Estimators of the average treatment effect

Here, we document all candidate estimators we will use in building synthetic causal estimators. For the most part, we write the estimators in the form

$$\sum_{i=1}^{n} \omega_i y_i,$$

where ω_i is a weight function that can in general depend on all the data. Let d_i be the treatment indicator, and let $\mathbf{x}_i$ denote covariates. We will denote the propensity score as e_i and the prognostic score as g_i.

- IPW:
$$\omega_i = \frac{d_i}{e_i \sum_{i=1}^{n} \frac{d_i}{e_i}} - \frac{1-d_i}{(1-e_i)\sum_{i=1}^{n} \frac{1-d_i}{1-e_i}}$$
- Regression
$$\omega_i = \left\{(X'X)^{-1}X'\right\}_{ji}$$
where this refers to the appropriate entry of the hat matrix and X is the design matrix including covariates and treatment.
- Regression within K strata
$$\omega_i = \sum_{k=1}^{K} K^{-1}\frac{n_k}{n} I\{e_i \in E_k\}\left\{(X(k)'X(k))^{-1}X(k)'\right\}_{ji}$$
where $\{E_k\}_{k=1,\ldots,K}$ is a partition of (0,1) and $X(k)$ is the design matrix including only those individuals for whom $e_i \in E_k$.
- Doubly robust
$$\omega_i = n^{-1}\left\{\frac{d_i}{e_i} - \frac{1-d_i}{1-e_i} - \sum_{j=1}^{n}\left(\frac{d_j - e_j}{e_j}h_{j1i} + \frac{d_j - e_j}{1-e_j}h_{j0i}\right)\right\}$$
where h_{jdi} corresponds to the ith entry in the matrix $X_{jd}(X'X)^{-1}X'$ and X_{jd} corresponds to the jth row of X with treatment set to value d.
- Stratified
$$\omega_i = K^{-1}\sum_{k=1}^{K}\frac{n_k}{n} I\{e_i \in E_k\}\left\{\frac{d_i}{\sum_{i=1}^{n} d_i I\{e_i \in E_k\}} - \frac{1-d_i}{\sum_{i=1}^{n}(1-d_i) I\{e_i \in E_k\}}\right\}$$
where $n_k = \sum_{i=1}^{n} I\{e_i \in E_k\}$.
- Matching ($M = 5$) with replacement on propensity
$$\omega_i = n^{-1}(2d_i - 1)\left(1 + M^{-1}\sum_{j:d_j \neq d_i} I\left\{\sum_{k=1}^{n} I\left\{|e_i - e_j| > |e_k - e_j|\right\}\right) < M\right\}$$
- Matching ($M = 5$) with replacement on prognostic
$$\omega_i = n^{-1}(2d_i - 1)\left(1 + M^{-1}\sum_{j:d_j \neq d_i} I\left\{\sum_{k=1}^{n} I\left\{|g_i - g_j| > |g_k - g_j|\right\} < M\right\}\right)$$

- Matching ($M=5$) with replacement on both scores

$$\omega_i = n^{-1}(2d_i - 1)\left(1 + M^{-1}\sum_{j:d_j \neq d_i} I\left\{\sum_{k=1}^{n} I\left\{\|z_i - z_j\| > \|z_k - z_j\|\right\} < M\right\}\right)$$

  where z_i is the vector of both scores.
- Caliper ($\delta = 0.05$) matching ($M=5$) with replacement on propensity score

$$\omega_i = n^{-1}(2d_i - 1) I\{|e_i - e_j| < \delta \quad \exists j\}\left(1 + M^{-1}\sum_{j:d_j \neq d_i} I\{|e_i - e_j| < \delta\} I\left\{\sum_{k=1}^{n} I\left\{|e_i - e_j| > |e_k - e_j|\right\} < M\right\}\right)$$

- Caliper ($\delta = 0.05$) matching ($M=5$) with replacement on prognostic score

$$\omega_i = n^{-1}(2d_i - 1) I\{|g_i - g_j| < \delta \quad \exists j\}\left(1 + M^{-1}\sum_{j:d_j \neq d_i} I\{|g_i - g_j| < \delta\} I\left\{\sum_{k=1}^{n} I\left\{|g_i - g_j| > |g_k - g_j|\right\} < M\right\}\right)$$

- Caliper ($\delta = 0.2$) matching ($M=5$) with replacement on both scores

$$\omega_i = n^{-1}(2d_i - 1) I\{\|z_i - z_j\| < \delta \quad \exists j\}\left(1 + M^{-1}\sum_{j:d_j \neq d_i} I\{\|z_i - z_j\| < \delta\} I\left\{\sum_{k=1}^{n} I\left\{\|z_i - z_j\| > \|z_k - z_j\|\right\} < M\right\}\right)$$

- Balancing

$$\omega_i = d_i p_K(\mathbf{x}_i) - (1 - d_i) q_K(x_i),$$

  where $p_K(\mathbf{x}_i)$ and $q_K(\mathbf{x}_i)$ are weights that ensure that $\sum_{i=1}^{n} d_i p_K(\mathbf{x}_i) u_K(\mathbf{x}_i) = \sum_{i=1}^{n}(1 - d_i) p_K(\mathbf{x}_i) u_K(\mathbf{x}_i) = n^{-1}\sum_{i=1}^{n} u_K(\mathbf{x}_i)$ for some contrasts encoded in $u_K(\mathbf{x}_i)$.
- Approximate residual balancing

$$\bar{\mathbf{x}}^{\intercal}(\widehat{\boldsymbol{\beta}}_t - \widehat{\boldsymbol{\beta}}_c) + \sum_{i=1}^{n} d_i \gamma_{t,i}(y_i - \mathbf{x}_i^{\intercal}\widehat{\boldsymbol{\beta}}_t) - \sum_{i=1}^{n}(1 - d_i)\gamma_{c,i}(y_i - \mathbf{x}_i^{\intercal}\widehat{\boldsymbol{\beta}}_c),$$

  where

$$\boldsymbol{\gamma}_t = \underset{\boldsymbol{\gamma}}{\operatorname{argmin}}\left\{(1 - \zeta)\|\boldsymbol{\gamma}\|_2^2 + \zeta\|\bar{\mathbf{x}} - \mathbf{x}_t^{\intercal}\boldsymbol{\gamma}\|_\infty^2\right\}$$

  subject to $\sum_{i:d_i=1}\gamma_i = 1$ and $0 \leq \gamma_i \leq n_t^{-2/3}$ and $\boldsymbol{\gamma}_c$ is defined similarly.