Supplemental Estimating Equations

Paul Zivich

June 5, 2024

This supplemental document provides the estimating functions for a few different extensions of the examples reviewed. Note that the notation is assumed to carry over from the workshop slides unless noted otherwise. To cite this document, please use the following DOI from Zenodo: https://doi.org/10.5281/zenodo.8017957

Let O_i denote the observed data for unit i. An M-estimator, $\hat{\theta}$ is the solution to the estimating equation $\sum_{i=1}^{n} \psi(O_i; \theta) = 0$, where $\psi(O_i; \theta)$ is the estimating function.

Average Treatment Effect in the Treated – G-computation

The following is an estimating function for the ATT with g-computation for the presented ZAPPS example, where $\theta = (\boldsymbol{\beta}, \mu_1, \mu_2.\delta_1, \delta_2)$.

$$\psi(O_i; \theta) = \begin{Bmatrix} Y_i - \text{expit}(\mathbf{Z}_i \hat{\boldsymbol{\beta}}^T) \\ \{Y_i - \text{expit}(\mathbf{Z}_i \hat{\boldsymbol{\beta}}^T)\} X_i \\ \{Y_i - \text{expit}(\mathbf{Z}_i \hat{\boldsymbol{\beta}}^T)\} W_i \\ X_i \{\text{expit}(\hat{\beta}_0 + \hat{\beta}_1 \times 1 + \hat{\beta}_2 W_i) - \hat{\mu}_1\} \\ X_i \{\text{expit}(\hat{\beta}_0 + \hat{\beta}_1 \times 0 + \hat{\beta}_2 W_i) - \hat{\mu}_2\} \\ (\hat{\mu}_1 - \hat{\mu}_2) - \hat{\delta}_1 \\ \ln(\hat{\mu}_1 / \hat{\mu}_2) - \hat{\delta}_1 \end{Bmatrix}$$

Note that the key distinction from the presented g-computation estimating equations presented in the workshop is that only those with $X_i = 1$ contribute non-zero values to the fourth and fifth estimating functions. In other words, the mean of the predicted values is only taken among those with $X_i = 1$.

Average Treatment Effect in the Treated – IPW

The following is an estimating function for the ATT with IPW for the presented ZAPPS example, where $\theta = (\boldsymbol{\alpha}, \mu_1, \mu_2.\delta_1, \delta_2)$.

$$\psi(O_i; \theta) = \begin{Bmatrix} X_i - \text{expit}(\hat{\alpha}_0 + \hat{\alpha}_1 W_i) \\ \{X_i - \text{expit}(\hat{\alpha}_0 + \hat{\alpha}_1 W_i)\} W_i \\ X_i \{Y_i - \mu_1\} \\ Y_i (1 - X_i) \times \frac{\text{expit}(\hat{\alpha}_0 + \hat{\alpha}_1 W_i)}{1 - \text{expit}(\hat{\alpha}_0 + \hat{\alpha}_1 W_i)} - \hat{\mu}_2 \\ (\hat{\mu}_1 - \hat{\mu}_2) - \hat{\delta}_1 \\ \ln(\hat{\mu}_1 / \hat{\mu}_2) - \hat{\delta}_1 \end{Bmatrix}$$

Note that the key distinction from the presented IPW estimating equations presented in the workshop is that the fourth is simply the mean among those with X_i and the fifth uses the inverse odds of treatment weights.

Trimming Propensity Scores

The following is an estimating function for the ATE with IPW and trimmed propensity scores for the presented ZAPPS example. Recall that trimming the propensity scores means we are drawing inference for only those between the investigator-chosen thresholds.

To make our expression a bit more compact, let $\pi(W_i; \hat{\alpha}) = \text{expit}(\hat{\alpha}_0 + \hat{\alpha}_1 W_i)$, t_l indicate the lower threshold for trimming and t_u indicate the upper threshold for trimming.

$$\psi(O_i;\theta) = \begin{Bmatrix} X_i - \pi(W_i;\hat{\alpha}) \\ \{X_i - \pi(W_i;\hat{\alpha})\}W_i \\ I(t_l \leq \pi(W_i;\hat{\alpha}) \leq t_u) \times \left\{\frac{Y_i X_i}{\pi(W_i;\hat{\alpha})} - \hat{\mu}_1\right\} \\ I(t_l \leq \pi(W_i;\hat{\alpha}) \leq t_u) \times \left\{\frac{Y_i(1-X_i)}{1-\pi(W_i;\hat{\alpha})} - \hat{\mu}_2\right\} \\ (\hat{\mu}_1 - \hat{\mu}_2) - \hat{\delta}_1 \\ \ln(\hat{\mu}_1/\hat{\mu}_2) - \hat{\delta}_1 \end{Bmatrix}$$

Note that it can be difficult to simultaneously find the roots of α and μ. To simplify the procedure in practice, it will often be easier to separate find $\hat{\alpha}$ separately and use those as the starting values for the root-finding procedure of the corresponding estimating equations for the above estimating function.

Truncated Propensity Scores

The following is an estimating function for the ATE with IPW and truncated propensity scores for the presented ZAPPS example. Recall that truncating the propensity scores means we are clipping the estimated propensity scores to be within a certain range.

Again, let $\pi(W_i;\hat{\alpha}) = \text{expit}(\hat{\alpha}_0 + \hat{\alpha}_1 W_i)$. Further, let

$$\pi^*(W_i;\alpha) = \begin{cases} s_l & \text{if } \pi(W_i;\alpha) < s_l \\ \pi(W_i;\alpha) & \text{if } s_l \leq \pi(W_i;\alpha) \leq s_u \\ s_u & \text{if } \pi(W_i;\alpha) > s_l \end{cases}$$

where s_l indicates the lower threshold for truncating and s_u indicates the upper threshold for truncating.

The estimating functions are

$$\psi(O_i;\theta) = \begin{Bmatrix} X_i - \pi(W_i;\hat{\alpha}) \\ \{X_i - \pi(W_i;\hat{\alpha})\}W_i \\ \frac{Y_i X_i}{\pi^*(W_i;\hat{\alpha})} - \hat{\mu}_1 \\ \frac{Y_i(1-X_i)}{1-\pi^*(W_i;\hat{\alpha})} - \hat{\mu}_2 \\ (\hat{\mu}_1 - \hat{\mu}_2) - \hat{\delta}_1 \\ \ln(\hat{\mu}_1/\hat{\mu}_2) - \hat{\delta}_1 \end{Bmatrix}$$

Again, it can be difficult to simultaneously find the roots of α and μ. To simplify the procedure in practice, it will often be easier to separate find $\hat{\alpha}$ separately and use those as the starting values for the root-finding procedure of the corresponding estimating equations for the above estimating function.

Others

A variety of estimating equation formulas are also actively maintained in the `delicatessen` documentation, available here `https://deli.readthedocs.io/en/latest/Reference/Estimating%20Equations.html`