions, the distinction between being true and making true vanishes.

Recall how we can calculate the post-intervention probability $\Pr\left(\cdot\mid\mathcal{W}_0\right)_{do\left(\vec{X}=\vec{x}\right)}$ from the pre-intervention probability $\Pr\left(\cdot\mid\mathcal{W}_0\right)$, if the set of exogenous variables $\mathcal{U}$ is independent in the sense of $\Pr\left(\cdot\mid\mathcal{W}_0\right)$ (which is the case if, and only if, this is so in the sense of Pr or any of the conditional probabilities $\Pr\left(\cdot\mid\mathcal{W}_0^{\vec{Y}=\vec{y}}\right)$): for any possible world w in $\mathcal{W}$,

$$\Pr\left(w\mid\mathcal{W}_0\right)_{do\left(\vec{X}=\vec{x}\right)}=\Pr{}^{*}\left([\![\vec{X}=X\left(\vec{w}\right)]\!]_{\mathcal{W}}\right)\times$$
$$\times\prod_{Y\in\mathcal{U}\cup\mathcal{V}\setminus\{X_1,\ldots,X_k\}}\Pr\left([\![Y=Y\left(w\right)]\!]_{\mathcal{W}}\mid[\![\vec{Pa}\left(Y\right)=\vec{Pa}\left(Y\right)\left(w\right)]\!]_{\mathcal{W}}\cap\mathcal{W}_0\right).$$

This equation is well-defined if all conditions in the conditional probabilities in the product have positive probability. This is not always the case, as $[\![\vec{Pa}\left(Y\right)=\vec{Pa}\left(Y\right)\left(w\right)]\!]_{\mathcal{W}}\cap\mathcal{W}_0$ may be empty – say, when we intervene on $\vec{Pa}\left(Y\right)$ and set them to values that they do not take on in any legal possible world. I assume that whichever precautions are taken to side-step this issue also apply to the following calculations. (In the present context, one can always consult the acyclic causal model, but the issue is more pressing when one considers pairs of directed acyclic graphs D and probability measures Pr such that Pr satisfies the causal Markov condition for D.)

Note that the conditional probabilities in the product take on only the extreme values 1 and 0 for endogenous variables Y; non-extreme conditional probabilities strictly between 0 and 1 are reserved for exogenous variables Y. Note also that we can rewrite this equation in the following way that I have not seen elsewhere (perhaps because it holds for acyclic causal models, but, unlike the manipulation theorem, not also for pairs of directed acyclic graphs D and probability measures Pr such that Pr satisfies the causal Markov condition for D). For any possible world w in $\mathcal{W}$,

$$\Pr(w \mid \mathcal{W}_0)_{do(\vec{X}=\vec{x})} = \Pr{}^*\left([\![\vec{X} = X(\vec{w})]\!]_{\mathcal{W}}\right) \times \Pr\left([\![\vec{\mathcal{U}} = \mathcal{U}(\vec{w})]\!]_{\mathcal{W}}\right) \times$$
$$\times \prod_{Y \in \mathcal{V} \setminus \{X_1,\ldots,X_k\}} \Pr\left([\![Y = Y(w)]\!]_{\mathcal{W}} \mid [\![\vec{\mathcal{U}}^+ = \mathcal{U}^+(\vec{w})]\!]_{\mathcal{W}} \cap \mathcal{W}_0\right),$$

where $\mathcal{U}^+ = \mathcal{U} \cup \{X_1, \ldots, X_k\}$. This holds even if the set of exogenous variables fails to be independent in the sense of $\Pr(\cdot \mid \mathcal{W}_0)$ (and Pr and all $\Pr\left(\cdot \mid \mathcal{W}_0^{\vec{Y}=\vec{y}}\right)$) and can be simplified as follows: for any possible world w in $\mathcal{W}$,

$$\Pr(w \mid \mathcal{W}_0)_{do(\vec{X}=\vec{x})} = \Pr\left([\![\vec{\mathcal{U}} = \mathcal{U}(\vec{w})]\!]_{\mathcal{W}}\right) \times$$
$$\times \prod_{Y \in \mathcal{V}} \Pr\left([\![Y = Y(w)]\!]_{\mathcal{W}} \mid [\![\vec{\mathcal{U}} = \mathcal{U}(\vec{w})]\!]_{\mathcal{W}} \cap [\![\vec{X} = \vec{x}]\!]_{\mathcal{W}} \cap \mathcal{W}_0\right).$$

The conditional probabilities in the product still take on only the extreme values 1 and 0 for endogenous variables Y, including $X_1, \ldots, X_k$; non-extreme conditional probabilities strictly between 0 and 1 are still reserved for exogenous variables Y. This brings to the fore that, in acyclic causal models, the exogenous variables are causally sufficient for the endogenous variables in the sense that a specification of the former – plus the endogenous variables intervened upon, if any – determines a specification of the latter.

