An Overview of The Back-Door and Front-Door Criteria

A Presentation Based On Sections 3.3 and 3.4 of Pearl's Causality

Mohammad Ali Javidian[1] Marco Valtorta[1]

[1]Department of Computer Science
University of South Carolina

June, 2018

Outline

Back-Door Criterion

Definition

A set of variables Z satisfies the *back-door criterion* relative to an ordered pair of variables (X_i, X_j) in a DAG G if:

Back-Door Criterion

Definition

A set of variables Z satisfies the *back-door criterion* relative to an ordered pair of variables (X_i, X_j) in a DAG G if:

(i) no node in Z is a descendant of X_i; and

Back-Door Criterion

Definition

A set of variables Z satisfies the *back-door criterion* relative to an ordered pair of variables (X_i, X_j) in a DAG G if:

(i) no node in Z is a descendant of X_i; and

(ii) Z blocks every path between X_i and X_j that contains an arrow into X_i.

Back-Door Criterion

Definition

A set of variables Z satisfies the *back-door criterion* relative to an ordered pair of variables (X_i, X_j) in a DAG G if:

(i) no node in Z is a descendant of X_i; and

(ii) Z blocks every path between X_i and X_j that contains an arrow into X_i.

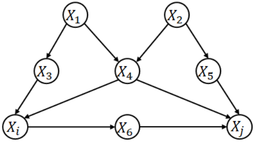


Figure: $S_1 = \{X_3, X_4\}$ and $S_2 = \{X_4, X_5\}$ would qualify under the back-door criterion, but $S_3 = \{X_4\}$ would not because X_4 does not d-separate X_i from X_j along the path $(X_i, X_3, X_1, X_4, X_2, X_5, X_j)$.

Outline

Back-Door Criterion

Back-Door Adjustment Theorem

If a set of variables Z satisfies the back-door criterion relative to (X, Y), then the causal effect of X on Y is identifiable and is given by the formula

Back-Door Criterion

Back-Door Adjustment Theorem

If a set of variables Z satisfies the back-door criterion relative to (X, Y), then the causal effect of X on Y is identifiable and is given by the formula

$$P(y|\hat{x}) = \sum_{z} P(y|x, z)P(z). \tag{1}$$

Outline

Proof of Theorem 1.

The proof originally offered in Pearl [2, 1993] is based on the observation that, when Z blocks all back-door paths from X to Y, setting $(X = x)$ or conditioning on $X = x$ has the same effect on Y.

Proof of Theorem 1.

The proof originally offered in Pearl [2, 1993] is based on the observation that, when Z blocks all back-door paths from X to Y, setting $(X = x)$ or conditioning on $X = x$ has the same effect on Y.

- The effect of an atomic intervention $do(X_i = x_i')$ is encoded by adding to G a link $F_i \rightarrow X_i$ (see the following Figure), where F_i is a new variable taking values in $\{do(x_i'), idle\}$, x_i' ranges over the domain of X_i, and *idle* represents *no intervention*. Thus, the new parent set of X_i in the augmented network is $PA_i' = PA_i \cup \{F_i\}$, and it is related to X_i by the conditional probability:

Proof of Theorem 1.

The proof originally offered in Pearl [2, 1993] is based on the observation that, when Z blocks all back-door paths from X to Y, setting $(X = x)$ or conditioning on $X = x$ has the same effect on Y.

- The effect of an atomic intervention $do(X_i = x_i')$ is encoded by adding to G a link $F_i \rightarrow X_i$ (see the following Figure), where F_i is a new variable taking values in $\{do(x_i'), idle\}$, x_i' ranges over the domain of X_i, and *idle* represents *no intervention*. Thus, the new parent set of X_i in the augmented network is $PA_i' = PA_i \cup \{F_i\}$, and it is related to X_i by the conditional probability:

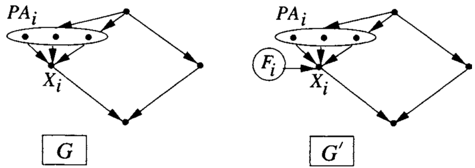


Figure: Representing external intervention F_i by an augmented network G'.

Proof of Theorem 1.

