Propensity score methods for continuous treatments

Walter Leite

Example

* Objective: Estimate the effect of Algebra nation usage on school means on Algebra 1 Exam.

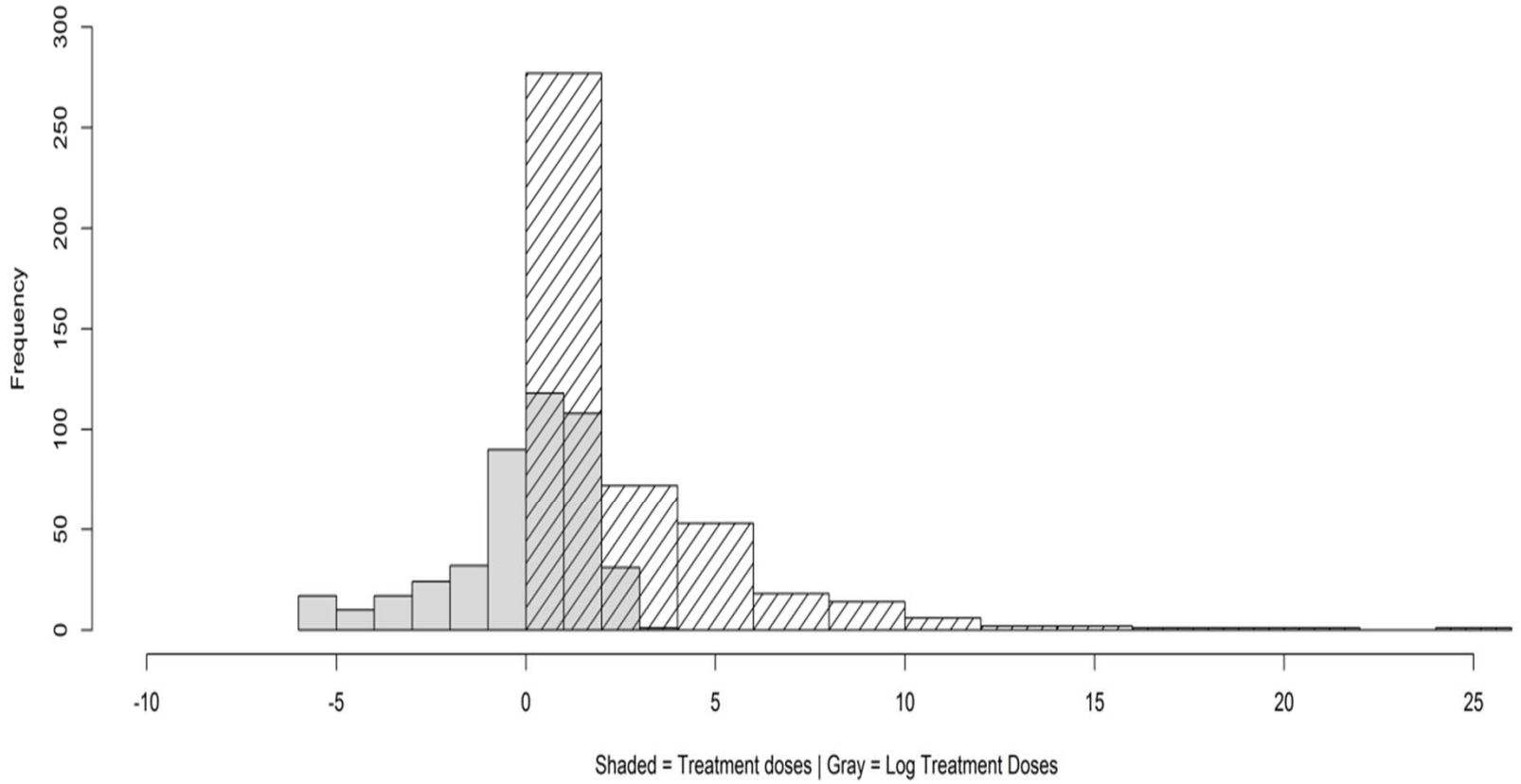
* Algebra Nation is a virtual learning environment with widespread use in Florida. School-level program implementation involves integration with school network, teacher training and providing free workbooks.