Among others, this highlights that, in acyclic causal models, any genuinely probabilistic feature of causation among endogenous variables (that is not due to probabilistic features of the intervention) derives from probabilistic features among exogenous variables (see Papineau 2022, ms). It also highlights that, in acyclic causal models, both pre- and post-intervention probabilities satisfy the following *determination condition*, even if the set of exogenous variables is not independent in the sense of any of these probabilities or Pr: each variable Y in $\mathcal{U} \cup \mathcal{V}$ is conditionally independent of its non-descendants given the exogenous variables $\mathcal{U}$ (with or without Y) plus the endogenous variables $\vec{X}$ intervened on, if any.

The determination condition has a consequence for causal inference. Consider exogenous variables $U_1, \ldots, U_m$ and endogenous variables $V_1, \ldots, V_n$ and assume that they are governed by some acylic causal model or other. Now consider what in statistics is called the marginal distribution: $\Pr(U_1, \ldots, U_m, V_1, \ldots, V_n)$. If we "observe" $\vec{X} = \vec{x}$, i.e., if we receive the information that $\vec{X} = \vec{x}$ is true (and no further information), we condition on $\vec{X} = \vec{x}$ to obtain the following new marginal distribution:

$$\Pr(U_1, \ldots, U_m, V_1, \ldots, V_n \mid x_1, \ldots, x_k).$$

By contrast, if we intervene on the endogenous variables $\vec{X}$ and set their values to $\vec{x}$, i.e., if we receive the information that $\vec{X} = \vec{x}$ has been made true (and no further information), we condition on $\vec{X} = \vec{x}$ *and that we are still in the same context, whichever one it is,* to obtain the following new *conditional* distribution:

$$\Pr\left(U_1, \ldots, U_m, V_1, \ldots, V_n \mid x_1, \ldots, x_k, U_1, \ldots, U_m\right).$$

This distribution always has the same conditions $U_1, \ldots, U_m$, no matter which acyclic causal model happens to be true. If we focused on the causal Markov instead of the determination condition, we would obtain the following conditional distribution:

$$\Pr\left(U_1, \ldots, U_m, V_1, \ldots, V_n \mid x_1, \ldots, x_k, Pa\left(X_1, \ldots, X_k\right)\right).$$

The latter conditional distribution has different conditions $Pa\left(X_1, \ldots, X_k\right)$ (even though $X_1, \ldots, X_k$ are fixed), depending on which acyclic causal model happens to be true.

That is, assuming an acylic causal model, as well as a probabilistic setting[2], the information that $\vec{X} = \vec{x}$ is true allows one to estimate marginal distributions, while the information that $\vec{X} = \vec{x}$ has been made true allows one to estimate conditional distributions with specified conditions.

3 Conclusion

The mathematics establishing them is entirely trivial, but that does not mean that my claims are trivial also philosophically. They show that, for an important class of conditionals, probabilities of conditionals are conditional probabilities. They show that, on at least one version of it (Meek & Glymour 1994), causal decision theory is a species of evidential decision theory (Jeffrey 1965/1983) – specifically, one that respects Carnap (1947)'s "principal of total evidence": expected utility is calculated with respect to the probability conditional on not just the evidence that an act is taken, but the decision maker's total evidence. Often, this includes the information that the decision maker herself brings about this act all by herself, i.e., by a hard intervention. And they reinforce a message that is at least implicit in the interventionist approach to causation (Spirtes et al. 1993/2000, Pearl 2000/2009, Woodward 2003): causal inference is statistical inference from correlations not between what is true and what is true, but between what is made true and what is true.[3]

[2] The situation is parallel in a non-probabilistic setting (Huber 2015, ms).

References

[1] Carnap, Rudolf (1947), On the Application of Inductive Logic. *Philosophy and Phenomenological Research* **8**, 133-148.

[2] Glymour, Clark & Danks, David & Glymour, Bruce & Eberhardt, Frederick & Ramsey, Joseph & Scheines, Richard & Spirtes, Peter & Teng, Choh Man & Zhang, Jiji (2010), Actual Causation: A Stone Soup Essay. *Synthese* **175**, 169-192.

[3] Halpern, Joseph Y. (2016), *Actual Causality*. Cambridge, MA: MIT Press.

[4] Herzberger, Hans G. (1979), Counterfactuals and Consistency. *Journal of Philosophy* **76**, 83-88.

---

[3]As a postscriptum, let me briefly address an issue I discuss in detail elsewhere. $IT\left(\vec{X}=\vec{x}\right)$ is defined only relative to an acyclic causal model $\mathcal{M}$. In the context of decision theory (Meek & Glymour 1994, Hitchcock 2016) one may want to allow for uncertainty over which acyclic causal model $\mathcal{M}$ is true. Stern (2017) offers one way of doing so by assigning degrees of certainty to pairs of directed acyclic graphs D (possibly determined by an acyclic causal model $\mathcal{M}$) and probability measures Pr such that Pr satisfies the causal Markov condition for D. Like Savage (1954)'s classical, as well as Lewis (1981)'s and Skyrms (1980, 1982)'s causal (Weirich 2020), the resulting interventionist decision theory fails to be partition-invariant: the recommendations of the theory depend on which set of mutually exclusive possible states of the world the decision maker considers.

