Equivalence of G-formula and IPTW

Kazuki Yoshida

@kaz_yos kaz-yos

2019–04–21

1 INTRODUCTION

1.1 IPTW and G-formula

- Here we will demonstrate the equivalence of the IPTW [Robins et al., 2000] and the g-formula [Robins, 1986] given saturated models.
- Special thanks to @edwardhkennedy!

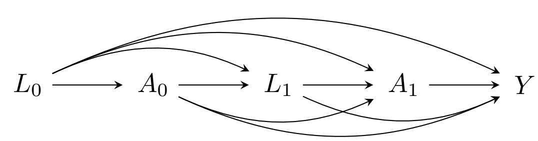
1.2 Notations

$$
\begin{aligned}
Y &: \text{Outcome measured at the end of the study} \\
Y^{a_0} &: \text{Counterfactual outcome with intervention at time 0 only} \\
Y^{a_0,a_1} &: \text{Counterfactual outcome with intervention at time 0 and 1} \\
L_0 &: \text{Baseline covariates} \\
A_0 &: \text{Baseline treatment assignment} \\
L_1 &: \text{Post-baseline covariates} \\
A_1 &: \text{Post-baseline treatment assignment}
\end{aligned}
$$

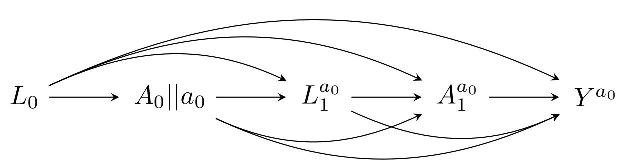
Here we assume both L_t and A_t are discrete.

1.3 Causal structure

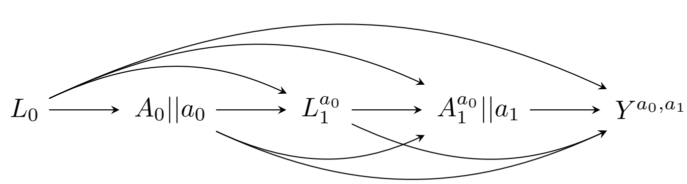
1. Original DAG



2. Single time point intervention SWIG



3. Two time point intervention SWIG



2 Single time point strategy

2.1 Identifiability Conditions

- We will follow the terminologies in [Hernan and Robins, 2019] (Chapter 3).
- 1. **Consistency**: the values of treatment under comparison correspond to well-defined interventions that, in turn, correspond to the versions of treatment in the data

$$Y_i = Y_i^{a_0} \text{ if } A_{0i} = a_0 \text{ for all } a_0$$

- 2. **Exchangeability**: the conditional probability of receiving every value of treatment, though not decided by the investigators, depends only on the measured covariates

$$A_0 \perp\!\!\!\perp Y^{a_0} | L_0 \text{ for all } a_0$$

- 3. **Positivity**: the conditional probability of receiving every value of treatment is greater than zero, i.e., positive

$$f(A_0 = a_0 | L_0 = l_0) > 0 \text{ for all } a_0, l_0 \text{ where } f(L_0 = l_0) > 0$$

2.2 G-formula

This follows the Technical Point 2.3 in [Hernan and Robins, 2019].

$$\begin{aligned}
&\quad \text{By iterative expectation}\\
E[Y^{a_0}] &= E[E[Y^{a_0}|L_0]]\\
&\quad \text{By conditional exchangeability: } Y^{a_0} \perp\!\!\!\perp A_0|L_0\\
&= E[E[Y^{a_0}|A_0, L_0]]\\
&\quad \text{By exchangeability, } E[Y^{a_0}|A_0, L_0] = E[Y^{a_0}|A_0 = a_0, L_0]\\
&= E[E[Y^{a_0}|A_0 = a_0, L_0]]\\
&\quad \text{By consistency}\\
&= E[E[Y|A_0 = a_0, L_0]]\\
&\quad \text{Make outer expectation explicit sum}\\
&= \sum_{l_0} E[Y|A_0 = a_0, L_0 = l_0] f(L_0 = l_0)\\
&= \text{Conditional mean averaged over } L_0
\end{aligned}$$

2.3 IPTW

This follows the Technical Point 2.3 in [Hernan and Robins, 2019].

