Inverse Weighting Methods

March 11, 2011

IPW Estimator

- The IPW estimator (also known as the Horvitz Thompson estimator used in survey methodology) for the average treatment effect is:

$$\widehat{ATE}_{IPW} = \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{T_i Y_i}{\hat{\pi}(X_i)} - \frac{(1 - T_i) Y_i}{1 - \hat{\pi}(X_i)} \right\}$$

where $\hat{\pi}(X_i)$ is the estimated propensity score.

IPW Estimator

- The IPW estimator (also known as the Horvitz Thompson estimator used in survey methodology) for the average treatment effect is:

$$\widehat{ATE}_{IPW} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{T_i Y_i}{\hat{\pi}(X_i)} - \frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}\right\}$$

  where $\hat{\pi}(X_i)$ is the estimated propensity score.

- Common alternative:

$$\left\{\sum_{i=1}^{n}\frac{T_i}{\hat{\pi}(X_i)}\right\}^{-1}\sum_{i=1}^{n}\frac{T_i Y_i}{\hat{\pi}(X_i)} - \left\{\sum_{i=1}^{n}\frac{1-T_i}{1-\hat{\pi}(X_i)}\right\}^{-1}\frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}$$

IPW Estimator

- The IPW estimator (also known as the Horvitz Thompson estimator used in survey methodology) for the average treatment effect is:

$$\widehat{ATE}_{IPW} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{T_i Y_i}{\hat{\pi}(X_i)} - \frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}\right\}$$

  where $\hat{\pi}(X_i)$ is the estimated propensity score.

- Common alternative:

$$\left\{\sum_{i=1}^{n}\frac{T_i}{\hat{\pi}(X_i)}\right\}^{-1}\sum_{i=1}^{n}\frac{T_i Y_i}{\hat{\pi}(X_i)} - \left\{\sum_{i=1}^{n}\frac{1-T_i}{1-\hat{\pi}(X_i)}\right\}^{-1}\frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}$$

- IPW estimator "upweights" treated units with a low probability of treatment and "down-weights" controls that have a high probability of being in treatment.

Unbiasedness of the IPW Estimator

$$E[\widehat{ATE}_{IPW}] = \frac{1}{n}\sum_{i=1}^{n}\left\{E\left[\frac{T_i Y_i}{\pi(X_i)}\right] - E\left[\frac{(1-T_i)Y_i}{1-\pi(X_i)}\right]\right\}$$

Unbiasedness of the IPW Estimator

$$
\begin{aligned}
E[\widehat{ATE}_{IPW}] &= \frac{1}{n}\sum_{i=1}^{n}\left\{E\left[\frac{T_i Y_i}{\pi(X_i)}\right] - E\left[\frac{(1-T_i)Y_i}{1-\pi(X_i)}\right]\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y(1)}{\pi(X_i)}E\left[T_i\right] - \frac{Y(0)}{1-\pi(X_i)}E\left[(1-T_i)\right]\right\}
\end{aligned}
$$

Unbiasedness of the IPW Estimator

$$
\begin{aligned}
E[\widehat{ATE}_{IPW}] &= \frac{1}{n}\sum_{i=1}^{n}\left\{E\left[\frac{T_iY_i}{\pi(X_i)}\right] - E\left[\frac{(1-T_i)Y_i}{1-\pi(X_i)}\right]\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y(1)}{\pi(X_i)}E\left[T_i\right] - \frac{Y(0)}{1-\pi(X_i)}E\left[(1-T_i)\right]\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y(1)}{\pi(X_i)}\pi(X_i) - \frac{Y(0)}{1-\pi(X_i)}1-\pi(X_i)\right\}
\end{aligned}
$$

Unbiasedness of the IPW Estimator

$$
\begin{aligned}
E[\widehat{ATE}_{IPW}] &= \frac{1}{n}\sum_{i=1}^{n}\left\{E\left[\frac{T_i Y_i}{\pi(X_i)}\right] - E\left[\frac{(1-T_i)Y_i}{1-\pi(X_i)}\right]\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y(1)}{\pi(X_i)}E\left[T_i\right] - \frac{Y(0)}{1-\pi(X_i)}E\left[(1-T_i)\right]\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y(1)}{\pi(X_i)}\pi(X_i) - \frac{Y(0)}{1-\pi(X_i)}1-\pi(X_i)\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{Y_i(1) - Y_i(0)\right\}
\end{aligned}
$$

Issues with the IPW Estimator

- Advantages:

Issues with the IPW Estimator

- Advantages:
  - More efficient (lower variance) than other matching estimators (Imbens, Hirano, and Ridder (2003)), standard variance estimators work (including the bootstrap).

Issues with the IPW Estimator

- Advantages:
  - More efficient (lower variance) than other matching estimators (Imbens, Hirano, and Ridder (2003)), standard variance estimators work (including the bootstrap).
  - Easy to implement: regression with weights

Issues with the IPW Estimator

- Advantages:
  - More efficient (lower variance) than other matching estimators (Imbens, Hirano, and Ridder (2003)), standard variance estimators work (including the bootstrap).
  - Easy to implement: regression with weights
- Disadvantages: boundary problems (propensity scores close to 0 or 1): " be a hero, don't divide by zero"

Issues with the IPW Estimator

- Advantages:
  - More efficient (lower variance) than other matching estimators (Imbens, Hirano, and Ridder (2003)), standard variance estimators work (including the bootstrap).
  - Easy to implement: regression with weights
- Disadvantages: boundary problems (propensity scores close to 0 or 1): " be a hero, don't divide by zero"
- Still possible to check balance.