- $$P(x_i|pa_i') = \begin{cases} P(x_i|pa_i) & \text{if } F_i = \text{idle} \\ 0 & \text{if } F_i = do(x_i') \text{ and } x_i \neq x_i' \\ 1 & \text{if } F_i = do(x_i') \text{ and } x_i = x_i' \end{cases} \quad (2)$$

Proof of Theorem 1.

- $$P(x_i|pa_i') = \begin{cases} P(x_i|pa_i) & \text{if } F_i = \text{idle} \\ 0 & \text{if } F_i = do(x_i') \text{ and } x_i \neq x_i' \\ 1 & \text{if } F_i = do(x_i') \text{ and } x_i = x_i' \end{cases} \quad (2)$$

- The effect of the intervention $do(x_i')$ is to transform the original probability function $P(x_1, \ldots, x_n)$ into a new probability function $P(x_1, \ldots, x_n|\hat{x}_i')$, given by

Proof of Theorem 1.

- $$P(x_i|pa_i') = \begin{cases} P(x_i|pa_i) & \text{if } F_i = \text{idle} \\ 0 & \text{if } F_i = do(x_i') \text{ and } x_i \neq x_i' \\ 1 & \text{if } F_i = do(x_i') \text{ and } x_i = x_i' \end{cases} \quad (2)$$

- The effect of the intervention $do(x_i')$ is to transform the original probability function $P(x_1, \ldots, x_n)$ into a new probability function $P(x_1, \ldots, x_n|\hat{x}_i')$, given by

$$P(x_1, \ldots, x_n|\hat{x}_i') = P'(x_1, \ldots, x_n|F_i = do(x_i')) \quad (3)$$

Proof of Theorem 1.

- $$P(x_i|pa_i') = \begin{cases} P(x_i|pa_i) & \text{if } F_i = \text{idle} \\ 0 & \text{if } F_i = do(x_i') \text{ and } x_i \neq x_i' \\ 1 & \text{if } F_i = do(x_i') \text{ and } x_i = x_i' \end{cases} \quad (2)$$

- The effect of the intervention $do(x_i')$ is to transform the original probability function $P(x_1, \ldots, x_n)$ into a new probability function $P(x_1, \ldots, x_n|\hat{x}_i')$, given by

$$P(x_1, \ldots, x_n|\hat{x}_i') = P'(x_1, \ldots, x_n|F_i = do(x_i')) \quad (3)$$

- By writing $P(y|\hat{x})$ in terms of the augmented probability function P' in accordance with eq. (3) and conditioning on Z we obtain:

- $$P(x_i|pa_i') = \begin{cases} P(x_i|pa_i) & \text{if } F_i = \text{idle} \\ 0 & \text{if } F_i = do(x_i') \text{ and } x_i \neq x_i' \\ 1 & \text{if } F_i = do(x_i') \text{ and } x_i = x_i' \end{cases} \quad (2)$$

- The effect of the intervention $do(x_i')$ is to transform the original probability function $P(x_1, \ldots, x_n)$ into a new probability function $P(x_1, \ldots, x_n|\hat{x}_i')$, given by

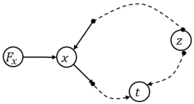
$$P(x_1, \ldots, x_n|\hat{x}_i') = P'(x_1, \ldots, x_n|F_i = do(x_i')) \quad (3)$$

- By writing $P(y|\hat{x})$ in terms of the augmented probability function P' in accordance with eq. (3) and conditioning on Z we obtain:

$$P(y|\hat{x}) = P'(y|F_x) = \sum_z P'(y|z, F_x)P'(z|F_x) = \sum_z P'(y|z, x, F_x)P'(z|F_x).$$

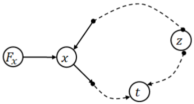
- According to the condition (i) of the definition of the back-door criterion, no node in Z is a descendant of x. So, all paths between F_x and each node $z \in Z$ have (at least) a collider i.e., $F_x \perp\!\!\!\perp z$ (see the following figure).

- According to the condition (i) of the definition of the back-door criterion, no node in Z is a descendant of x. So, all paths between F_x and each node $z \in Z$ have (at least) a collider i.e., $F_x \perp\!\!\!\perp z$ (see the following figure).



Proof of Theorem 1.