$$\begin{aligned}
&\quad \text{By iterative expectation}\\
E[Y^{a_0}] &= E[E[Y^{a_0}|L_0]]\\
&\quad \text{Insert a carefully-crafted expression that is 1.}\\
&= E\left[\frac{f(A_0 = a_0|L_0)}{f(A_0 = a_0|L_0)} E[Y^{a_0}|L_0]\right]\\
&\quad \text{Using probability = expectation of indicator}\\
&= E\left[\frac{E[I(A_0 = a_0)|L_0]}{f(A_0 = a_0|L_0)} E[Y^{a_0}|L_0]\right]\\
&\quad \text{Conditional exchangeability: } Y^{a_0} \perp\!\!\!\perp A_0|L_0\\
&\quad \text{This allows merging the two inner expectations.}\\
&= E\left[\frac{1}{f(A_0 = a_0|L_0)} E[I(A_0 = a_0) Y^{a_0}|L_0]\right]\\
&\quad \text{By stability, } g(L_0) = E[g(L_0)|L_0].\\
&\quad \text{i.e., a function of } L_0 \text{ only (IPTW expression) can go into } E[\cdot|L_0]\\
&= E\left[E\left[\frac{1}{f(A_0 = a_0|L_0)} I(A_0 = a_0) Y^{a_0} \middle| L_0\right]\right]\\
&\quad \text{Reversing iterative expectation (tower property)}
\end{aligned}$$

$$= E\left[\frac{1}{f(A_0 = a_0|L_0)} I(A_0 = a_0) Y^{a_0}\right]$$

By consistency, $I(A_0 = a_0)Y^{a_0} = I(A_0 = a_0)Y = Y$ for $A_0 = a_0$.
Also, $I(A_0 = a_0)Y^{a_0} = 0 = I(A_0 = a_0)Y$ for $A_0 \neq a_0$.
Thus, $I(A_0 = a_0)Y^{a_0} = I(A_0 = a_0)Y$ regardless of A_0.

$$= E\left[\frac{1}{f(A_0 = a_0|L_0)} I(A_0 = a_0) Y\right]$$

$= \text{IPTW mean of } Y$ for group $A_0 = a_0$

- Law of Total Expectation (Iterative Expectation): https://en.wikipedia.org/wiki/Law_of_total_expectation
- Stability and Tower Property: https://en.wikipedia.org/wiki/Conditional_expectation#Basic_properties
- Indicator function: https://www.statlect.com/fundamentals-of-probability/indicator-functions

3 Multiple time point strategy

3.1 Identifiability Conditions

- We will follow the terminologies in [Hernan and Robins, 2019] (Chapter 19, Technical Point 19.2).
- 1. **Consistency**: the values of treatment under comparison correspond to well-defined interventions that, in turn, correspond to the versions of treatment in the data

$$Y_i = Y_i^{a_0,a_1} \text{ if } A_{0i} = a_0, A_{1i} = a_1 \text{ for all } a_0, a_1$$

- 2. **Exchangeability**: the conditional probability of receiving every value of treatment, though not decided by the investigators, depends only on the measured covariates

$$A_0 \perp\!\!\!\perp Y^{a_0,a_1}|L_0 \text{ for all } a_0, a_1$$
$$A_1 \perp\!\!\!\perp Y^{a_0,a_1}|L_1, A_0 = a_0, L_0 \text{ for all } a_0, a_1$$

- 3. **Positivity**: the conditional probability of receiving every value of treatment is greater than zero, i.e., positive

$$f(A_0 = a_0|L_0 = l_0) > 0 \text{ for all } a_0, l_0 \text{ where } f(L_0 = l_0) > 0$$
$$f(A_1 = a_1|L_1 = l_1, A_0 = a_0, L_0 = l_0) > 0 \text{ for all } a_1, l_1, a_0, l_0 \text{ where } f(L_1 = l_1, A_0 = a_0, L_0 = l_0) > 0$$

3.2 G-formula

This follows the Technical Point 2.3 in [Hernan and Robins, 2019].

