

ON BAYESIAN G-COMPUTATION AND INSTRUMENTAL VARIABLES

P. GUSTAFSON AND H. CAMPBELL

1. INTRODUCTION

After a marked change in treatment strategy, there is an obvious desire to infer the benefit of this change, often in terms of longer-term outcomes than can be assessed in clinical trials. Thus there is a need to use observational data, but this is typically accompanied by concern about confounding by indication - if patients start on the new therapy because they aren't doing well, straight-up comparisons of ever versus never treated patients are not fair. One strategy is to apply modern causal inference tools that can appropriately deal with this. Another, seemingly cruder strategy, would involve treating calendar period of reaching eligibility for treatment as an instrumental variable. Each approach involves very strong assumptions, though interestingly they are different assumptions, one not stronger than the other. So we want to compare, at least in terms of power to detect treatment efficacy in the first instance.

Frequentist with logistic model: $Pr(A = 1|B = b, C = c) = predict(logistic(Pr(A = 1|B = b, C = c)))$

Bayesian with Uniform priors: $Pr(A = 1|B = b, C = c) \sim Beta(1 + \sum_{i=1}^n 1_{(A=1, B=b, C=c)}(x_i), 1 + \sum_{i=1}^n 1_{(A=0, B=b, C=c)}(x_i))$

Assumption	IV	g-comp
H not a confounder	✓	✗
C is the only confounder	✗	✓

Target value to estimate alongside associated p -value:

$$(1) \quad T_{(1,1)-(0,0)} = Pr(Y_2 = 1|do(X_1 = 1, X_2 = 1)) - Pr(Y_2 = 1|do(X_1 = 0, X_2 = 0))$$

With Y_1 , without H :

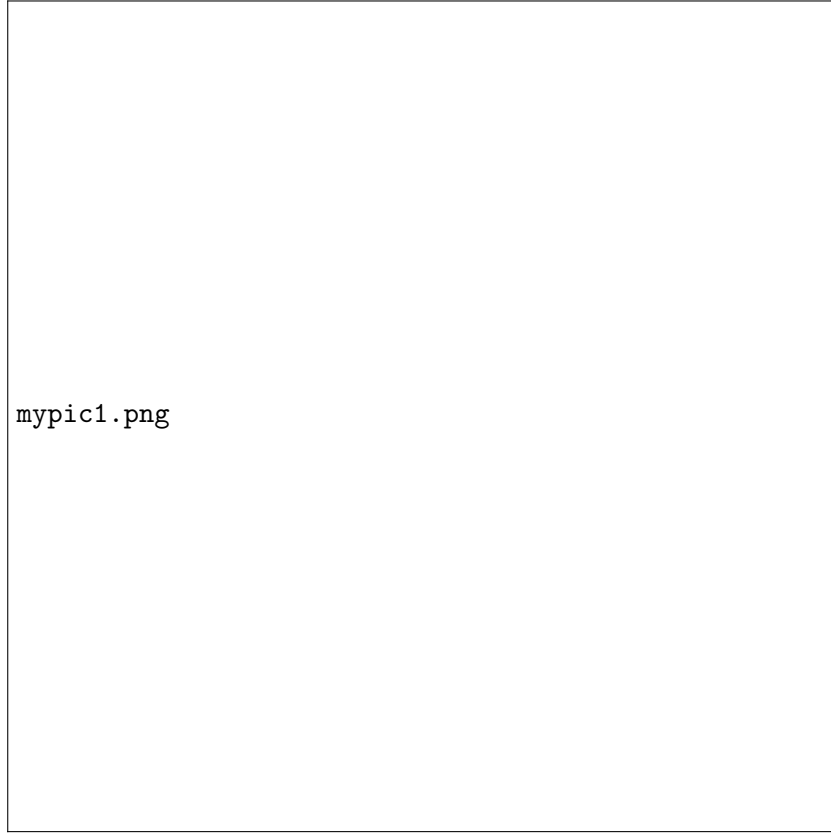


FIGURE 1. Okay

$$\begin{aligned}
 Pr(Y_2 = 1 | do(X_1 = x_1, X_2 = x_2)) &= \sum_{y_1=0}^1 \sum_{c_1=0}^1 \sum_{c_2=0}^1 Pr(C_1 = c_1) \\
 &\quad \cdot Pr(C_2 = c_2 | C_1 = c_1, X_1 = x_1) \\
 &\quad \cdot Pr(Y_1 = y_1 | C_1 = c_1, X_1 = x_1) \\
 &\quad \cdot Pr(Y_2 = 1 | C_2 = c_2, X_2 = x_2, Y_1 = y_1)
 \end{aligned}$$

With Y_1 , and with H :

