Balancing Covariates via Propensity Score Weighting

Kari Lock Morgan

Department of Statistical Science
Duke University
kari@stat.duke.edu

March 17, 2014

Joint work with Fan Li and Alan Zaslavsky
Outline

1. Introduce a general framework for a class of balancing weights that unifies most of the existing weights
2. Propose overlap weight and show some optimality properties
3. Illustrate with examples
Data: random sample of \( n \) units from a population.

Treatment status \( Z_i (= 0, 1) \) and covariates \( X_i \) are observed.

Potential outcomes are \( Y_i(0) \), \( Y_i(1) \), but only \( Y_i(Z_i) \) is observed.

Propensity score is \( e(x) = \Pr(Z_i = 1|X_i = x) \).
Population density of the covariates $X$ is $f(x)$

Density for group $Z = z$ is $f_z(x) = \Pr(X = x | Z = z)$

Then $f_1(x) \propto f(x)e(x)$ and $f_0(x) \propto f(x)(1 - e(x))$

Choose weights such that $f_1(x)w_1(x) \propto f_0(x)w_0(x)$
Balancing weights

- We propose the following class of balancing weights:

\[
\begin{align*}
    w_1(x) & \propto \frac{h(x)}{e(x)}, \\
    w_0(x) & \propto \frac{h(x)}{1-e(x)},
\end{align*}
\]

where \(h(\cdot)\) is a pre-specified function.

- The weighted covariate distributions in the two treatment groups have the same target density \(\propto f(x)h(x)\):

\[
\begin{align*}
    f_1(x)w_1(x) & \propto f(x)e(x)\frac{h(x)}{e(x)} = f(x)h(x), \\
    f_0(x)w_0(x) & \propto f(x)(1-e(x))\frac{h(x)}{1-e(x)} = f(x)h(x).
\end{align*}
\]
### Examples of target population and balancing weights

<table>
<thead>
<tr>
<th>target population</th>
<th>$h(x)$</th>
<th>estimand</th>
<th>weight $(w_1, w_0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>combined</td>
<td>1</td>
<td>ATE</td>
<td>$(\frac{1}{e(x)}, \frac{1}{1-e(x)})$ [HT]</td>
</tr>
<tr>
<td>treated</td>
<td>$e(x)$</td>
<td>ATT</td>
<td>$(1, \frac{e(x)}{1-e(x)})$</td>
</tr>
<tr>
<td>control</td>
<td>$1 - e(x)$</td>
<td>ATC</td>
<td>$(\frac{1-e(x)}{e(x)}, 1)$</td>
</tr>
<tr>
<td>truncated</td>
<td>$1(\alpha &lt; e(x) &lt; 1 - \alpha)$</td>
<td>ATTrunc</td>
<td>$(\frac{1(\alpha &lt; e(x) &lt; 1 - \alpha)}{e(x)}, \frac{1(\alpha &lt; e(x) &lt; 1 - \alpha)}{1-e(x)})$</td>
</tr>
<tr>
<td>combined</td>
<td>$e(x)(1 - e(x))$</td>
<td>ATO</td>
<td>$(1 - e(x), e(x))$</td>
</tr>
</tbody>
</table>
Estimands and Estimators

- Conditional average treatment effect (ATE)

\[
\tau(x) \equiv \mathbb{E}(Y(1)|X = x) - \mathbb{E}(Y(0)|X = x).
\]

- Estimand is average (ATE) over a target population with density \( \propto f(x)h(x) \):

\[
\tau_h \equiv \frac{\int \tau(dx)f(x)h(x)\mu(dx)}{\int f(x)h(x)\mu(dx)}.
\]

- \( \tau_h \) can be estimated by weighted averages:

\[
\hat{\tau}_h^{w} = \frac{\sum_{i:Z_i=1} Y_i w_1(x_i)}{\sum_{i:Z_i=1} w_1(x_i)} - \frac{\sum_{i:Z_i=0} Y_i w_0(x_i)}{\sum_{i:Z_i=0} w_0(x_i)}.
\]
Two large-sample results on $\hat{\tau}_h$