An alternative is to generalize acyclic causal models to acyclic models of causality (Huber ms). Unlike in causal models, in acyclic models of causality each possible world has its own "causal laws" (possibly, but not necessarily an acyclic causal model) and directed acyclic graph. Consequently, $IT\left(\vec{X}=\vec{x}\right)$ now says that what would be the case if $\vec{X}=\vec{x}$ were made true by an intervention that sets the values of $\vec{X}$ to $\vec{x}$ is the case – without assuming any particular acyclic causal model or possible world to be true. Since $IT\left(\vec{X}=\vec{x}\right)$ still implies $\vec{X}=\vec{x}$,

$$\Pr\left(\cdot \mid [\![IT\left(\vec{X}=\vec{x}\right)]\!]_{\mathcal{W}}\right)=\Pr\left(\cdot \mid [\![IT\left(\vec{X}=\vec{x}\right)]\!]_{\mathcal{W}}\cap[\![\vec{X}=\vec{x}]\!]_{\mathcal{W}}\right).$$

This has the desirable consequence that, as in Jeffrey (1965/1983)'s evidential decision theory, one can arrive at a formula for calculating expected utility that is partition-invariant (Joyce 1999: sct. 5.5, 2000).

In defining $IT\left(\vec{X}=\vec{x}\right)$, we now quantify over all possible worlds w that are compatible with the agent's causal assumptions when no intervention takes place. This delivers the same result as the definition in section 2 if one assumes that $\mathcal{M}$ is true and no intervention takes place. It still reduces to $\vec{X}=\vec{x}$ if no causal assumptions are made, but this is as it should be.

[5] Hitchcock, Christopher R. (2016), Conditioning, Intervening, and Decision. *Synthese* **193**, 1157-1176.

[6] Hitchcock, Christopher R. (2018), Causal Models. In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.

[7] Huber, Franz (2015), What Should I Believe About What Would Have Been the Case? *Journal of Philosophical Logic* **44**, 81-110.

[8] Huber, Franz (2021), *Belief and Counterfactuals. A Study in Means-End Philosophy*. New York, NY: Oxford University Press.

[9] Huber, Franz (ms), *Causality, Counterfactuals, and Belief. More Means-End Philosophy*. Unpublished manuscript.

[10] Jeffrey, Richard C. (1965/1983), *The Logic of Decision*. 2nd ed. Chicago: University of Chicago Press.

[11] Joyce, James M. (1999), *The Foundations of Causal Decision Theory*. New York: Cambridge University Press.

[12] Joyce, James M. (2000), Why We Still Need the Logic of Decision. *Philosophy of Science* **67**, S1-S13.

[13] Lewis, David K. (1973), *Counterfactuals*. Cambridge, MA: Harvard University Press.

[14] Lewis, David K. (1976), Probabilities of Conditionals and Conditional Probabilities. *The Philosophical Review* **85**, 297-315.

[15] Lewis, David K. (1981), Causal Decision Theory. *Australasian Journal of Philosophy* **59**, 5-30.

[16] Meek, Christopher & Glymour, Clark (1994), Conditioning and Intervening. *British Journal for the Philosophy of Science* **45**, 1001-1021.

[17] Papineau, David (2022), The Statistical Nature of Causation. *The Monist* **105**, 247-275.

[18] Papineau, David (ms), The Causal Structure of Reality. Unpublished manuscript.

[19] Pearl, Judea (2000/2009), *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge: Cambridge University Press.

[20] Pearl, Judea (2017), Physical and Metaphysical Counterfactuals: Evaluating Disjunctive Actions. *Journal of Causal Inference* **5**, 46-54.

[21] Savage, Leonard J. (1954), *The Foundations of Statistics*. New York: Wiley.

[22] Spirtes, Peter & Glymour, Clark & Scheines, Richard (1993/2000), *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press.

[23] Skyrms, Brian (1980), *Causal Necessity. A Pragmatic Investigation of the Necessity of Laws*. New Haven, CT: Yale University Press.

[24] Skyrms, Brian (1982), Causal Decision Theory. *Journal of Philosophy* **79**, 695-711.

[25] Stalnaker, Robert C. (1968), A Theory of Conditionals. In N. Rescher (ed.), *Studies in Logical Theory*. *American Philosophical Quarterly Monograph Series* **4**. Oxford: Blackwell, 98-112.

[26] Stern, Reuben (2017), Interventionist Decision Theory. *Synthese* **194**, 4133-4153.

[27] Weirich, Paul (2020), Causal Decision Theory. In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.

[28] Woodward, James (2003), *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.