Improving double protected estimation in causal inference models with longitudinal data.

Julieta Molina.
Phd student, Instituto de Cálculo, UBA - CONICET.

Topics

- Double protected estimation.
- Longitudinal data - Causal Analysis.
- Marginal structural mean model.

Double protected estimation in missing data.

Robins JM, Rotnitzky A, Scharfstein D. (1999)

Missing data:

- $Y_i \in \mathbb{R}$ scalar outcome of interest.
- $A_i \in \{0, 1\}$, $A_i = 1$ if Y_i is observed and $A_i = 0$ otherwise.
- $L_i \in \mathbb{R}^m$ vector of additional variables.
- Observed data $O_1, \cdots, O_n$ i.i.d., where

$$O_i = \begin{cases} L_i, Y_i, A_i & \text{if } A_i = 1 \\ L_i, A_i & \text{if } A_i = 0. \end{cases}$$

Aim: estimation of the unknown parameter $\beta_0 \in \mathbb{R}$

$$E[Y] = \beta_0.$$

Assumptions: Missing at Random (MAR)

Y is independent of A, given L.

Identificability :

- MAR guarantees that

$$\beta_0 = E[Y] = E\left[\frac{A}{P(A=1|L)}Y\right].$$

$\widehat{\beta}_{\text{ipw}}$ estimator.

Missingness model: (MM)

$$P(A = 1|L) = f(L\,, \gamma_0). \tag{1}$$

We have that under MM:

$$E\left[\frac{A}{f(L\,, \gamma_0)}(Y - \beta_0)\right] = 0. \tag{2}$$

Inverse probability weighted estimator (ipw): $\widehat{\beta}_{\text{ipw}}$ solves

$$E_n\left[\frac{A}{f(L\,, \widehat{\gamma})}(Y - \beta)\right] = 0, \tag{3}$$

where $\widehat{\gamma}$ is the maximum likelihood estimator for γ_0 under MM.

$\widehat{\beta}_{\tau,\,\mathrm{ipw}}$ estimator.

For all measurable functions τ we have:

$$E\left[\tau(A,L) - E_{\gamma_0}[\tau(A,L)|L]\right] = 0. \tag{4}$$

Augmented inverse probability weighted estimator: $\widehat{\beta}_{\tau,\,\mathrm{ipw}}$ solves

$$E_n\left[\frac{A}{f(L\,,\widehat{\gamma})}(Y-\beta)\right] - E_n\left[\tau(A,L) - E_{\widehat{\gamma}}[\tau(A,L)|L]\right] = 0. \tag{5}$$

Local efficiency:

$$\tau_0(A, L) = \frac{A}{f(L, \gamma_0)} \{E[Y|A = 1, L] - \beta_0\},$$

is the function that yields the estimator $\widehat{\beta}_{\tau,\text{ipw}}$ with smallest asymptotic variance.

$$E_n[\frac{A}{f(L\,,\widehat{\gamma})}(Y-\beta)] - E_n\left[\tau(A,L) - E_{\widehat{\gamma}}[\tau(A,L)|L]\right] = 0, \qquad (6)$$

$$\tau_0(A,L) = \frac{A}{f(L\,,\gamma_0)}\{E[Y|A=1,L] - \beta_0\}, \qquad (7)$$

$$E[Y|L,A=1] = m(L,\eta_0), \quad P(A=1|L) = f(L\,\gamma_0), \qquad (8)$$

$$\tau_0(A,L) - E_{\gamma_0}[\tau_0(A,L)|L] = (\frac{A}{f(L\,,\gamma_0)} - 1)\{m(L,\eta_0) - \beta_0\}, \qquad (9)$$

$$\tilde{\tau}_0(A,L) - E_{\widehat{\gamma}}[\tilde{\tau}_0(A,L)|L] = (\frac{A}{f(L\,,\widehat{\gamma})} - 1)\{m(L,\widehat{\eta}) - \beta_0\}. \qquad (10)$$

Double protected estimator: $\widehat{\beta}_{dp}(\widehat{\eta}, \widehat{\gamma})$ solves

$$E_n \left[\frac{A}{f(L\,, \widehat{\gamma})}(Y - \beta) + (1 - \frac{A}{f(L\,, \widehat{\gamma})})\{m(L, \widehat{\eta}) - \beta\} \right] = 0, \qquad (11)$$

where

- $\widehat{\eta}$ is an estimator of η_0 under OR model:

$$E[Y|L, A = 1] = m(L, \eta_0). \qquad (12)$$

- $\widehat{\gamma}$ is the ml estimator of γ_0 under M model:

$$P(A = 1|L) = f(L\,\gamma_0). \qquad (13)$$

Theorem

- if $P(A=1|L)=f(L\,,\gamma_0)$ and $E[Y|l,A=1]=m(L,\eta_0)$ then, $\widehat{\beta}_{dp}(\widehat{\eta},\widehat{\gamma})$ is is consistent and asymptotically normal (can) and has asymptotic variance equal to the smallest asymptotic variance of all estimators $\widehat{\beta}_{\tau,ipw}$.