- According to the condition (i) of the definition of the back-door criterion, no node in Z is a descendant of x. So, all paths between F_x and each node $z \in Z$ have (at least) a collider i.e., $F_x \perp\!\!\!\perp z$ (see the following figure).



- Therfore, $P'(z|F_x) = P'(z) = P(z)$.

Proof of Theorem 1.

- If all back-door paths from X to Y are blocked, then all paths from F_X to Y must go through the children of X, and those would be blocked if we condition on X. The implication is that Y is independent of F_X given X,

Proof of Theorem 1.

- If all back-door paths from X to Y are blocked, then all paths from F_X to Y must go through the children of X, and those would be blocked if we condition on X. The implication is that Y is independent of F_X given X,

$$P(y|x, F_X = do(x)) = P(y|x) = P(y|x, F_X = \text{idle}) \tag{4}$$

- If all back-door paths from X to Y are blocked, then all paths from F_X to Y must go through the children of X, and those would be blocked if we condition on X. The implication is that Y is independent of F_X given X,

$$P(y|x, F_X = do(x)) = P(y|x) = P(y|x, F_X = \text{idle}) \tag{4}$$

- According to the back-door condition (ii), Z blocks every path between X and Y that contains an arrow into X. This observation together with eq. (4) implies that $Y \perp\!\!\!\perp F_x|(X, Z)$ i.e.,

- If all back-door paths from X to Y are blocked, then all paths from F_X to Y must go through the children of X, and those would be blocked if we condition on X. The implication is that Y is independent of F_X given X,

$$P(y|x, F_X = do(x)) = P(y|x) = P(y|x, F_X = \text{idle}) \tag{4}$$

- According to the back-door condition (ii), Z blocks every path between X and Y that contains an arrow into X. This observation together with eq. (4) implies that $Y \perp\!\!\!\perp F_x|(X, Z)$ i.e.,

$$P'(y|z, x, F_x) = P'(y|z, x) = P(y|z, x).$$

Outline

Front-Door Criterion

Definition

A set of variables Z satisfies the *front-door criterion* relative to an ordered pair of variables (X, Y) in a DAG G if:

Front-Door Criterion

Definition

A set of variables Z satisfies the *front-door criterion* relative to an ordered pair of variables (X, Y) in a DAG G if:

(i) Z intercepts all directed paths from X to Y;

Front-Door Criterion

Definition

A set of variables Z satisfies the *front-door criterion* relative to an ordered pair of variables (X, Y) in a DAG G if:

(i) Z intercepts all directed paths from X to Y;

(ii) there is no unblocked back-door path from X to Z; and

Front-Door Criterion

Definition

A set of variables Z satisfies the *front-door criterion* relative to an ordered pair of variables (X, Y) in a DAG G if:

(i) Z intercepts all directed paths from X to Y;

(ii) there is no unblocked back-door path from X to Z; and

(iii) all back-door paths from Z to Y are blocked by X.

Front-Door Criterion

Definition

A set of variables Z satisfies the *front-door criterion* relative to an ordered pair of variables (X, Y) in a DAG G if:

(i) Z intercepts all directed paths from X to Y;

(ii) there is no unblocked back-door path from X to Z; and

(iii) all back-door paths from Z to Y are blocked by X.

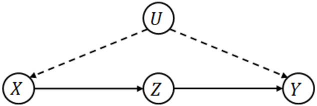


Figure: A diagram representing the front-door criterion.

Outline

1. The Back-Door Criterion
   - Definition 1. (Back-Door)
   - Theorem 1. (Back-Door Adjustment)
   - Proof of Theorem 1.
2. The Front-Door Criterion
   - Definition 2. (Front-Door)
   - Theorem 2. (Front-Door Adjustment)
3. *do* Calculus
4. Symbolic Derivation
5. Example

Front-Door Criterion

Front-Door Adjustment Theorem

If a set of variables Z satisfies the front-door criterion relative to (X, Y) and if $P(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by the formula