By iterative expectation

$$E[Y^{a_0,a_1}] = E[E[Y^{a_0,a_1}|L_0]]$$

By conditional exchangeability: $Y^{a_0,a_1} \perp\!\!\!\perp A_0|L_0$

$$= E[E[Y^{a_0,a_1}|A_0, L_0]]$$

By exchangeability, $E[Y^{a_0,a_1}|A_0, L_0] = E[Y^{a_0,a_1}|A_0 = a_0, L_0]$

$$= E[E[Y^{a_0,a_1}|A_0 = a_0, L_0]]$$

By iterative expectation

$$= E[E[E[Y^{a_0,a_1}|L_1, A_0 = a_0, L_0]|A_0 = a_0, L_0]]$$

By conditional exchangeability: $Y^{a_0,a_1} \perp\!\!\!\perp A_1|L_1, A_0, L_0$

$$= E[E[E[Y^{a_0,a_1}|A_1, L_1, A_0 = a_0, L_0]|A_0 = a_0, L_0]]$$

By exchangeability,
$E[Y^{a_0,a_1}|A_1, L_1, A_0 = a_0, L_0] = E[Y^{a_0,a_1}|A_1 = a_1, L_1, A_0 = a_0, L_0]$

$$= E[E[E[Y^{a_0,a_1}|A_1 = a_1, L_1, A_0 = a_0, L_0]|A_0 = a_0, L_0]]$$

By consistency

$$= E[E[E[Y|A_1 = a_1, L_1, A_0 = a_0, L_0]|A_0 = a_0, L_0]]$$

$$\begin{aligned}
&\text{Make outer expectations explicit sums}\\
&= \sum_{l_0} \sum_{l_1} E[Y|A_1 = a_1, L_1 = l_1, A_0 = a_0, L_0 = l_0]\\
&\times f(L_1 = l_1|A_0 = a_0, L_0 = l_0) f(L_0 = l_0)
\end{aligned}$$

3.3 IPTW

This follows the Technical Point 2.3 in [Hernan and Robins, 2019] and an input from @edwardhkennedy.

$$\begin{aligned}
&\text{By iterative expectation}\\
E\left[Y^{a_0,a_1}\right] &= E\left[E\left[Y^{a_0,a_1}|L_0\right]\right]\\
&\text{Insert a carefully-crafted expression that is 1.}\\
&= E\left[\frac{f(A_0 = a_0|L_0)}{f(A_0 = a_0|L_0)} E\left[Y^{a_0,a_1}|L_0\right]\right]\\
&\text{Using probability = expectation of indicator}\\
&= E\left[\frac{E\left[I(A_0 = a_0)|L_0\right]}{f(A_0 = a_0|L_0)} E\left[Y^{a_0,a_1}|L_0\right]\right]\\
&\text{By conditional exchangeability: } Y^{a_0,a_1} \perp\!\!\!\perp A_0|L_0\\
&\text{Thus, product of expectation = expectation of product}\\
&= E\left[\frac{1}{f(A_0 = a_0|L_0)} E\left[I(A_0 = a_0)Y^{a_0,a_1}|L_0\right]\right]\\
&\text{Law of total expectation}\\
&= E\left[\frac{1}{f(A_0 = a_0|L_0)} E\left[I(A_0 = a_0)Y^{a_0,a_1}|A_0 = a_0, L_0\right]\right.\\
&\qquad \left. + \frac{1}{f(A_0 = a_0|L_0)} E\left[I(A_0 = a_0)Y^{a_0,a_1}|A_0 \neq a_0, L_0\right]\right]\\
&\text{Indicator in second inner expectation = 0}\\
&= E\left[\frac{1}{f(A_0 = a_0|L_0)} E\left[I(A_0 = a_0)Y^{a_0,a_1}|A_0 = a_0, L_0\right]\right]\\
&\text{By iterative expectation}\\
&= E\left[\frac{1}{f(A_0 = a_0|L_0)} E\left[E\left[I(A_0 = a_0)Y^{a_0,a_1}|L_1, A_0 = a_0, L_0\right]|A_0 = a_0, L_0\right]\right]\\
&\text{Insert a carefully-crafted expression that is 1.}\\
&= E\left[\frac{1}{f(A_0 = a_0|L_0)} \times\right.\\
&\qquad \left. E\left[\frac{f(A_1 = a_1|L_1, A_0 = a_0, L_0)}{f(A_1 = a_1|L_1, A_0 = a_0, L_0)} E\left[I(A_0 = a_0)Y^{a_0,a_1}|L_1, A_0 = a_0, L_0\right] \middle| A_0 = a_0, L_0\right]\right]\\
&\text{Using probability = expectation of indicator}\\
&= E\left[\frac{1}{f(A_0 = a_0|L_0)} \times\right.\\
&\qquad \left. E\left[\frac{E\left[I(A_1 = a_1)|L_1, A_0 = a_0, L_0\right]}{f(A_1 = a_1|L_1, A_0 = a_0, L_0)} E\left[I(A_0 = a_0)Y^{a_0,a_1}|L_1, A_0 = a_0, L_0\right] \middle| A_0 = a_0, L_0\right]\right]\\
&\text{By conditional exchangeability: } Y^{a_0,a_1} \perp\!\!\!\perp A_1|L_1, A_0 = a_0, L_0\\
&\text{Thus, product of expectation = expectation of product}\\
&= E\left[\frac{1}{f(A_0 = a_0|L_0)} \times\right.\\
&\qquad \left. E\left[\frac{E\left[I(A_0 = a_0)I(A_1 = a_1)Y^{a_0,a_1}|L_1, A_0 = a_0, L_0\right]}{f(A_1 = a_1|L_1, A_0 = a_0, L_0)} \middle| A_0 = a_0, L_0\right]\right]\\
&\text{IPTW at time 1 is a constant given } L_1, A_0 = a_0, L_0
\end{aligned}$$

