Intervening is Conditioning

Franz Huber, University of Toronto

Abstract

The thesis of this short note is that post-intervention probabilities can be considered to be certain 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 $X_1 = x_1 \wedge \ldots \wedge X_k = x_k \boxdot\!\!\rightarrow \phi$, or simply $\vec{X} = \vec{x} \boxdot\!\!\rightarrow \phi$, is true in $\mathcal{M} = \langle \mathcal{S}, \mathcal{F} \rangle$ in $\vec{u}$, $\mathcal{M}, \vec{u} \models \vec{X} = \vec{x} \boxdot\!\!\rightarrow \phi$ 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 probability measure $\Pr_{\mathcal{U}}$ over the power-set of $R(\mathcal{U})$. To avoid technical complications, assume that all variables take on at most finitely many values. We can extend $\Pr_{\mathcal{U}}$ to a unique probability measure $\Pr_{\mathcal{M}}$ over the power-set of $\mathcal{W}$ by allocating the probability of context $\vec{u}$ to the unique possible world $w_{\vec{u}}$ that is legal in $\mathcal{M}$ and assigning probability zero to all possible worlds that are illegal in $\mathcal{M}$.

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

The post-intervention probability $\Pr_{\mathcal{M}_{\vec{X}=\vec{x}}}$ relative to acyclic causal model $\mathcal{M}$ after intervening on the endogenous variables $\vec{X}$ and setting their values to $\vec{x}$ can be defined to be the unique probability measure over the power-set of $\mathcal{W}$ that extends $\Pr_{\mathcal{U}}$ in a manner analogous manner to $\Pr_{\mathcal{M}}$, viz. by allocating the probability of context $\vec{u}$ to the unique possible world $w_{\vec{u}}^{\vec{X}=\vec{x}}$ that is legal in $\mathcal{M}_{\vec{X}=\vec{x}}$ and assigning probability zero to all possible worlds that are illegal in $\mathcal{M}_{\vec{X}=\vec{x}}$. It can also be calculated from the pre-intervention probability $\Pr_{\mathcal{M}}$ as follows (see Spirtes et al. 1993/2000: 51's manipulation theorem): for any possible world w in $\mathcal{W}$,

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

if the conditional probabilities in the latter product are all defined. Here, $\vec{Y}(w)$ are the values of the variables $\vec{Y}$ in w, $[\![Y = Y(w)]\!]$ is the proposition over $\mathcal{W}$ expressed by the sentence $Y = Y(w)$, and the intervention-function $\Pr^*$ takes on value 1 for $\vec{X}(w) = \vec{x}$ and value 0 for $\vec{X}(w) \neq \vec{x}$. (In the interest of readability, I do not distinguish between singletons of contexts or possible worlds and their elements.) The post-intervention probability satisfies the causal Markov condition for the directed acyclic graph determined by the acyclic causal model $\mathcal{M}_{\vec{X}=\vec{x}}$ if the pre-intervention probability satisfies the causal Markov condition for the directed acyclic graph that is determined by $\mathcal{M}$, i.e., if the set of exogenous variables $\mathcal{U}$ is independent in the sense of $\Pr_{\mathcal{M}}$.

Note that, for every context $\vec{u}$, as well as any two interventions on endogenous variables $\vec{X}$ and $\vec{Y}$:

$$\Pr_{\mathcal{M}_{\vec{X}=\vec{x}}}\left(w_{\vec{u}}^{\vec{X}=\vec{x}}\right) = \Pr_{\mathcal{U}}(\vec{u}) = \Pr_{\mathcal{M}_{\vec{Y}=\vec{y}}}\left(w_{\vec{u}}^{\vec{Y}=\vec{y}}\right)$$

That is, the post-intervention probability $\Pr_{\mathcal{M}_{\vec{X}=\vec{x}}}$ re-allocates the probability of context $\vec{u}$ away from the unique possible world $w_{\vec{u}}$ that is legal in $\mathcal{M}$ to the unique possible world $w_{\vec{u}}^{\vec{X}=\vec{x}}$ that is legal in $\mathcal{M}_{\vec{X}=\vec{x}}$. This means that the post-intervention probability $\Pr_{\mathcal{M}_{\vec{X}=\vec{x}}}$ is what Lewis (1976: 310) calls *the image of* the pre-intervention probability $\Pr_{\mathcal{M}}$ *on* $\vec{X} = \vec{x}$ (modulo the fact Lewis 1976 works with sentences rather than propositions). This imaging probability is the pre-intervention probability of counterfactuals $\boxright$ with antecedent $\vec{X} = \vec{x}$ which validate conditional excluded middle, $\Pr_{\mathcal{M}}\left([\![\vec{X} = \vec{x} \boxright \cdot]\!]\right)$.