Front-Door Criterion

Front-Door Adjustment Theorem

If a set of variables Z satisfies the front-door criterion relative to (X, Y) and if $P(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by the formula

$$P(y|\hat{x}) = \sum_{z} P(z|x) \sum_{x'} P(y|x', z)P(x'). \tag{5}$$

Rules of *do* Calculus

Preliminary Notation

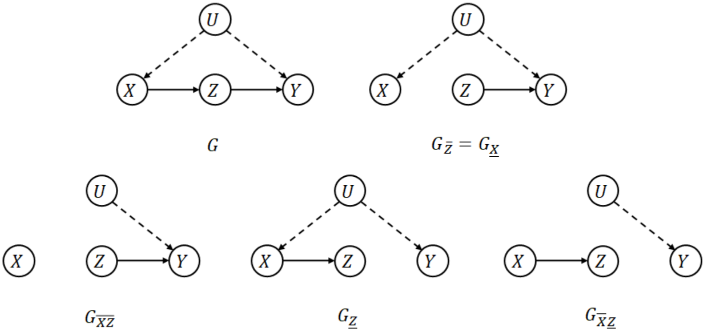


Figure: Subgraphs of G used in the derivation of causal effects.

Inference Rules

Rules of *do* Calculus

Rule 1 (Insertion/deletion of observations):

$$P(y|\hat{x}, z, w) = P(y|\hat{x}, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}}.$$

Inference Rules

Rules of *do* Calculus

Rule 1 (Insertion/deletion of observations):

$$P(y|\hat{x}, z, w) = P(y|\hat{x}, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}}.$$

Rule 2 (Action/observation exchange):

$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}.$$

Inference Rules

Rules of *do* Calculus

Rule 1 (Insertion/deletion of observations):

$$P(y|\hat{x}, z, w) = P(y|\hat{x}, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}}.$$

Rule 2 (Action/observation exchange):

$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}.$$

Rule 3 (Insertion/deletion of actions):

$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}, \overline{Z(W)}}}.$$

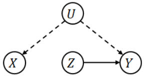
where $Z(W)$ is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

Symbolic Derivation of Causal Effects: An Example
Step 1: Compute $P(z|\hat{x})$

- $X \perp\!\!\!\perp Z$ in $G_{\underline{X}}$ because the path from X to Z is blocked by the converging arrows at Y.

Symbolic Derivation of Causal Effects: An Example
Step 1: Compute $P(z|\hat{x})$

- $X \perp\!\!\!\perp Z$ in $G_{\underline{X}}$ because the path from X to Z is blocked by the converging arrows at Y.

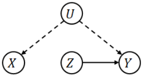


$G_{\underline{X}}$

Symbolic Derivation of Causal Effects: An Example

Step 1: Compute $P(z|\hat{x})$

- $X \perp\!\!\!\perp Z$ in $G_{\underline{X}}$ because the path from X to Z is blocked by the converging arrows at Y.



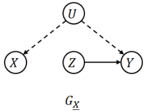
$G_{\underline{X}}$

- G satisfies the applicability condition for Rule 2:

$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}.$$

Symbolic Derivation of Causal Effects: An Example
Step 1: Compute $P(z|\hat{x})$

- $X \perp\!\!\!\perp Z$ in $G_{\underline{X}}$ because the path from X to Z is blocked by the converging arrows at Y.



- G satisfies the applicability condition for Rule 2:

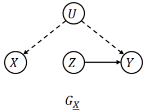
$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}.$$

- In Rule 2, set $y = z, x = \phi, z = x, w = \phi$:

Symbolic Derivation of Causal Effects: An Example

Step 1: Compute $P(z|\hat{x})$

- $X \perp\!\!\!\perp Z$ in $G_{\underline{X}}$ because the path from X to Z is blocked by the converging arrows at Y.



- G satisfies the applicability condition for Rule 2:

$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}.$$

- In Rule 2, set $y = z, x = \emptyset, z = x, w = \emptyset$:

$$P(z|\hat{x}) = P(z|x) \quad \text{because} \quad (Z \perp\!\!\!\perp X)_{G_{\underline{X}}}.$$

Symbolic Derivation of Causal Effects: An Example

Step 2: Compute $P(y|\hat{z})$

- $P(y|\hat{z}) = \sum_x P(y|x, \hat{z})P(x|\hat{z})$.