- if $P(A=1|L)=f(L\,,\gamma_0)$ or $E[Y|L,A=1]=m(L,\eta_0)$ then, $\widehat{\beta}_{dp}(\widehat{\eta},\widehat{\gamma})$ is can for β_0 (**doble protected estimator**).

Improving $\widehat{\beta}_{dp}(\widehat{\eta}, \widehat{\gamma})$

Rotnitzky, Lei, Sued and Robins (2009), TAN, Z. (2008), Cao, W., TSIATIS, A. & DAVIDIAN, M. (2009), constructed estimators $\widehat{\beta}_{\text{idp}}$ with the following properties:

- They are double protected for β_0 for OR and M models.
- If both models are correct, they have asymptotic variance equal to the smallest asymptotic variance of all estimators $\widehat{\beta}_{\tau\,\text{ipw}}$.
- They lie in the range of β_0.

- Under model M, they are more efficient than $\widehat{\beta}_{\text{ipw}}$.

Longitudinal Data : for $1 \leq i \leq n$

$$L_{0i}\,,\,A_{0i}\,,\,L_{1i}\,,\ldots L_{ki}\;,\ldots\;,\,L_{si}\,,\,A_{si}\,,\,L_{s+1_i}, \quad \text{iid},$$

where

- $L_{ki} \in \mathbb{R}^{m_k} \quad 0 \leq k \leq s$ pre-treatment variables.
- $A_{ki} \quad 0 \leq k \leq s$ treatment variables.
- $Y_i = \varphi(L_{0i}, L_{1i}, \ldots L_{s+1_i})$ outcome of interest.

Causal Analysis.

Regime g

- $g = (g_1, g_2 \ldots g_s)$.
- $g_k(\ell_0, \ell_1, \cdots, \ell_k) =$ treatment at time k.
- $G = \{g_x : x \in \mathcal{X}\}$.

Contrafactual variables:

We have $(L_0,\ A_0,\ L_1 \ldots\ A_s\, L_{s+1})$ (Factual variables).

Contrafactual variables: given a regime g

$L_{k,g}$ = answer at time k which would be observed in the world where everybody is forced to follow the regime g.

Consistency assumption:

$$A_k = g_k(L_0, L_1, \ldots L_k) \quad \text{then} \quad L_{k+1}, g = L_{k+1}. \tag{14}$$

In causal Analysis we try to identify and estimate, from factual variables and longitudinal data respectively, causal parameters like $E[Y_g]$.

Relation with missing data and example:

- $s = 0$
- L_0, A_0, Y where $A_0 = 1$(take the drug) or $A_0 = 0$ (don't take the drug)
- $g = 1$, ie: $g =$every body takes the drug, so $g(L_0) \equiv 1$ (static regimen.)
- $E[Y_g]$.
- if $A_0 = 1 = g(L_0)$ then $Y_g = Y$.

So, if we assume NUC (no unmeasure confounders - equivalent to MAR), this is a missing at random problem.

Marginal structural mean model

- $L_{0i}\ A_{0i}\ L_{1i}\ \ldots L_{ki}\ \ldots L_{si}\ A_{si}\ L_{s+1_i}$, iid.

- $G = \{g_x : x \in \mathcal{X}\}$.

$$E\left[Y_{g_x}|Z\right] = h\left(x, Z; \beta_0\right) \quad x \in \mathcal{X}, \tag{15}$$

- $\beta_0 \in R^{p\times 1}$ is unknown,Z subset of components of L_0, $h\left(\cdot,\cdot,\cdot\right)$ known smooth function.

- Murphy SA, van der Laan MJ, Robins JM, CPPRG (2001) proposed estimators for β_0 in marginal structural mean models with the same properties that $\widehat{\beta}_{dp}(\widehat{\eta}, \widehat{\gamma})$.

- Lately M.Sued, A. Rotnitzky and myself have improved those estimators following the same spirit as the one used to construct $\widehat{\beta}_{idp}(\widehat{\eta}, \widehat{\gamma})$.

References

- Murphy SA, van der Laan MJ, Robins JM, CPPRG (2001). Marginal mean models for dynamic regimes. Journal of the American Statistical Associataion. 96(456):1410-1423.

- Robins JM, Rotnitzky A, Scharfstein D. (1999). Sensitivity Analysis for Selection Bias and Unmeasured Confounding in Missing Data and Causal Inference Models. In: Statistical Models in Epidemiology: The Environment and Clinical Trials. Halloran, M.E. and Berry, D., eds. IMA Volume 116, NY: Springer-Verlag, pp. 1-92.

- Tan, Z. (2008). Comment: Improved Local Efficiency and Double Robustness. *Interntl J. of Biostatist.* 4, article 10.

- Cao, W., Tsiatis, A. & Davidian, M. (2009). Improving efficiency and robustness of the doubly robust estimator for a population mean with incomplete data. *Biometrika* **96** 723-34.