DE GRUYTER Epidemiol. Methods 2014; 3(1): 113–120

Julieta Molina, Lucio Pantazis and Mariela Sued*

Some Considerations on the Back Door Theorem and Conditional Randomization

Abstract: In this work, we propose a different “surgical modified model” for the construction of counterfactual variables under non-parametric structural equation models. This approach allows the simultaneous representation of counterfactual responses and observed treatment assignment, at least when the intervention is done in one node. Using the new proposal, the d-separation criterion is used to verify conditions related with ignorability or conditional ignorability, and a new proof of the back door theorem is provided under this framework

Keywords: causal inference, graphical methods, back door theorem

DOI 10.1515/em-2013-0018

1 Introduction

The main objective of this short report is to discuss the relationship between the back door theorem (Pearl 2000, 79) and the conditional randomization (or exchangeability) assumption. We will relate these two concepts through the d-separation rule, constructing both counterfactual variables and observed treatments in the same graph.

During the last years, several authors have been concerned with this problem. The twin directed acyclic graphs (DAGs), presented by Balke and Pearl (1994), allow simultaneous construction of observed and counterfactual variables.

Recently, Richardson and Robins (2013) presented a graphical theory based on single world intervention graphs (SWIGs), unifying causal directed graphs and potential outcomes. The present study, less ambitious, can be considered as a complementary work focusing in particular on the back door theorem, one of the most popular criteria to identify the distribution of counterfactual variables. We can present our results in a simple way, accessible to those who may be nonexperts in the mathematical technicalities, but still familiar with the field and with non-parametric structural equation models (NPSEM) (Pearl 2000, ch. 7). In this setting, Pearl has proposed a modified NPSEM where potential outcomes are defined by replacing the equations related to the treatment nodes by the constants corresponding to the desired intervention. Therefore, the observed treatment assignment and counterfactual variables do not occur together, neither in the model for the observed data nor in the modified model.

We propose here a new modified NPSEM model containing both treatments and counterfactual variables. Unlike the case of the twin graph, variables in this new model factorize according to the back door theorem graph, namely the DAG, where arrows emerging from nodes associated with the intervention are removed. We use this fact to prove that, for univariate treatment, the graphical assumptions of Pearl’s back door theorem (Pearl 2000, 79) imply conditional exchangeability. In this way, we establish a new proof of Pearl’s back door theorem and thus of identifiability of the mean of the counterfactual variables, a proof which can be understood by both proponents of the counterfactual and graphical approaches to causality.

It should be said that even if Richardson and Robins’ graphs differ from the graphs in this paper, our perspectives are rather similar. One difference is that, although they assume an underlying NPSEM, they do

***Corresponding author: Mariela Sued,** Universidad de Buenos Aires, Intendente Güiraldes 2160, Ciudad Universitaria, Buenos Aires C1428EGA, Argentina, E-mail: msued@dm.uba.ar
Julieta Molina: E-mail: julietechka@gmail.com, **Lucio Pantazis:** E-mail: luciopantazis@gmail.com, Universidad de Buenos Aires, Intendente Güiraldes 2160, Ciudad Universitaria, Buenos Aires C1428EGA, Argentina

not assume independence of the disturbances (as we do). Instead, they prefer to work with the Finest Fully Randomized Causally Interpretable Structured Tree Graphs model (FFRCISTG), introduced in Robins (1986), but in the sense of Definition 2 (p. 23) presented in Richardson and Robins (2013).

In this work, we start assuming NPSEM with independent errors (NPSEM-IE). Although NPSEM-IE are less general than FFRCISTG models, we decided to present our proposal under this setting, considering that these models are commonly used. However, in Section 4 we show that all the results presented in this work remain valid if FFRCISTG models are assumed instead.

This work is organized as follows. In Section 2 we present a simple example with a three-node DAG, explaining the main idea to construct jointly the observed treatment assignment and counterfactual variables. We check in this example that the assumptions of the back door theorem imply conditional randomization. In Section 3 we generalize these results, first for the case of an intervention on one node, and then for many nodes. In Section 4 we discuss our results in the framework of FFRCISTG models.

To conclude this introduction, we would like to establish a subtle difference frequently omitted. Given a DAG G, we use $\mathbf{V} = \{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n\}$ to denote the nodes of a graph, whereas random variables associated with a given node $\mathbf{V}_i$ are denoted by V_i or by some perturbation of V_i, like $V_{i,t}$ or V_i^t, as it will be explained later on.