Symbolic Derivation of Causal Effects: An Example

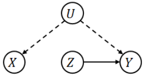
Step 2: Compute $P(y|\hat{z})$

- $P(y|\hat{z}) = \sum_x P(y|x,\hat{z})P(x|\hat{z})$.
- $X \perp\!\!\!\perp Z$ in $G_{\overline{Z}}$ because the path from X to Z is blocked by the converging arrows at Y.

Symbolic Derivation of Causal Effects: An Example

Step 2: Compute $P(y|\hat{z})$

- $P(y|\hat{z}) = \sum_x P(y|x,\hat{z})P(x|\hat{z})$.
- $X \perp\!\!\!\perp Z$ in $G_{\overline{Z}}$ because the path from X to Z is blocked by the converging arrows at Y.

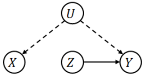


$G_{\overline{Z}}$

Symbolic Derivation of Causal Effects: An Example

Step 2: Compute $P(y|\hat{z})$

- $P(y|\hat{z}) = \sum_x P(y|x,\hat{z})P(x|\hat{z})$.
- $X \perp\!\!\!\perp Z$ in $G_{\overline{Z}}$ because the path from X to Z is blocked by the converging arrows at Y.



$G_{\overline{Z}}$

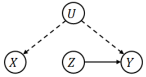
- G satisfies the applicability condition for Rule 3:

$$P(y|\hat{x},\hat{z},w) = P(y|\hat{x},w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X,W)_{G_{\overline{X},\overline{Z(W)}}}.$$

Symbolic Derivation of Causal Effects: An Example

Step 2: Compute $P(y|\hat{z})$

- $P(y|\hat{z}) = \sum_x P(y|x,\hat{z})P(x|\hat{z})$.
- $X \perp\!\!\!\perp Z$ in $G_{\overline{Z}}$ because the path from X to Z is blocked by the converging arrows at Y.



$G_{\overline{Z}}$

- G satisfies the applicability condition for Rule 3:

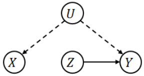
$$P(y|\hat{x},\hat{z},w) = P(y|\hat{x},w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X,W)_{G_{\overline{X},\overline{Z(W)}}}.$$

- In Rule 3, set $y = x, x = \emptyset, z = z, w = \emptyset$:

Symbolic Derivation of Causal Effects: An Example

Step 2: Compute $P(y|\hat{z})$

- $P(y|\hat{z}) = \sum_x P(y|x,\hat{z})P(x|\hat{z})$.
- $X \perp\!\!\!\perp Z$ in $G_{\overline{Z}}$ because the path from X to Z is blocked by the converging arrows at Y.



$G_{\overline{Z}}$

- G satisfies the applicability condition for Rule 3:

$$P(y|\hat{x},\hat{z},w) = P(y|\hat{x},w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X,W)_{G_{\overline{X},\overline{Z(W)}}}.$$

- In Rule 3, set $y = x, x = \emptyset, z = z, w = \emptyset$:

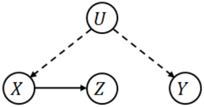
$$P(x|\hat{z}) = P(x) \quad \text{because} \quad (Z \perp\!\!\!\perp X)_{G_{\overline{Z}}}.$$

Symbolic Derivation of Causal Effects: An Example
Step 2 (continued): Compute $P(y|\hat{z})$

- $(Z \perp\!\!\!\perp Y|X)_{G_{\underline{Z}}}$ because the path from Z to Y is blocked by X.

Symbolic Derivation of Causal Effects: An Example
Step 2 (continued): Compute $P(y|\hat{z})$

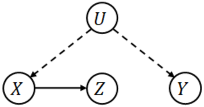
- $(Z \perp\!\!\!\perp Y|X)_{G_{\underline{Z}}}$ because the path from Z to Y is blocked by X.



$G_{\underline{Z}}$

Symbolic Derivation of Causal Effects: An Example
Step 2 (continued): Compute $P(y|\hat{z})$

- $(Z \perp\!\!\!\perp Y|X)_{G_{\underline{Z}}}$ because the path from Z to Y is blocked by X.

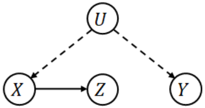


$G_{\underline{Z}}$

- G satisfies the applicability condition for Rule 2:
  $P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}$.