What I just described is a special case of Pearl (2000/2009: ch. 3)'s *do*-operator, which turns pre-intervention into post-intervention probabilities, except that it is defined also if no acyclic causal model is assumed and one is given merely a directed acyclic graph (possibly with double-arrows) and probability measure satisfying the causal Markov condition for it. Pearl (2017) aims at enriching the set of sentences for which the *do*-operator is defined. (Pearl 2017 also notes the close relationship between intervening and imaging, though arrives at this result in a slightly different way.) In the present context of acyclic causal models, this aim amounts to enriching the set of antecedents for which the interventionist or structural counterfactuals from section 1 are defined.

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.)

Recall how we can calculate the post-intervention probability $\Pr_{\mathcal{M}_{\vec{X}=\vec{x}}}$ from the pre-intervention probability $\Pr_{\mathcal{M}}$, if, for $Y \in \mathcal{U} \cup \mathcal{V} \setminus \{X_1, \ldots, X_k\}$, the conditional probabilities

$$\Pr_{\mathcal{M}}\left([\![Y = Y(w)]\!] \mid [\![Pa(Y) = Pa(Y)(w)]\!]\right)$$

are all defined. The latter need not be the case. $[\![Pa(Y) = Pa(Y)(w)]\!]$ receives probability zero from $\Pr_{\mathcal{M}}$ if we intervene on $Pa(Y)$ and set them to values that they do not take on in any possible world that is legal in $\mathcal{M}$. I assume that whichever precautions are taken to side-step this issue also apply to the following considerations. (In the present context, one can always consult the acyclic causal model, but the issue is more pressing when all one has is a directed acyclic graph and a probability measure that satisfies the causal Markov condition for it.)

These conditional probabilities 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. We can rewrite the relevant 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 and probability measures such that the latter satisfy the causal Markov condition for the former). For any possible world w in $\mathcal{W}$,

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

In fact, this holds even if the set of exogenous variables fails to be independent in the sense of $\Pr_{\mathcal{M}}$. The conditional probabilities in the product still take on only the extreme values 1 and 0 for endogenous variables, including $X_1, \ldots, X_k$; non-extreme conditional probabilities strictly between 0 and 1 are still reserved for exogenous variables. 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 on, 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 highlights also that, in acyclic causal models, both pre- and post-intervention probabilities satisfy the following *causal determination condition*, even if the set of exogenous variables is not independent in the sense of any of these probabilities.

> **Causal Determination Condition** Each exogenous or endogenous variable is conditionally independent of its causal non-descendants or non-effects given all of the exogenous variables, as well as all of the endogenous variables intervened on, if any.

The causal determination condition holds in acyclic causal models for the exact same reason as the causal Markov condition holds in acyclic causal models with independent exogenous variables (Pearl 2000/2009: 30, Pearl & Verma 1994: 792, Steel 2005: 22): in an acyclic causal model, the value of every variable is uniquely determined by a specification of the values of all exogenous variables plus the endogenous variables intervened on, if any.

The causal 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, but it is not specified which one. Now consider what in statistics is called a marginal distribution over these variables:

$$\Pr := \Pr(U_1, \ldots, U_m, V_1, \ldots, V_n)$$

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:

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

This conditional distribution has the same conditions $U_1, \ldots, U_m$, no matter which acyclic causal model is true. We can use it to obtain a new marginal distribution in the following manner, where the sum ranges over all specifications $u_1, \ldots, u_m$ of the values of the exogenous variables $U_1, \ldots, U_m$, respectively:

$$\Pr^{\vec{x}} := \sum_{u_1,\ldots,u_m} \Pr(U_1, \ldots, U_m, V_1, \ldots, V_n \mid x_1, \ldots, x_k, u_1, \ldots, u_m) \Pr(u_1, \ldots, u_m)$$

If we focus on the causal Markov instead of the causal determination condition, we obtain the following conditional distribution:

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

The latter has different conditions $Pa(X_1, \ldots, X_k)$, even though $X_1, \ldots, X_k$ are fixed, depending on which acyclic causal model is true. To determine it, further causal assumptions are needed. If the acyclic causal model $\mathcal{M}$ is given and Pr is the pre-intervention probability $\Pr_{\mathcal{M}}$, we get the post-intervention probability $\Pr_{\mathcal{M}_{\vec{X}=\vec{x}}}$ in this manner, where the sum now ranges over all specifications pa of the values of the direct causes or causal parents $Pa(X_1, \ldots, X_k)$ of the variables $X_1, \ldots, X_k$ intervened on:

$$\Pr_{\mathcal{M}_{\vec{X}=\vec{x}}} = \sum_{pa} \Pr(U_1, \ldots, U_m, V_1, \ldots, V_n \mid x_1, \ldots, x_k, pa) \Pr(pa)$$