2 Toy example – main idea

In the potential outcome framework (Rubin 1974), the identifiability of the average treatment effect is guaranteed under the assumption of conditional randomization (or ignorability). It states that there exists a vector L of observed variables such that Y_a and A are independent given L, for $a = t, c$, where A is a binary treatment variable taking values in $\{t, c\}$ while Y_a denotes the potential outcomes under treatment level a. More precisely, Y_a is the outcome variable that would have been observed in a hypothetical world in which all individuals received treatment level a. Denoting by Y the observed outcome and assuming that it satisfies $Y = Y_t I_{A=t} + Y_c I_{A=c}$, we get that the average treatment effect ($\mathrm{ATE} = E[Y_t] - E[Y_c]$) is identified by the distribution of observed data (L, A, Y) by the formula $\mathrm{ATE} = E[E[Y|A = t, L]] - E[E[Y|A = c, L]]$.

The d-separation criterion (Pearl 2000, 18) is a graphical tool designed to check independence and conditional independence between coordinates (or sub-vectors) of a random vector whose distribution satisfies the Markov factorization with respect to a given DAG. Then, one is tempted to use such a tool to decide whether conditional ignorability can be assumed for the problem under consideration, studying the DAG associated with it.

For those who are familiar with DAGs, the back door theorem is a famous result used to identify the distribution of the counterfactual variables, and its assumptions give rise to the same formula presented under conditional exchangeability for identifying the average treatment effect. So, we asked ourselves whether the graphical conditions required by the back door theorem allow to prove conditional exchangeability using the d-separation criterion. To answer this question, we need to construct both counterfactual and treatment in the same DAG, in particular in the DAG involved in the back door theorem. To do so, we propose a simple modification to the approach presented by Pearl to define counterfactual variables. In the coming example, we outline the basic idea of our construction, which is generalized in the following section.

Assume that the causal diagram associated with the problem of interest is given by the DAG G (Figure 1).

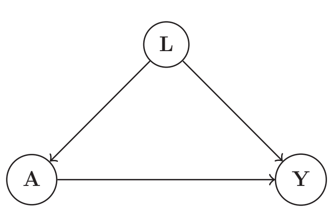

Figure 1 The original DAG G

In terms of NPSEM (Pearl 2000, ch. 7 or 2009) this means that there exists a set of functions $F = \{f_{\mathbf{L}}, f_{\mathbf{A}}, f_{\mathbf{Y}}\}$ and jointly independent disturbances $U = \{U_{\mathbf{L}}, U_{\mathbf{A}}, U_{\mathbf{Y}}\}$, which give rise to factual variables according to the following recursive system:

$$L = f_{\mathbf{L}}(U_{\mathbf{L}})\ , \quad A = f_{\mathbf{A}}(L, U_{\mathbf{A}})\ , \quad Y = f_{\mathbf{Y}}(L, A, U_{\mathbf{Y}})\ . \qquad [1]$$

We use $M = (F, U)$ to denote the model which defines the factual variables. To emulate the intervention $do(a)$, Pearl (2000, ch. 7 or 2009) considers a model M_a where the function $f_{\mathbf{A}}$ is replaced by the constant a, while the disturbances remain unchanged: $M_a = (F_a, U)$, with $F_a = \{f_{a,\mathbf{L}}, f_{a,\mathbf{A}}, f_{a,\mathbf{Y}}\}$, where

$$f_{a,\mathbf{L}} = f_{\mathbf{L}}\ , \quad f_{a,\mathbf{A}} = a\ , \quad f_{a,\mathbf{Y}} = f_{\mathbf{Y}}\ . \qquad [2]$$

The variables obtained iterating the functions in model M_a using the same vector of disturbances $U = \{U_{\mathbf{L}}, U_{\mathbf{A}}, U_{\mathbf{Y}}\}$ are denoted with the subindex a: L_a, A_a and Y_a. In this way, the counterfactual response of interest at level a is given by Y_a.