Augmented Inverse Propensity Weighting Estimator

- Model both the propensity to be treated AND the outcome.

Augmented Inverse Propensity Weighting Estimator

- Model both the propensity to be treated AND the outcome.
- This estimator is *doubly* robust in the sense that it is consistent for the ATE whenever

Augmented Inverse Propensity Weighting Estimator

- Model both the propensity to be treated AND the outcome.
- This estimator is *doubly* robust in the sense that it is consistent for the ATE whenever
  1. The propensity score model is correctly specified.

Augmented Inverse Propensity Weighting Estimator

- Model both the propensity to be treated AND the outcome.
- This estimator is *doubly* robust in the sense that it is consistent for the ATE whenever
  1. The propensity score model is correctly specified.
  2. The outcome regression is correctly specified.

Augmented Inverse Propensity Weighting Estimator

- Model both the propensity to be treated AND the outcome.
- This estimator is *doubly* robust in the sense that it is consistent for the ATE whenever
  1. The propensity score model is correctly specified.
  2. The outcome regression is correctly specified.
- If one of the models is correctly specified, then the estimator is consistent.

Augmented Inverse Propensity Weighting Estimator

- Model both the propensity to be treated AND the outcome.
- This estimator is *doubly* robust in the sense that it is consistent for the ATE whenever
  1. The propensity score model is correctly specified.
  2. The outcome regression is correctly specified.
- If one of the models is correctly specified, then the estimator is consistent.
- If both models are correct, then the AIPW achieves the semiparametric efficiency bound.

Augmented Inverse Propensity Weighting Estimator

- Model both the propensity to be treated AND the outcome.
- This estimator is *doubly* robust in the sense that it is consistent for the ATE whenever
  1. The propensity score model is correctly specified.
  2. The outcome regression is correctly specified.
- If one of the models is correctly specified, then the estimator is consistent.
- If both models are correct, then the AIPW achieves the semiparametric efficiency bound.
- If weights are highly variable, then AIPW is not robust.

Augmented Inverse Propensity Weighting Estimator

- Model both the propensity to be treated AND the outcome.
- This estimator is *doubly* robust in the sense that it is consistent for the ATE whenever
  1. The propensity score model is correctly specified.
  2. The outcome regression is correctly specified.
- If one of the models is correctly specified, then the estimator is consistent.
- If both models are correct, then the AIPW achieves the semiparametric efficiency bound.
- If weights are highly variable, then AIPW is not robust.
- Many estimators have this double robustness property, including regression on a matched dataset.

AIPW Estimator

$$\begin{aligned}\widehat{ATE} &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{T_i Y_i}{\hat{\pi}(X_i)} - \frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}\right] - \frac{T_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\ &\times \quad [(1-\hat{\pi}(X_i))E(Y_i|T_i=1,X_i) + \hat{\pi}(X_i)E(Y_i|T_i=0,X_i)]\end{aligned}$$

- First line corresponds to basic IPW estimator.

AIPW Estimator

$$
\begin{aligned}
\widehat{ATE} &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{T_iY_i}{\hat{\pi}(X_i)} - \frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}\right] - \frac{T_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\
&\times \quad [(1-\hat{\pi}(X_i))E(Y_i|T_i=1,X_i) + \hat{\pi}(X_i)E(Y_i|T_i=0,X_i)]
\end{aligned}
$$

- First line corresponds to basic IPW estimator.
- Second line adjusts this estimator by a weighted average of the two regression estimators.

AIPW Estimator

$$\begin{aligned}\widehat{ATE} &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{T_i Y_i}{\hat{\pi}(X_i)} - \frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}\right] - \frac{T_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\ &\times \quad [(1-\hat{\pi}(X_i))E(Y_i|T_i=1,X_i) + \hat{\pi}(X_i)E(Y_i|T_i=0,X_i)]\end{aligned}$$

- First line corresponds to basic IPW estimator.
- Second line adjusts this estimator by a weighted average of the two regression estimators.
- The adjustment set X_i need not be the same ones used in the propensity score.