Symbolic Derivation of Causal Effects: An Example
Step 2 (continued): Compute $P(y|\hat{z})$

- $(Z \perp\!\!\!\perp Y|X)_{G_{\underline{Z}}}$ because the path from Z to Y is blocked by X.

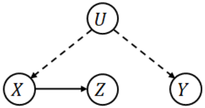


$G_{\underline{Z}}$

- G satisfies the applicability condition for Rule 2:
  $P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}$.
- In Rule 2, set $y = y, x = \emptyset, z = z, w = x$:

Symbolic Derivation of Causal Effects: An Example
Step 2 (continued): Compute $P(y|\hat{z})$

- $(Z \perp\!\!\!\perp Y|X)_{G_{\underline{Z}}}$ because the path from Z to Y is blocked by X.



$G_{\underline{Z}}$

- G satisfies the applicability condition for Rule 2:
  $P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}$.
- In Rule 2, set $y = y, x = \emptyset, z = z, w = x$:

$$P(y|x, \hat{z}) = P(y|x, z) \quad \text{because} \quad (Z \perp\!\!\!\perp Y|X)_{G_{\underline{Z}}}.$$

Symbolic Derivation of Causal Effects: An Example

Step 2 (continued): Compute $P(y|\hat{z})$

- $P(y|\hat{z}) = \sum_x P(y|x, \hat{z})P(x|\hat{z}) = \sum_x P(y|x, z)P(x)$.

Symbolic Derivation of Causal Effects: An Example

Step 2 (continued): Compute $P(y|\hat{z})$

- $P(y|\hat{z}) = \sum_x P(y|x,\hat{z})P(x|\hat{z}) = \sum_x P(y|x,z)P(x)$.
- This formula is a special case of the back-door formula in Theorem 1.

Symbolic Derivation of Causal Effects: An Example

Step 3: Compute $P(y|\hat{x})$

- $P(y|\hat{x}) = \sum_z P(y|z, \hat{x})P(z|\hat{x})$.

Symbolic Derivation of Causal Effects: An Example

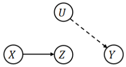
Step 3: Compute $P(y|\hat{x})$

- $P(y|\hat{x}) = \sum_z P(y|z,\hat{x})P(z|\hat{x})$.
- $(Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}\underline{Z}}}$ because there is no outgoing edge from Z and no incoming edge to X in $G_{\overline{X}\underline{Z}}$.

Symbolic Derivation of Causal Effects: An Example

Step 3: Compute $P(y|\hat{x})$

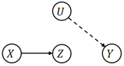
- $P(y|\hat{x}) = \sum_z P(y|z,\hat{x})P(z|\hat{x})$.
- $(Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}\underline{Z}}}$ because there is no outgoing edge from Z and no incoming edge to X in $G_{\overline{X}\underline{Z}}$.



$G_{\overline{X}\underline{Z}}$

Symbolic Derivation of Causal Effects: An Example
Step 3: Compute $P(y|\hat{x})$

- $P(y|\hat{x}) = \sum_z P(y|z,\hat{x})P(z|\hat{x})$.
- $(Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}\underline{Z}}}$ because there is no outgoing edge from Z and no incoming edge to X in $G_{\overline{X}\underline{Z}}$.

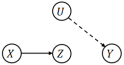


- G satisfies the applicability condition for Rule 2:

$$P(y|\hat{x},\hat{z},w) = P(y|\hat{x},z,w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X,W)_{G_{\overline{X}\underline{Z}}}.$$

Symbolic Derivation of Causal Effects: An Example
Step 3: Compute $P(y|\hat{x})$

- $P(y|\hat{x}) = \sum_z P(y|z,\hat{x})P(z|\hat{x})$.
- $(Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}\underline{Z}}}$ because there is no outgoing edge from Z and no incoming edge to X in $G_{\overline{X}\underline{Z}}$.