Our proposal to represent counterfactual variables consists in the use of a new system of functions, in which the value a is inserted in lieu of the variable corresponding to the node **A**, every time this one is required by the recursion. To do so we change the functions related to each node having **A** as parent. In the present example, $M^a = (F^a, U)$, with $F^a = \{f^a_{\mathbf{L}}, f^a_{\mathbf{A}}, f^a_{\mathbf{Y}}\}$, where

$$f^a_{\mathbf{L}} = f_{\mathbf{L}} \quad f^a_{\mathbf{A}} = f_{\mathbf{A}}\ , \quad f^a_{\mathbf{Y}}(\ell, u) = f_{\mathbf{Y}}(\ell, a, u)\ . \qquad [3]$$

Note that this new set of functions is compatible with the DAG $G_{\underline{\mathbf{A}}}$, where arrows emerging from **A** are removed (Figure 2).

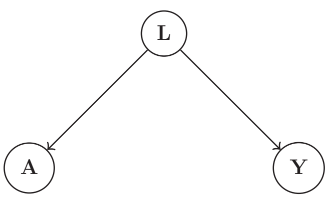

Figure 2 $G_{\underline{\mathbf{A}}}$, constructed removing in G arrows emerging from **A**

The variables constructed iterating the functions in F^a and using the same vector of disturbances $\{U_{\mathbf{L}}, U_{\mathbf{A}}, U_{\mathbf{Y}}\}$ are denoted by the supraindex a: L^a, A^a and Y^a. Then, we get that the distribution of (L^a, A^a, Y^a) is compatible with $G_{\underline{\mathbf{A}}}$.

The following Lemma and Corollary summarize the main results of this section.

Lemma 1 *1.* $L_a = L^a = L$, $A^a = A$ *and* $Y_a = Y^a$.
2. **A** *and* **Y** *are d-separated by* **L** *in* $G_{\underline{\mathbf{A}}}$ *and so, since the distribution of* (L^a, A^a, Y^a) *is compatible with* $G_{\underline{\mathbf{A}}}$*, we get that* A^a *is independent of* Y^a *given* L^a*.*

Corollary 2 *For the causal DAG given in Figure 1, we get that A is independent of* Y_a *given L. Thus, conditional randomization holds.*

3 Intervention with constant regimes

3.1 Interventions on a single node

Consider a causal DAG G with nodes $\mathbf{V}_1, \ldots, \mathbf{V}_n$, labeled in a compatible way with G. Recall that in the graph terminology, we say that $\mathbf{V}_i$ is a parent of $\mathbf{V}_j$ if an arrow points from $\mathbf{V}_i$ to $\mathbf{V}_j$. We use $PA_G(\mathbf{V}_j)$ to denote the set of parents of $\mathbf{V}_j$ in G. If $\mathbf{V}_i$ has a directed path to $\mathbf{V}_k$ we say that $\mathbf{V}_i$ is an ancestor of $\mathbf{V}_k$, and use $An_G(\mathbf{V}_k)$ to denote the set of ancestors of $\mathbf{V}_k$ in G.

Consider a collection of independent random variables $U = \{U_1, \ldots, U_n\}$. Let $\mathcal{V}_i$ denote the common support of any random variable associated with the node $\mathbf{V}_i$ and let $\mathcal{U}_i$ denote the support of U_i. A set of functions $F = \{f_i : i \geq 1\}$ is said to be compatible with G, if for each $i = 1, \ldots n$, we get that

$$f_i : \prod_{\mathbf{V}_j \in PA_G(\mathbf{V}_i)} \mathcal{V}_j \times \mathcal{U}_i \longrightarrow \mathcal{V}_i \,. \qquad [4]$$

Given a set $F = \{f_i : i \geq 1\}$ of compatible functions with G, and independent $U = \{U_1, \ldots, U_n\}$, factual variables are defined by the recurrence

$$V_i = f_i(PA_i, U_i),$$

where PA_i are the random variables (already defined by the recurrence) associated with the nodes in $PA_G(\mathbf{V}_i)$. Note that, by construction, the distribution of $(V_1, \ldots, V_n)$ is compatible with G, meaning that it satisfies the Markovian factorization induced by G. We use $M = (F, U)$ to denote the model that gives rise to factual variables.

In order to represent an intervention at level a for a given node $\mathbf{A}$, Pearl (2000, ch. 7 or 2009) defined the “Surgically modified model” $M_a = (F_a, U)$, considering $F_a = \{f_{a,i} : i \geq 1\}$, where $f_{a,i} = f_i$ if $\mathbf{V}_i \neq \mathbf{A}$ and for $\mathbf{V}_j = \mathbf{A}$, $f_{a,j} = a$. Counterfactual variables are defined by this new set of functions and the same disturbances $\{U_1, \ldots, U_n\}$, by the recurrence

$$V_{a,i} = f_{a,i}(PA_{a,i}, U_i) \,,$$

where $PA_{a,i}$ are the random variables (already defined by the recurrence) associated with the nodes in $PA_G(\mathbf{V}_i)$.