AIPW Estimator

- AIPW estimator is more robust to small weights than the IPW estimator. To see why, set $T_i = 1$. The estimator becomes:

$$
\begin{aligned}
& \tfrac{Y_i}{\hat{\pi}(X_i)} - \tfrac{1}{\hat{\pi}(X_i)} \\
\times \quad & [(1 - \hat{\pi}(X_i))E(Y_i | T_i = 1, X_i) + \hat{\pi}(X_i)E(Y_i | T_i = 0, X_i)] \\
= \quad & \left[\tfrac{Y_i}{\hat{\pi}(X_i)} - \tfrac{[1-\hat{\pi}(X_i)E(Y|T=1,X)]}{\hat{\pi}(X_i)}\right] - E(Y|T=0,X)
\end{aligned}
$$

AIPW Estimator

- AIPW estimator is more robust to small weights than the IPW estimator. To see why, set $T_i = 1$. The estimator becomes:

$$
\begin{aligned}
& \frac{Y_i}{\hat{\pi}(X_i)} - \frac{1}{\hat{\pi}(X_i)} \\
\times \quad & [(1 - \hat{\pi}(X_i))E(Y_i | T_i = 1, X_i) + \hat{\pi}(X_i)E(Y_i | T_i = 0, X_i)] \\
= \quad & \left[\frac{Y_i}{\hat{\pi}(X_i)} - \frac{[1-\hat{\pi}(X_i)E(Y|T=1,X)]}{\hat{\pi}(X_i)}\right] - E(Y|T = 0, X)
\end{aligned}
$$

- When $\hat{\pi}(X_i)$ is close to 0, then $\frac{Y_i}{\hat{\pi}(X_i)}$ will get large in absolute value. This is a problem.

AIPW Estimator

- AIPW estimator is more robust to small weights than the IPW estimator. To see why, set $T_i = 1$. The estimator becomes:

$$
\begin{aligned}
& \tfrac{Y_i}{\hat{\pi}(X_i)} - \tfrac{1}{\hat{\pi}(X_i)} \\
\times \quad & [(1-\hat{\pi}(X_i))E(Y_i|T_i=1,X_i) + \hat{\pi}(X_i)E(Y_i|T_i=0,X_i)] \\
= \quad & \left[\tfrac{Y_i}{\hat{\pi}(X_i)} - \tfrac{[1-\hat{\pi}(X_i)E(Y|T=1,X)]}{\hat{\pi}(X_i)}\right] - E(Y|T=0,X)
\end{aligned}
$$

- When $\hat{\pi}(X_i)$ is close to 0, then $\frac{Y_i}{\hat{\pi}(X_i)}$ will get large in absolute value. This is a problem.
- However, it is mitigated by the fact that $frac[1-\hat{\pi}(X_i)E(Y|T=1,X)]\hat{\pi}(X_i)$ gets large at the same rate as $\frac{Y_i}{\hat{\pi}(X_i)}$

Unbiasedness of the AIPW Estimator

$$
\begin{aligned}
\widehat{ATE} &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{T_i Y_i}{\hat{\pi}(X_i)} - \frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}\right] - \frac{T_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\
&\times \left[(1-\hat{\pi}(X_i))E(Y_i|T_i=1,X_i) + \hat{\pi}(X_i)E(Y_i|T_i=0,X_i)\right]
\end{aligned}
$$

Unbiasedness of the AIPW Estimator

$$
\begin{aligned}
\widehat{ATE} &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{T_iY_i}{\hat{\pi}(X_i)} - \frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}\right] - \frac{T_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\
&\times \quad [(1-\hat{\pi}(X_i))E(Y_i \mid T_i = 1, X_i) + \hat{\pi}(X_i)E(Y_i \mid T_i = 0, X_i)] \\
&= \frac{1}{n}\sum_{i=1}^{n}\{Y_i(1) - Y_i(0)\} - \frac{T_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\
&\times \quad [(1-\hat{\pi}(X_i))E(Y_i \mid T_i = 1, X_i) + \hat{\pi}(X_i)E(Y_i \mid T_i = 0, X_i)] \\
&= \frac{1}{n}\sum_{i=1}^{n}\{Y_i(1) - Y_i(0)\} - \frac{\pi_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\
&\times \quad [(1-\hat{\pi}(X_i))E(Y_i \mid T_i = 1, X_i) + \hat{\pi}(X_i)E(Y_i \mid T_i = 0, X_i)]
\end{aligned}
$$

Unbiasedness of the AIPW Estimator

$$
\begin{aligned}
\widehat{ATE} &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{T_iY_i}{\hat{\pi}(X_i)} - \frac{(1-T_i)Y_i}{1-\hat{\pi}(X_i)}\right] - \frac{T_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\
&\times \quad [(1-\hat{\pi}(X_i))E(Y_i|T_i=1,X_i) + \hat{\pi}(X_i)E(Y_i|T_i=0,X_i)] \\
&= \frac{1}{n}\sum_{i=1}^{n}\{Y_i(1) - Y_i(0)\} - \frac{T_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\
&\times \quad [(1-\hat{\pi}(X_i))E(Y_i|T_i=1,X_i) + \hat{\pi}(X_i)E(Y_i|T_i=0,X_i)] \\
&= \frac{1}{n}\sum_{i=1}^{n}\{Y_i(1) - Y_i(0)\} - \frac{\pi_i - \hat{\pi}_i(X_i)}{\hat{\pi}(X_i)(1-\hat{\pi}(X_i))} \\
&\times \quad [(1-\hat{\pi}(X_i))E(Y_i|T_i=1,X_i) + \hat{\pi}(X_i)E(Y_i|T_i=0,X_i)] \\
&= \frac{1}{n}\sum_{i=1}^{n}\{Y_i(1) - Y_i(0)\}
\end{aligned}
$$