Here is why. Rule 2 (action/observation exchange) of Pearl (2000/2009: sct. 3.4)'s *do*-calculus implies that, for any specification uv^- of the values of $UV^- := \{U_1, \ldots, U_m, V_1, \ldots, V_n\} \setminus (\{X_1, \ldots, X_k\} \cup Pa(X_1, \ldots, X_k))$ and any specification pa of the values of $Pa(X_1, \ldots, X_k)$, the conditional post-intervention probability of uv^- given pa equals its conditional pre-intervention probability given $x_1, \ldots, x_k, pa$ (because UV and $X_1, \ldots, X_k$ are d-separated by $PA(X_1, \ldots, X_k)$ after all arrows out of $X_1, \ldots, X_k$ are removed). Furthermore, both conditional probabilities of any specification of values for $\{X_1, \ldots, X_k\} \cup Pa(X_1, \ldots, X_k)$ equal 1 or both equal 0. So, for any specification uv of the values of $UV := \{U_1, \ldots, U_m, V_1, \ldots, V_n\}$ and any specification pa of the values of $Pa(X_1, \ldots, X_k)$, the conditional post-intervention probability of uv given pa equals its conditional pre-intervention probability given $x_1, \ldots, x_k, pa$. Finally, Rule 3 (insertion/deletion of actions) of Pearl (2000/2009: sct. 3.4)'s *do*-calculus implies that, for any specification pa of the values of $Pa(X_1, \ldots, X_k)$, the post-intervention probability of pa equals its pre-intervention probability (because $Pa(X_1, \ldots, X_k)$ and $X_1, \ldots, X_k$ are d-separated by the empty set after all arrows into $X_1, \ldots, X_k$ are removed). The claim then follows from the law of total probability.

3 Consequences

The upshot of this is three-fold. First, intervening is a form of conditioning, viz. one that respects Carnap (1947)'s "principle of total evidence" and obtains a new marginal distribution via obtaining a new conditional distribution whose condition specifies "the total evidence" received, viz. not merely that $\vec{X} = \vec{x}$ is true but the stronger claim that $\vec{X} = \vec{x}$ (is true and) has been made true. This is so even if all we have is a directed acyclic graph and probability measure satisfying the causal Markov condition for it, but no acyclic causal model. In other words, post-intervention probabilities can be considered to be certain conditional probabilities.

Second, in the context of acyclic causal (i.e., structural) models, intervening is a form of conditioning that obtains a new marginal distribution via obtaining a new conditional distribution whose conditions are all of the exogenous variables. Apart from the classification of the variables into exogenous and endogenous, no further causal assumptions are needed. Specifically, no acyclic causal model or directed acyclic graph needs to be specified. In other words, causal inference is possible without assuming a causal model.

Third, on at least one version of it, causal decision theory (Meek & Glymour 1994, Hitchcock 2016) is rendered a species of evidential decision theory (Jeffrey 1965/1983) that respects Carnap (1947)'s "principal of total evidence": expected utility is calculated with respect to the probability conditional on not merely 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.

In addition, a partition-invariant formulation of causal decision theory now is possible. This point requires a bit of background. In decision theory 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, 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. Not so with $\Pr^{\vec{x}}$. As $\Pr^{\vec{x}}(\cdot) = \Pr^{\vec{x}}(\cdot \mid x_1, \ldots, x_k)$, Joyce (1999: sct. 5.5)'s considerations apply and one can arrive at a formula for calculating expected utility that is partition-invariant, as in Jeffrey (1965/2000)'s evidential decision theory (Joyce 2000).

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

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

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

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

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

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

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

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

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

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

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

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

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

[17] Pearl, Judea & Verma, Thomas S. (1994), A Theory of Inferred Causation. In D. Prawitz & B. Skyrms & D. Westerståhl (eds.), *Proceedings of the Ninth International Congress of Logic, Methodology, and Philosophy of Science.* Amsterdam: Elsevier, 789-811.

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

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

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

[21] Steel, Daniel (2005), Indeterminism and the Causal Markov Condition. *British Journal for the Philosophy of Science* **56**, 3-26.

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