Before presenting our proposal for constructing counterfactual variables, recall that given a DAG G and a node $\mathbf{A}$ in G, $G_{\underline{\mathbf{A}}}$ is the graph obtained by removing from G all arrows emerging from $\mathbf{A}$. We will now introduce a new set of functions $F^a = \{f_i^a : i \geq 1\}$, compatible with $G_{\underline{\mathbf{A}}}$, which will allow the simultaneous definition of both the observed assignment random variable A associated with the node $\mathbf{A}$ and the counterfactual responses. To achieve this, if $\mathbf{A} \notin PA_G(\mathbf{V}_i)$ we get that $PA_{G_{\underline{\mathbf{A}}}}(\mathbf{V}_i) = PA_G(\mathbf{V}_i)$ and define f_i^a being equal to f_i. When $\mathbf{A} \in PA_G(\mathbf{V}_i)$, f_i^a is obtained by fixing the value a in the original function f_i. To be more precise, if $\mathbf{A} \in PA_G(\mathbf{V}_i)$, without loss of generality, it holds

$$f_i : \prod_{\mathbf{V}_j \in PA_G(\mathbf{V}_i) \setminus \mathbf{A}} \mathcal{V}_j \times \mathcal{A} \times \mathcal{U}_i \longrightarrow \mathcal{V}_i, \qquad [5]$$

where $\mathcal{A}$ denotes the set of possible values which can be assumed by the variables associated with node $\mathbf{A}$. Since $PA_{G_{\underline{\mathbf{A}}}}(\mathbf{V}_i) = PA_G(\mathbf{V}_i) \setminus \mathbf{A}$ and F^a should be compatible with $G_{\underline{\mathbf{A}}}$, we need f_i^a to satisfy the following condition:

$$f_i^a : \prod_{\mathbf{V}_j \in PA_G(\mathbf{V}_i) \setminus \mathbf{A}} \mathcal{V}_j \times \mathcal{U}_i \longrightarrow \mathcal{V}_i. \qquad [6]$$

Then, for $\bar{v}_i \in \prod_{\mathbf{V}_j \in PA_G(\mathbf{V}_i) \setminus \mathbf{A}} \mathcal{V}_j$, $u \in \mathcal{U}_i$, we define

$$f_i^a(\bar{v}_i, u) = f_i(\bar{v}_i, a, u).$$

Let V_i^a denote the variables obtained by the recurrence based on these new functions:

$$V_i^a = f_i^a(PA_i^a, U_i) \,,$$

where PA_i^a are the random variables (already defined by the recurrence) associated with the nodes in $PA_{G_{\underline{\mathbf{A}}}}(\mathbf{V}_i)$. Note that the distribution of $(V_1^a, \ldots, V_n^a)$ is compatible with $G_{\underline{\mathbf{A}}}$. Let $M^a = (F^a, U)$.

The following Lemma explains how variables defined under models M, M_a and M^a are related.

Lemma 3 *The random variables associated with both modified models $M_a = (F_a, U)$ and $M^a = (F^a, U)$ are the same, with the exception of those associated with node* $\mathbf{A}$:

$$V_{i,a} = V_i^a \quad \textit{if} \quad \mathbf{V}_\mathrm{i} \neq \mathbf{A} \, .$$

Variables associated with the node $\mathbf{A}$ *defined by $M = (F, U)$ and $M^a = (F^a, U)$, respectively, are equal*:

$$A = A^a \, .$$

Moreover, if $\mathbf{V}_i$ is not a descendent of $\mathbf{A}$, *we get that*

$$V_i = V_{a,i} = V_i^a \, .$$

Finally, under the assumption that the U_i are mutually independent, the joint distribution of the vector $(V_1^a, \ldots, V_n^a)$ factors according to $G_{\underline{\mathbf{A}}}$ (i.e. the variables are Markov with respect to $G_{\underline{\mathbf{A}}}$).

To conclude this section, we state the back door theorem, which was originally presented in Pearl (1993) and can be found, as most of the results presented in this work, in Pearl (2009). A new proof of this result is provided.