$$= E\left[\frac{1}{f(A_0 = a_0|L_0)}\times\right.$$
$$\left. E\left[E\left[\frac{I(A_0 = a_0)I(A_1 = a_1)Y^{a_0,a_1}}{f(A_1 = a_1|L_1, A_0 = a_0, L_0)}\middle|L_1, A_0 = a_0, L_0\right]\middle|A_0 = a_0, L_0\right]\right]$$

Reverse iterative expectation (tower property)

$$= E\left[\frac{1}{f(A_0 = a_0|L_0)}E\left[\frac{I(A_0 = a_0)I(A_1 = a_1)Y^{a_0,a_1}}{f(A_1 = a_1|L_1, A_0 = a_0, L_0)}\middle|A_0 = a_0, L_0\right]\right]$$

Add a second term that is zero by indicator $I(A_0 = a_0)$

$$= E\left[\frac{1}{f(A_0 = a_0|L_0)}E\left[\frac{I(A_0 = a_0)I(A_1 = a_1)Y^{a_0,a_1}}{f(A_1 = a_1|L_1, A_0 = a_0, L_0)}\middle|A_0 = a_0, L_0\right]\right.$$
$$\left.+\frac{1}{f(A_0 = a_0|L_0)}E\left[\frac{I(A_0 = a_0)I(A_1 = a_1)Y^{a_0,a_1}}{f(A_1 = a_1|L_1, A_0 = a_0, L_0)}\middle|A_0 \neq a_0, L_0\right]\right]$$

Thus, we can drop conditioning on A_0.

$$= E\left[\frac{1}{f(A_0 = a_0|L_0)}E\left[\frac{I(A_0 = a_0)I(A_1 = a_1)Y^{a_0,a_1}}{f(A_1 = a_1|L_1, A_0 = a_0, L_0)}\middle|L_0\right]\right]$$

IPTW at time 0 is a constant given L_0

$$= E\left[E\left[\frac{I(A_0 = a_0)I(A_1 = a_1)Y^{a_0,a_1}}{f(A_0 = a_0|L_0)f(A_1 = a_1|L_1, A_0 = a_0, L_0)}\middle|L_0\right]\right]$$

Reverse iterative expectation (tower property)

$$= E\left[\frac{I(A_0 = a_0)I(A_1 = a_1)Y^{a_0,a_1}}{f(A_0 = a_0|L_0)f(A_1 = a_1|L_1, A_0 = a_0, L_0)}\right]$$

By consistency and presence of indicators

$I(A_0 = a_0)I(A_1 = a_1)Y^{a_0,a_1} = I(A_0 = a_0)I(A_1 = a_1)Y$ for all A_0, A_1

That is, consistency where $A_0 = a_0, A_1 = a_1$, otherwise both sides are zeros.

$$= E\left[\frac{I(A_0 = a_0)I(A_1 = a_1)Y}{f(A_0 = a_0|L_0)f(A_1 = a_1|L_1, A_0 = a_0, L_0)}\right]$$

$=$ IPTW mean of Y for group $A_0 = a_0, A_1 = a_1$

4 Bibliography Part

4.1 Bibliography

[Hernan and Robins, 2019] Hernan, M. A. and Robins, J. M. (2019). *Causal Inference*. Chapman & Hall/CRC.

[Robins, 1986] 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(9):1393–1512.

[Robins et al., 2000] Robins, J. M., Hernán, M. A., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–560.