$G_{\overline{X}\underline{Z}}$

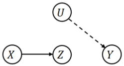
- G satisfies the applicability condition for Rule 2:

$$P(y|\hat{x},\hat{z},w) = P(y|\hat{x},z,w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X,W)_{G_{\overline{X}\underline{Z}}}.$$

- In Rule 2, set $y = y, x = x, z = z, w = \emptyset$:

Symbolic Derivation of Causal Effects: An Example
Step 3: Compute $P(y|\hat{x})$

- $P(y|\hat{x}) = \sum_z P(y|z,\hat{x})P(z|\hat{x})$.
- $(Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}\underline{Z}}}$ because there is no outgoing edge from Z and no incoming edge to X in $G_{\overline{X}\underline{Z}}$.



$G_{\overline{X}\underline{Z}}$

- G satisfies the applicability condition for Rule 2:

$$P(y|\hat{x},\hat{z},w) = P(y|\hat{x},z,w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X,W)_{G_{\overline{X}\underline{Z}}}.$$

- In Rule 2, set $y = y, x = x, z = z, w = \emptyset$:

$$P(y|z,\hat{x}) = P(y|\hat{z},\hat{x}) \quad \text{because} \quad (Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}\underline{Z}}}.$$

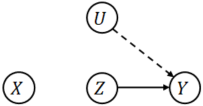
Symbolic Derivation of Causal Effects: An Example
Step 3 (continued): Compute $P(y|\hat{x})$

- $(Y \perp\!\!\!\perp X|Z)_{G_{\overline{XZ}}}$ because there is no incoming edge to X and no outgoing edge from X in $G_{\overline{XZ}}$.

Symbolic Derivation of Causal Effects: An Example

Step 3 (continued): Compute $P(y|\hat{x})$

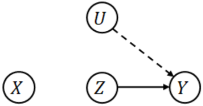
- $(Y \perp\!\!\!\perp X|Z)_{G_{\overline{XZ}}}$ because there is no incoming edge to X and no outgoing edge from X in $G_{\overline{XZ}}$.



$G_{\overline{XZ}}$

Symbolic Derivation of Causal Effects: An Example
Step 3 (continued): Compute $P(y|\hat{x})$

- $(Y \perp\!\!\!\perp X|Z)_{G_{\overline{XZ}}}$ because there is no incoming edge to X and no outgoing edge from X in $G_{\overline{XZ}}$.



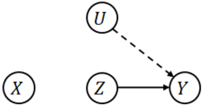
$G_{\overline{XZ}}$

- G satisfies the applicability condition for Rule 3:

$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X},\overline{Z(W)}}}.$$

Symbolic Derivation of Causal Effects: An Example
Step 3 (continued): Compute $P(y|\hat{x})$

- $(Y \perp\!\!\!\perp X|Z)_{G_{\overline{XZ}}}$ because there is no incoming edge to X and no outgoing edge from X in $G_{\overline{XZ}}$.



$G_{\overline{XZ}}$

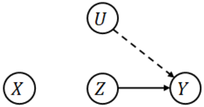
- G satisfies the applicability condition for Rule 3:

$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}, \overline{Z(W)}}}.$$

- In Rule 3, set $y = y, x = z, z = x, w = \emptyset$:

Symbolic Derivation of Causal Effects: An Example
Step 3 (continued): Compute $P(y|\hat{x})$

- $(Y \perp\!\!\!\perp X|Z)_{G_{\overline{XZ}}}$ because there is no incoming edge to X and no outgoing edge from X in $G_{\overline{XZ}}$.



$G_{\overline{XZ}}$

- G satisfies the applicability condition for Rule 3:

$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}, \overline{Z(W)}}}.$$

- In Rule 3, set $y = y, x = z, z = x, w = \emptyset$:

$$P(y|\hat{z}, \hat{x}) = P(y|\hat{z}) \quad \text{because} \quad (Y \perp\!\!\!\perp Z|X)_{G_{\overline{XZ}}}.$$

Symbolic Derivation of Causal Effects: An Example

Step 3 (continued): Compute $P(y|\hat{x})$

- $P(y|\hat{x}) = \sum_z P(y|z,\hat{x})P(z|\hat{x}) = \sum_z P(z|x) \sum_{x'} P(y|x',z)P(x')$.

Example: Smoking and the Genotype Theory

- The tobacco industry has managed to forestall antismoking legislation by arguing that the observed correlation between smoking and lung cancer could be explained by some sort of carcinogenic genotype (U) that involves inborn craving for nicotine.