Theorem 4 ***The Back Door Criterion*** *Consider a set of nodes $\mathbf{L} \subset \{\mathbf{V}_1, \ldots, \mathbf{V}_n\}$, such that $\mathbf{L} \cap \mathbf{A} = \emptyset$. Assume that the following conditions hold*:
1. No element of $\mathbf{L}$ *is a descendent of* $\mathbf{A}$ *in G,*
2. $\mathbf{L}$ *blocks all back door paths from* $\mathbf{A}$ *to* $\mathbf{Y}$ *in G.*

Then, Y_a is independent of A given L and so

$$P(Y_a = y) = \sum_\ell P(Y = y | A = a, L = \ell) P(L = \ell) \, .$$

Proof: To prove that conditional ignorability holds, meaning that Y_a is independent of A given L, we note that under the assumption of Theorem 4, considering the results presented in Lemma 3, we get that

1. If no element of $\mathbf{L}$ is a descendent of $\mathbf{A}$ in G, then $L = L_a = L^a$.
2. If $\mathbf{L}$ blocks all back door paths from $\mathbf{A}$ to $\mathbf{Y}$ in G, then $\mathbf{A}$ and $\mathbf{Y}$ are d-separated by $\mathbf{L}$ in $G_{\underline{\mathbf{A}}}$, and so A^a and Y^a are independent given L^a.

Finally, resorting again to the results stated in Lemma 3, we also know that $A^a = A$ and $Y^a = Y_a$. So, if $\mathbf{L}$ satisfies both conditions [1] and [2], we can conclude that Y_a is independent of A given L. This means that conditional ignorability holds, as we meant to prove. Thus, the distribution of the counterfactual variables can be identified by the formula

$$P(Y_a = y) = \sum_\ell P(Y = y | A = a, L = \ell) P(L = \ell) \, .$$

□

3.2 Interventions on multiple nodes

Assume now that we wish to intervene in a set of nodes $\mathbf{A}_{set} = \{\mathbf{A}_1, \ldots, \mathbf{A}_k\}$. Consider $a_i \in \mathcal{A}_i$, where $\mathcal{A}_i$ denotes the support of variables associated with node $\mathbf{A}_i$, and let $a = (a_1, \ldots, a_k)$. Following the new surgically modified model, we will change the functions related to those nodes whose parents include some $\mathbf{A}_j$.

As in the one node case, given a DAG G, let $M = (F, U)$ denote the model (compatible with G) for factual variables $(V_1, \ldots, V_n)$. Let $(V_{a,1}, \ldots, V_{a,n})$ denote the vector of variables determined by the model $M_a = (F_a, U)$ proposed by Pearl, with $F_a = \{f_{a,i} : i \geq 1\}$, where $f_{a,i} = f_i$ if V_i does not belong to the set $\mathbf{A}_{set}$, and when $\mathbf{V}_j = \mathbf{A}_i$ for some i, $f_{a,j} = a_i$.

We will now generalize our construction presented for single node intervention in this new scenario. To do so, we consider $M^a = (F^a, U)$, for $F^a = \{f_i^a : i \geq 1\}$, compatible with $G_{\underline{\mathbf{A}_{set}}}$, the graph obtained removing in G all arrows emerging from the set $\mathbf{A}_{set}$. Note that the set of parents of a given node $\mathbf{V}_i$ in $G_{\underline{\mathbf{A}_{set}}}$ is obtained by eliminating from the set of parents of $\mathbf{V}_i$ in the original DAG G, all nodes in $\mathbf{A}^i_{set} = \mathbf{A}_{set} \cap PA_G(\mathbf{V}_i)$; namely, we have that $PA_{G_{\underline{\mathbf{A}_{set}}}}(\mathbf{V}_i) = PA_G(\mathbf{V}_i) \backslash \mathbf{A}^i_{set}$. Therefore, the definition of f_i^a depends on whether the set $\mathbf{A}^i_{set}$ is empty or not. Now, if $\mathbf{A}^i_{set} = \emptyset$, we get that $PA_{G_{\underline{\mathbf{A}_{set}}}}(\mathbf{V}_i) = PA_G(\mathbf{V}_i)$ and we define $f_i^a = f_i$. When $\mathbf{A}^i_{set} \neq \emptyset$, we can assume that

