RAC plots and misspecification

Rachael Caelie (Rocky) Aikens

3/2/2021

Intro: Null expectations about AC plots

One interesting fact that has emerged in our exploration of AC plots is: *not all assignment-control plots are possible.* In particular, if two points have the same horizontal position (same propensity score), they must have the same probability of treatment, by definition. In pictures, this means that the red treated individuals must be scattered evenly across horizontal level-sets.

One interesting implication for this is that an AC plot may be able to indicate in some cases whether the propensity score has been misspecified.

Set-up

The results that follow depict several simulated datasets. The primary generative model for these is based on Aikens, Greaves, and Baiocchi (2020) and is specified as follows:

$$
\begin{aligned}
X_i &\sim Normal(0, I_{10}) \\
T_i &\sim Bernoulli\left(\frac{1}{1+exp(\phi(X_i))}\right) \\
Y_i(0) &= \Psi(X_i) + \epsilon_i \\
\epsilon_i &\sim N(0,1)
\end{aligned}
$$

where $\phi(X)$ and $\Psi(X)$ represent the true propensity and prognostic score functions, given by

$$
\begin{aligned}
\phi(X_i) &= c_1 X_{i1} - c_0 \\
\Psi(X_i) &= \rho X_{i1} + \sqrt{(1-\rho^2)} X_{i2}.
\end{aligned}
$$

Where c_1, c_0, and ρ are constants. In particular, the form for the prognostic function above guarantees that $\rho = Corr(\phi(X), \Psi(X))$.

Aikens, Rachael C, Dylan Greaves, and Michael Baiocchi. 2020. "A Pilot Design for Observational Studies: Using Abundant Data Thoughtfully." *Statistics in Medicine*. Wiley Online Library.