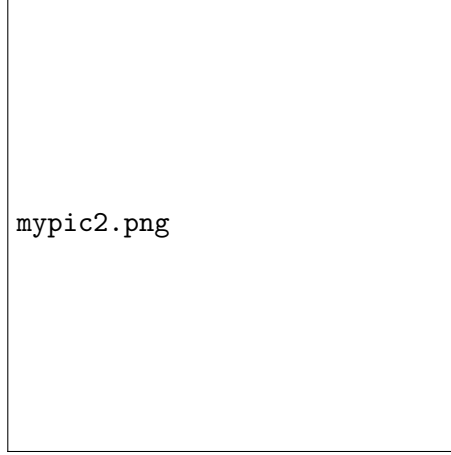


FIGURE 2. Okay

$$\begin{aligned}
 Pr(Y_2 = 1|do(X_1 = x_1, X_2 = x_2)) &= \sum_{y_1=0}^1 \sum_{c_1=0}^1 \sum_{c_2=0}^1 \sum_{h=0}^1 Pr(H = h) \cdot Pr(C_1 = c_1|H = h) \\
 &\quad \cdot Pr(C_2 = c_2|C_1 = c_1, X_1 = x_1, H = h) \\
 &\quad \cdot Pr(Y_1 = y_1|C_1 = c_1, X_1 = x_1, H = h) \\
 &\quad \cdot Pr(Y_2 = 1|C_2 = c_2, X_2 = x_2, Y_1 = y_1, H = h)
 \end{aligned}$$

Without Y_1 , without H :

$$\begin{aligned}
 Pr(Y_2 = 1|do(X_1 = x_1, X_2 = x_2)) &= \sum_{c_1=0}^1 \sum_{c_2=0}^1 Pr(C_1 = c_1) \\
 &\quad \cdot Pr(C_2 = c_2|C_1 = c_1, X_1 = x_1) \\
 &\quad \cdot Pr(Y_2 = 1|C_2 = c_2, X_2 = x_2)
 \end{aligned}$$

Without Y_1 , with H :

$$\begin{aligned}
 Pr(Y_2 = 1|do(X_1 = x_1, X_2 = x_2)) &= \sum_{c_1=0}^1 \sum_{c_2=0}^1 \sum_{h=0}^1 Pr(H = h) \cdot Pr(C_1 = c_1|H = h) \\
 &\quad \cdot Pr(C_2 = c_2|C_1 = c_1, X_1 = x_1, H = h) \\
 &\quad \cdot Pr(Y_2 = 1|C_2 = c_2, X_2 = x_2, H = h)
 \end{aligned}$$

2. METHODS

- run down of the simple sandbox of two time-points, everything binary - g-computation (perhaps motivated several different ways) - IV idea

- Comparison in a focussed scenario - confounding by indication present (C encourages treatment start) - different possible forms of treatment effect (through C and/or around C) - both sets of assumptions met

- Comparison across a broad range of scenarios

- set up distribution giving rise to a wide array of states of the world - look for patterns, e.g., is the power of the g-computation driven only by $\theta_{11}-\theta_{00}$, or are there other important factors

- Violations of assumptions

- any defensible sense of one procedure being more robust than the other - how quickly is power lost as we relax assumptions

3. APPLICATION

I don't know whether we'd find a good 'real-data' example, but we could address the question of how much data we would need (how many subjects, over how long a time period) in order to say something useful about the efficacy of beta-interferon for MS in terms of long-term disability.

Random citation [1] embedded in text.

4. DISCUSSION

REFERENCES

- [1] O. Saarela, D. A. Stephens, E. E. Moodie, and M. B. Klein, “On bayesian estimation of marginal structural models,” *Biometrics*, 2015.
- [2] L. Borsi, *Estimating the Causal Effect of Switching to Second-line Antiretroviral HIV Treatment Using G-computation*. PhD thesis, ETH, Department of Mathematics, 2012.
- [3] M. Kuroki and Z. Cai, “Instrumental variable tests for directed acyclic graph models,”
- [4] S. Greenland, “An introduction to instrumental variables for epidemiologists,” *International journal of epidemiology*, vol. 29, no. 4, pp. 722–729, 2000.
- [5] K. Peter, *Marginal structural models and causal inference*. PhD thesis, Master thesis, ETH, 2011.
- [6] B. Chakraborty and E. E. Moodie, “G-computation: Parametric estimation of optimal dtrs,” in *Statistical Methods for Dynamic Treatment Regimes*, pp. 101–112, Springer, 2013.
- [7] S. L. Taubman, J. M. Robins, M. A. Mittleman, and M. A. Hernán, “Intervening on risk factors for coronary heart disease: an application of the parametric g-formula,” *International Journal of Epidemiology*, vol. 38, no. 6, pp. 1599–1611, 2009.
- [8] J. M. Snowden, S. Rose, and K. M. Mortimer, “Implementation of g-computation on a simulated data set: demonstration of a causal inference technique,” *American Journal of Epidemiology*, p. kwq472, 2011.