$$f_i : \prod_{\mathbf{V}_j \in PA_G(\mathbf{V}_i) \backslash \mathbf{A}^i_{set}} \mathcal{V}_j \times \prod_{\mathbf{A}_j \in \mathbf{A}^i_{set}} \mathcal{A}_j \times \mathcal{U}_i \longrightarrow \mathcal{V}_i, \qquad [7]$$

and consider

$$f_i^a : \prod_{\mathbf{V}_j \in PA_G(\mathbf{V}_i) \backslash \mathbf{A}^i_{set}} \mathcal{V}_j \times \mathcal{U}_i \longrightarrow \mathcal{V}_i, \qquad [8]$$

where for $\bar{v}_i \in \prod_{\mathbf{V}_j \in PA_G(\mathbf{V}_i) \backslash \mathbf{A}^i_{set}} \mathcal{V}_j$, $u \in \mathcal{U}_i$, we define

$$f_i^a(\bar{v}_i, u) = f_i(\bar{v}_i, a^i, u),$$

including in a^i all the coordinates of the vector $a = (a_1, \ldots, a_k)$ corresponding to the set $\mathbf{A}^i_{set}$: $a^i = (a_j : \mathbf{A}_j \in \mathbf{A}^i_{set})$. In other words, when $PA_G(\mathbf{V}_i) \cap \mathbf{A}_{set} \neq \emptyset$, each time the value of the variable related to the node $\mathbf{A}_j$ is required by the original function f_i (meaning that $\mathbf{A}_j \in \mathbf{A}^i_{set}$), we construct the function f_i^a fixing in f_i the value a_j.

Let $(V_1^a, \ldots, V_n^a)$ denote the vector of variables obtained by the recurrence based on these new functions (F^a) and disturbances U. Once more, we get that the distribution of $(V_1^a, \ldots, V_n^a)$ is compatible with $G_{\underline{\mathbf{A}_{set}}}$. The results are presented in what follows.

Lemma 5 *Let $A = (A_1, \ldots, A_k)$ and $A^a = (A_1^a, \ldots, A_k^a)$ denote the random variables related to the nodes $\mathbf{A}_1, \ldots, \mathbf{A}_k$, according to model M and M^a, respectively. If $\mathbf{W} \cap \mathbf{A}_{set} = \emptyset$, then the following version of the consistency assumption holds*:

$$\{A^a = a\ ,\ W^a = w\} = \{A = a\ ,\ W = w\}\ .$$

The random variables associated with both modified models $M_a = (F_a, U)$ and $M^a = (F^a, U)$ are the same, with the exception of those associated with nodes in $\mathbf{A}_{set}$:

$$V_{i,a} = V_i^a \ \ \text{if} \ \ \mathbf{V}_i \notin \mathbf{A}_{set}.$$

Under the assumption that the U_i are mutually independent, the joint distribution of the vector $(V_1^a, \ldots, V_n^a)$ factors according to $G_{\underline{\mathbf{A}_{set}}}$ (i.e. the variables are Markov with respect to $G_{\underline{\mathbf{A}_{set}}}$).

Finally, we include a new proof of the back door theorem, using the independences deduced from its assumptions and Lemma 5.

Theorem 6 ***Back Door Criterion: Many Nodes*** *Consider a set of nodes $\mathbf{L} \subset \{\mathbf{V}_1, \ldots, \mathbf{V}_n\}$, such that $\mathbf{L} \cap \mathbf{A}_{set} = \emptyset$. Assume that the following conditions hold*:

1. No element of $\mathbf{L}$ is a descendent of $\mathbf{A}_{set}$,

2. $\mathbf{L}$ blocks all back door paths from $\mathbf{A}_{set}$ to $\mathbf{Y}$ in G.

Then,

$$P(Y_a = y) = \sum_{\ell} P(Y = y | A = a, L = \ell) P(L = \ell)\ ,$$

with $a = (a_1, \ldots a_k)$.

Proof: Under the present assumptions we get that

1. If no element of $\mathbf{L}$ is a descendent of $\mathbf{A}_{set}$, then $L = L_a = L^a$.
2. If $\mathbf{L}$ blocks all back door paths from $\mathbf{A}_{set}$ to $\mathbf{Y}$ in G, then $\mathbf{A}_{set}$ and $\mathbf{Y}$ are d-separated by $\mathbf{L}$ in $G_{\underline{\mathbf{A}_{set}}}$, and so A^a and Y^a are independent given L^a.