**Result 1.** *Given the normalizing constraint* 
\[
\int f(x)h(x)\mu(dx) = 1, \text{ the large-sample variance of the estimator } \hat{\tau}_h \text{ is:}
\]
\[
\nabla[\hat{\tau}_h] = \int f(x)h(x)^2 \left[ \frac{v_1(x)}{e(x)} + \frac{v_0(x)}{1 - e(x)} \right] \mu(dx)/N,
\]

*where $v_z(x)$ is the variance of $Y$ in a neighborhood $dx$ of $x$ in the $Z = z$ group.*

**Result 2.** *Assuming $v_0(x) \equiv v_1(x) \equiv v$, the function $h(x) = e(x)(1 - e(x))$ gives the smallest asymptotic variance for the weighted estimator $\hat{\tau}_h$, and*

\[
\min\{\nabla[\hat{\tau}_h]\} = \frac{v}{N} \int f(x)e(x)(1 - e(x))\mu(dx).
\]
We propose a new weight by letting \( h(x) = e(x)(1 - e(x)) \), leading to the overlap weights:

\[
\begin{align*}
    w_1(x) &\propto 1 - e(x), \\
    w_0(x) &\propto e(x).
\end{align*}
\]

Target population \( f(x)e(x)(1 - e(x)) \) defined by overlap of covariates

“Marginal” units who may or may not receive the treatment (Rosenbaum, 2012).
Figure: Densities for the treatment group, $f_1(x)$, control group, $f_0(x)$, and overlap population, $f(x)h(x)$. 
Result 3. When the propensity scores are estimated from a logistic regression model with main effects, $\text{logit}\{e(x_i)\} = \beta_0 + \beta'x_i$, the overlap weights lead to exact balance in any included covariate between treatment and control groups. That is,

$$\frac{\sum_i x_{i,k}Z_i(1 - \hat{e}_i)}{\sum_i Z_i(1 - \hat{e}_i)} = \frac{\sum_i x_{i,k}(1 - Z_i)\hat{e}_i}{\sum_i(1 - Z_i)\hat{e}_i}.$$
Advantages of the overlap weight

**Statistical advantages**
- Minimizes asymptotic variance of the weighted average estimator among all balancing weights.
- Perfect (exact small-sample) balance for means of included covariates in logistic propensity score model.
- Weights are bounded (unlike HT, etc.).
- Avoids artificially truncating weights or eliminating cases.

**Scientific advantages**
- Clinical equipoise.
- The “marginal units” are likely the group who are responsive to policy intervention.
Simulated Example

- Simulate $n_0 = n_1 = 1000$ units.
- A single covariate: $X_i \sim N(0, 1) + 2Z_i$.

Figure: Original covariate distributions within each treatment group, and weighted covariate distributions with overlap, HT, ATT weights.

<table>
<thead>
<tr>
<th></th>
<th>Unweighted</th>
<th>Overlap</th>
<th>HT</th>
<th>ATT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\overline{X}_1$</td>
<td>1.98</td>
<td>1.01</td>
<td>0.74</td>
<td>1.98</td>
</tr>
<tr>
<td>$\overline{X}_0$</td>
<td>0.03</td>
<td>1.01</td>
<td>1.19</td>
<td>2.22</td>
</tr>
</tbody>
</table>
Simulated Example

- A single covariate: $X_i \sim N(0, 1) + 2Z_i$.
- Outcome model with additive treatment effect:
  $Y_i \sim N(X_i, 1) + \tau Z_i$, and $\tau = 1$.
- Use the nonparametric estimator $\hat{\tau}_h^W$ with different weights:

<table>
<thead>
<tr>
<th></th>
<th>Unweighted</th>
<th>Overlap</th>
<th>HT</th>
<th>ATT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\tau}$</td>
<td>2.945</td>
<td>1.000</td>
<td>0.581</td>
<td>0.640</td>
</tr>
<tr>
<td>$SE(\hat{\tau})$</td>
<td>0.054</td>
<td>0.038</td>
<td>0.386</td>
<td>0.402</td>
</tr>
</tbody>
</table>
Goal: estimate racial disparity in medical expenditures after balancing covariates (Le Cook et al., 2010)

Race is not manipulable so comparisons are descriptive, not causal

Data: 2009 Medical Expenditure Panel Survey: 9830 non-Hispanic Whites, 4020 Blacks, 1446 Asians, 5280 Hispanics

Three independent comparisons; comparing non-Hispanic Whites to each minority group

Logistic regression to estimate propensity scores, 31 covariates (5 continuous, 26 binary)

Ignore survey weights here, but weighting allows easy incorporation of survey weights
Figure: Covariate balance (absolute standardized bias) with no weights, overlap weights, and HT weights.
Racial Disparity in Medical Expenditure

- Estimated Propensity Score
- Overlap
- HT

- White
- Black

- White
- Asian

- White
- Hispanic
One Asian woman has over 30% of the weight! (out of 1446 Asians)

78 year old Asian lady with a BMI of 55.4: $e(x) = 0.9998$

Common practice:
- Eliminate cases with propensity scores close to 0 or 1
- Truncate propensity scores or weights
- Can lead to ad hoc changes to target population
- Results can be very sensitive to truncation choice

The overlap weights avoid these extreme weights and avoid an abrupt threshold for elimination or truncation
### Table: Unweighted, overlap weights, and HT weighting estimates (SE) for difference in yearly medical expenditure.

<table>
<thead>
<tr>
<th></th>
<th>Unweighted</th>
<th>Overlap</th>
<th>HT</th>
</tr>
</thead>
<tbody>
<tr>
<td>White - Black</td>
<td>$786 (222)</td>
<td>$824 (185)</td>
<td>$856 (200)</td>
</tr>
<tr>
<td>White - Asian</td>
<td>$2764 (209)</td>
<td>$1227 (205)</td>
<td>$2167 (640)</td>
</tr>
<tr>
<td>White - Hispanic</td>
<td>$2599 (174)</td>
<td>$1212 (171)</td>
<td>$596 (323)</td>
</tr>
</tbody>
</table>
Right heart catheterization (RHC)

- Diagnostic to measure cardiac function.
- Observational data (Murphy and Cluff 1990).
  - \( n = 5735: \) 2184 RHC \( (Z_i = 1) \), 3551 control \( (Z_i = 0) \).
  - Outcome: survival at 30 days after admission.
  - Covariates: 53 binary/categorical variables.
- Extensively studied in literature.
  - Most focused on ATT: e.g. Connors et al. (1996), Hirano and Imbens (2001).
  - Crump et al. (2009): ATE for a truncated population with good overlap.
Right heart catheterization (RHC)

Table: Estimated treatment effect (in %) with different weights

<table>
<thead>
<tr>
<th></th>
<th>unweighted</th>
<th>overlap</th>
<th>HT</th>
<th>ATT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\tau}_h$</td>
<td>7.36</td>
<td>6.54</td>
<td>5.93</td>
<td>5.81</td>
</tr>
<tr>
<td>SE($\hat{\tau}_h$)</td>
<td>1.27</td>
<td>1.32</td>
<td>2.46</td>
<td>2.67</td>
</tr>
</tbody>
</table>
Unified framework for use of weighting to balance covariates for any target population.

The general class of balancing weights balance covariates and include many of the existing weights.

We propose a new type of weights, the overlap weights

- Target population emphasizes overlap in covariates.
- Minimizes asymptotic variance of weighted estimators.
- Exact balance for covariate means using logistic propensity score estimates.
- Simulated and real examples display the balancing property of the overlap weights.