Example: Smoking and the Genotype Theory

- The tobacco industry has managed to forestall antismoking legislation by arguing that the observed correlation between smoking and lung cancer could be explained by some sort of carcinogenic genotype (U) that involves inborn craving for nicotine.

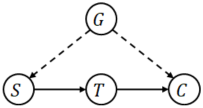


Figure: A diagram representing the story of smoking and the genotype (X=S=Smoking, Z=T=Tar, Y=C=Cancer, and U=G=Genotype (unobserved)).

Example: Smoking and the Genotype Theory

continued

	Group Type	$P(x, z)$ Group Size (% of Population)	$P(Y = 1 \mid x, z)$ % of Cancer Cases in Group
$X = 0, Z = 0$	Nonsmokers, No tar	47.5	10
$X = 1, Z = 0$	Smokers, No tar	2.5	90
$X = 0, Z = 1$	Nonsmokers, Tar	2.5	5
$X = 1, Z = 1$	Smokers, Tar	47.5	85

A hypothetical data set: 95% of smokers and 5% of nonsmokers have developed high levels of tar in their lungs. Moreover, 81% of subjects with tar deposits have developed lung cancer, compared to only 14% among those with no tar deposits.

Example: Smoking and the Genotype Theory

continued

	Group Type	$P(x, z)$ Group Size (% of Population)	$P(Y = 1 \mid x, z)$ % of Cancer Cases in Group
$X = 0, Z = 0$	Nonsmokers, No tar	47.5	10
$X = 1, Z = 0$	Smokers, No tar	2.5	90
$X = 0, Z = 1$	Nonsmokers, Tar	2.5	5
$X = 1, Z = 1$	Smokers, Tar	47.5	85

A hypothetical data set: 95% of smokers and 5% of nonsmokers have developed high levels of tar in their lungs. Moreover, 81% of subjects with tar deposits have developed lung cancer, compared to only 14% among those with no tar deposits.

- These results seem to prove that smoking is a major contributor to lung cancer. However, the tobacco industry might argue that the table tells a different story that smoking actually decreases ones risk of lung cancer. Their argument goes as follows. If you decide to smoke, then your chances of building up tar deposits are 95%, compared to 5% if you decide not to smoke. In order to evaluate the effect of tar deposits, we look separately at two groups, smokers and nonsmokers. The table shows that tar deposits have a protective effect in both groups: in smokers, tar deposits lower cancer rates from 90% to 85%; in nonsmokers, they lower cancer rates from 10% to 5%. Thus, regardless of whether I have a natural craving for nicotine, I should be seeking the protective effect of tar deposits in my lungs, and smoking offers a very effective means of acquiring those deposits.

Example: Smoking and the Genotype Theory

continued

To settle the dispute between the two interpretations, we now apply the front-door formula (eq. (5)) to the data in the Table of the previous slide:

Example: Smoking and the Genotype Theory
continued

To settle the dispute between the two interpretations, we now apply the front-door formula (eq. (5)) to the data in the Table of the previous slide:

$$P(y|\hat{x}) = \sum_{z} P(z|x) \sum_{x'} P(y|x', z) P(x').$$

Example: Smoking and the Genotype Theory

continued

To settle the dispute between the two interpretations, we now apply the front-door formula (eq. (5)) to the data in the Table of the previous slide:

$$P(y|\hat{x}) = \sum_{z} P(z|x) \sum_{x'} P(y|x', z) P(x').$$

$$\begin{aligned}
P(Y = 1 \,|\, do(X = 1)) &= .05(.10 \times .50 + .90 \times .50) \\
&\quad + .95(.05 \times .50 + .85 \times .50) \\
&= .05 \times .50 + .95 \times .45 = .4525, \\
P(Y = 1 \,|\, do(X = 0)) &= .95(.10 \times .50 + .90 \times .50) \\
&\quad + .05(.05 \times .50 + .85 \times .50) \\
&= .95 \times .50 + .05 \times .45 = .4975.
\end{aligned}$$

Reference For Further Reading

J. Pearl.
Causality. Models, reasoning, and inference.
Cambridge University Press, 2009.

J. Pearl.
Comment: Graphical Models, Causality and Intervention.
Statistical Science, 8(3):266–269, 1993.