Finally, if $\mathbf{L}$ satisfies the previous conditions, by Lemma 5, we get that $\{A^a = a,\ L^a = \ell\} = \{A = a,\ L = \ell\}$ for any ℓ, $\{Y^a = y, A^a = a,\ L^a = \ell\} = \{Y = y, A = a,\ L = \ell\}$ for any (ℓ, y) and so (under positivity),

$$\begin{aligned} P(Y_a = y) = P(Y^a = y) &= \sum_{\ell} P(Y^a = y | L^a = \ell) P(L^a = \ell) \\ &= \sum_{\ell} P(Y^a = y | L^a = \ell, A^a = a) P(L^a = \ell) = \sum_{\ell} P(Y = y | A = a, L = \ell) P(L = \ell)\ . \end{aligned}$$

□

4 FFRCISTG models

In the previous results, we used the rules of d-separation to detect independence or conditional independence between variables of a random vector. To do so, given a graph G, all we required from the joint distribution of our vector was compatibility with G. When variables are constructed following a NPSEM-IE, the Markov factorization induced by G holds automatically, and that is why our results are valid when the errors are independent.

However, the Markov factorization remains true under weaker conditions. For instance, let $v = (v_1, \ldots, v_n) \in \prod_{j=1}^{n} \mathcal{V}_j$ and call $v_{pa_G(\mathbf{V}_i)}$ the subvector of v containing the coordinates related to the nodes in the set $PA_G(\mathbf{V}_i)$, namely $v_{pa_G(\mathbf{V}_i)} = (v_j : \mathbf{V}_j \in PA_G(\mathbf{V}_i))$. If

$$\left\{ f_i\left(v_{pa_G(\mathbf{V}_i)}, U_i\right) : \mathbf{V}_i \in G \right\} \text{ are independent,} \quad \text{for all } v \in \prod\nolimits_{j=1}^{n} \mathcal{V}_j\ , \tag{9}$$

then, the distribution of the vector whose variables are constructed with $M = (F, U)$ is compatible with the graph G. This condition mainly defines the FFRCISTG models (Richardson and Robins 2013).

It is worth noting that if $M = (F, U)$ satisfies condition [9] relative to G, the intervened model $M^a = (F^a, U)$, defined in Section 3.2, also satisfies condition [9] relative to $G_{\underline{\mathbf{A}}_{set}}$, since

$$\left\{ f_i^a\left(v_{pa_{G_{\underline{\mathbf{A}}_{set}}}(\mathbf{V}_i)}, U_i\right) : \mathbf{V}_i \in G_{\underline{\mathbf{A}}_{set}} \right\} = \left\{ f_i\left(v^a_{pa_G(\mathbf{V}_i)}, U_i\right) : \mathbf{V}_i \in G \right\}$$

where $v^a_{pa_G(\mathbf{V}_i)}$ denotes the vector that results from replacing v_j with a_j for $\{j : \mathbf{A}_j \in \mathbf{A}_{set}\}$. Then, the distribution of the variables constructed using the model M^a, requiring that the errors satisfy only condition [1], factors according to the graph $G_{\underline{\mathbf{A}}_{set}}$, allowing the use of d-separation rules, and thus extending our results to this new model.

Acknowledgments: We are grateful to the referees and, especially to the editor, for their comments that have largely contributed to improve the original manuscript.

This research was partially supported by grants from the University of Buenos Aires, the National Council for Science and Technology of Argentina (CONICET) and the National Agency for Scientific Promotion (ANPCyT).

References

Balke, A., and Pearl, J. (1994). Probabilistic evaluation of counterfactual queries. UCLA Cognitive Systems Laboratory, Technical Report (R-213-A). In: Proceedings of the Twelfth National Conference on Artificial Intelligence, Seattle, WA, Volume I, 230–237.

Pearl, J. (1993). Comment: Graphical models, causality and intervention. Statistical Science, 8:266–269.

Pearl, J. (2000). Causality. Cambridge: Cambridge University Press.

Pearl, J. (2009). Causal inference in statistics: An overview. Statistics Surveys, 3:96–146. URL http://ftp.cs.ucla.edu/pub/stat_ser/r350.pdf

Richardson, T., and Robins, J. (2013). Single World Intervention Graphs (SWIGs): A unification of the counterfactual and graphical approaches to causality. Working Paper Number 128, Center for Statistics and the Social Sciences University of Washington.

Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period – Application to control of the healthy worker survivor effect. Mathematical Modelling, 7:1393–1512.

Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology, 66:688–701.