Modern Statistical Learning Methods for Observational Biomedical Data

Lab 0: Identification of an average treatment effect

David Benkeser
Emory University

Marco Carone
Univ. of Washington

MODULE 8

Summer Institute in Statistics for Clinical and Epidemiological Research

July 2023

Contents of this Lab

The G-computation and IPTW identification formulas

For the population described in the table below, calculate $E(Y \mid A = a)$ and $E[Y(a)]$ corresponding to treatment $(a = 1)$ and control $(a = 0)$ using:

(i) the G-computation formula;

(ii) the IPTW formula.

(Green and orange values provide counts of individuals with $Y = 1$ and $Y = 0$, resp.)

	$W = 0$	$W = 1$	$W = 2$	
$A = 1$	$20 + 12$	$49 + 11$	$89 + 7$	
$A = 0$	$94 + 34$	$30 + 30$	$6 + 18$	

The G-computation and IPTW identification formulas

	$W = 0$	$W = 1$	$W = 2$	
$A = 1$	$20 + 12$	$49 + 11$	$89 + 7$	188
$A = 0$	$94 + 34$	$30 + 30$	$6 + 18$	212
	160	120	120	400

The G-computation and IPTW identification formulas

$$\bar{Q}_n(1,0) = \frac{12}{32} = 0.375 \quad \Big| \quad \bar{Q}_n(1,1) = \frac{11}{60} = 0.183 \quad \Big| \quad \bar{Q}_n(1,2) = \frac{7}{96} = 0.073$$

$$\bar{Q}_n(0,0) = \frac{34}{128} = 0.266 \quad \Big| \quad \bar{Q}_n(0,1) = \frac{30}{60} = 0.5 \quad \Big| \quad \bar{Q}_n(0,2) = \frac{18}{24} = 0.75$$

$$g_n(0) = \frac{32}{160} = 0.2 \quad \Big| \quad g_n(1) = \frac{60}{120} = 0.5 \quad \Big| \quad g_n(2) = \frac{96}{120} = 0.8$$

$$P_n(W=0) = \frac{160}{400} = 0.4 \quad \Big| \quad P_n(W=1) = \frac{120}{400} = 0.3 \quad \Big| \quad P_n(W=2) = \frac{120}{400} = 0.3$$

The G-computation and IPTW identification formulas

$$
\begin{aligned}
\gamma_{n,G} &= \left[\bar{Q}_n(1,0)P_n(W=0) + \bar{Q}_n(1,1)P_n(W=1) + \bar{Q}_n(1,2)P_n(W=2)\right] \\
&\qquad - \left[\bar{Q}_n(0,0)P_n(W=0) + \bar{Q}_n(0,1)P_n(W=1) + \bar{Q}_n(0,2)P_n(W=2)\right] \\
&= \left[\frac{12}{32}\cdot\frac{160}{400} + \frac{11}{60}\cdot\frac{120}{400} + \frac{7}{96}\cdot\frac{120}{400}\right] - \left[\frac{34}{128}\cdot\frac{160}{400} + \frac{30}{60}\cdot\frac{120}{400} + \frac{18}{24}\cdot\frac{120}{400}\right] \\
&= 0.227 - 0.481 = -0.254
\end{aligned}
$$

$$
\begin{aligned}
\gamma_{n,IPTW} &= \left[\frac{P_n(Y=1,A=1,W=0)}{g_n(0)} + \frac{P_n(Y=1,A=1,W=1)}{g_n(1)} + \frac{P_n(Y=1,A=1,W=2)}{g_n(2)}\right] \\
&\qquad - \left[\frac{P_n(Y=1,A=0,W=0)}{1-g_n(0)} + \frac{P_n(Y=1,A=0,W=1)}{1-g_n(1)} + \frac{P_n(Y=1,A=0,W=2)}{1-g_n(2)}\right] \\
&= \left[\frac{12}{400}\Big/\frac{32}{160} + \frac{11}{400}\Big/\frac{60}{120} + \frac{7}{400}\Big/\frac{96}{120}\right] \\
&\qquad - \left[\frac{34}{400}\Big/\left(1-\frac{32}{160}\right) + \frac{30}{400}\Big/\left(1-\frac{60}{120}\right) + \frac{18}{400}\Big/\left(1-\frac{96}{120}\right)\right] \\
&= 0.227 - 0.481 = -0.254
\end{aligned}
$$