* Many measures of treatment dose possible. We will use the log of ratio of total logins per examinee.

* The outcome is school-level mean scores on Algebra I End-of-Course assessment.

* The sample has 448 schools.

Log transformation of treatment doses



Covariates for example

* 13 covariates are selected. Examples:
  - school-level means in the Algebra I EOC exam in 2012 the number of students,
  - percentage of students eligible for free lunch,
  - dummy indicators of whether the school was a charter, magnet, or Title I in 2012

Propensity score methods for continuous treatments

* Hirano and Imbens (2004) generalized propensity scores.
* Robins, Hernán and Brumback (2000) inverse probability weights.
* Imai and Van Dyk (2004) propensity score function

Rubin's Causal Model for continuous treatments

* For each dose of treatment, there is a vector of potential outcomes for all doses.

* The individual treatment effect at a dose is the mean of the potential outcomes.

$$\mu(z) = E\left[Y_i(z)\right]$$

Weak Unconfoundness for treatment doses

* The potential outcomes of the treatment at dose z is independent of treatment dosage assignment given covariates.

$$Y_i(z) \perp Z_i(\mathrm{z}) \mid \mathrm{X}$$

Steps of Hirano and Imbens (2004) method

1) Model the continuous treatment indicator as a function of covariates;

2) Obtain generalized propensity scores;

3) Model the outcomes as a function of treatment and generalized propensity scores.

4) Estimate the average potential outcome at each treatment dose of interest (the dose response function).

Step 1: Model for continuous treatment

* This is a parametric model of the relationship between treatment dose and covariates.

$$Z_i = \beta_0 + \beta X_i + \varepsilon_i$$

$$e_i \sim N(0, \sigma_i^2)$$

Step 2: The generalized propensity score

* The generalized propensity score for each individual *i* is the conditional density of the treatment evaluated at the individual's specific values of Z and X.

$$r(z,x) = f_{Z|X}(z \mid x)$$

Conditional density function of the normal distribution

Model for treatment dose

$$r(z,x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \left(Z_i - \beta_0 - \beta X_i \right)^2 \right)$$

Weak unconfoundess and the generalized propensity score

* If weak unconfoundess holds for observed covariates, it also holds given a correctly-specified generalized propensity score (GPS).

$$f_Z(\mathrm{z} \mid \mathrm{r}(\mathrm{Z}, \mathrm{X}), \mathrm{Y}(\mathrm{Z})) = f_Z(\mathrm{z} \mid \mathrm{r}(\mathrm{Z}, \mathrm{X}))$$

Evaluation of covariate balance

* stratify based on GPS and fit one regression for each covariate with the treatment dose as the outcome and with GPS strata and the covariate as predictors.

* standardized regression coefficients measure of the effect size of the covariates on treatment dose.

* Covariate balance is achieved if the effect sizes of the covariates and covariate-by-GPS interactions are small.

Results of covariate balance

Variable	Baseline	GPS Strata
Charter	0.321	**0.346**
Magnet	0.459	**0.319**
Title I school	0.105	**0.024**
Rural school	0.321	**0.034**
Location size	0.397	**0.215**
Number of students	0.000	**0.000**
Senior high	0.064	**0.000**
Number of students 2014	0.106	**0.038**
Mean scale 2012	0.117	**0.077**
Level 1 percent 2012	0.108	**0.064**
Level 5 percent 2012	0.084	**0.093**

Step 3: The model for the outcome

$$E\left[Y_i \mid Z_i, r_i(Z,X)\right] = \gamma_0 + \gamma_1 Z_i + \gamma_2 Z_i^2 + \gamma_3 r_i(Z,X) + \gamma_4 r_i(Z,X)^2 + \gamma_4 Z_i r_i(Z,X)$$

* The coefficients from this outcome model do not have a causal interpretation
* This is an intermediate step for obtaining the dose response function.

Step 4: Obtain individual treatment effects and plot dose response function

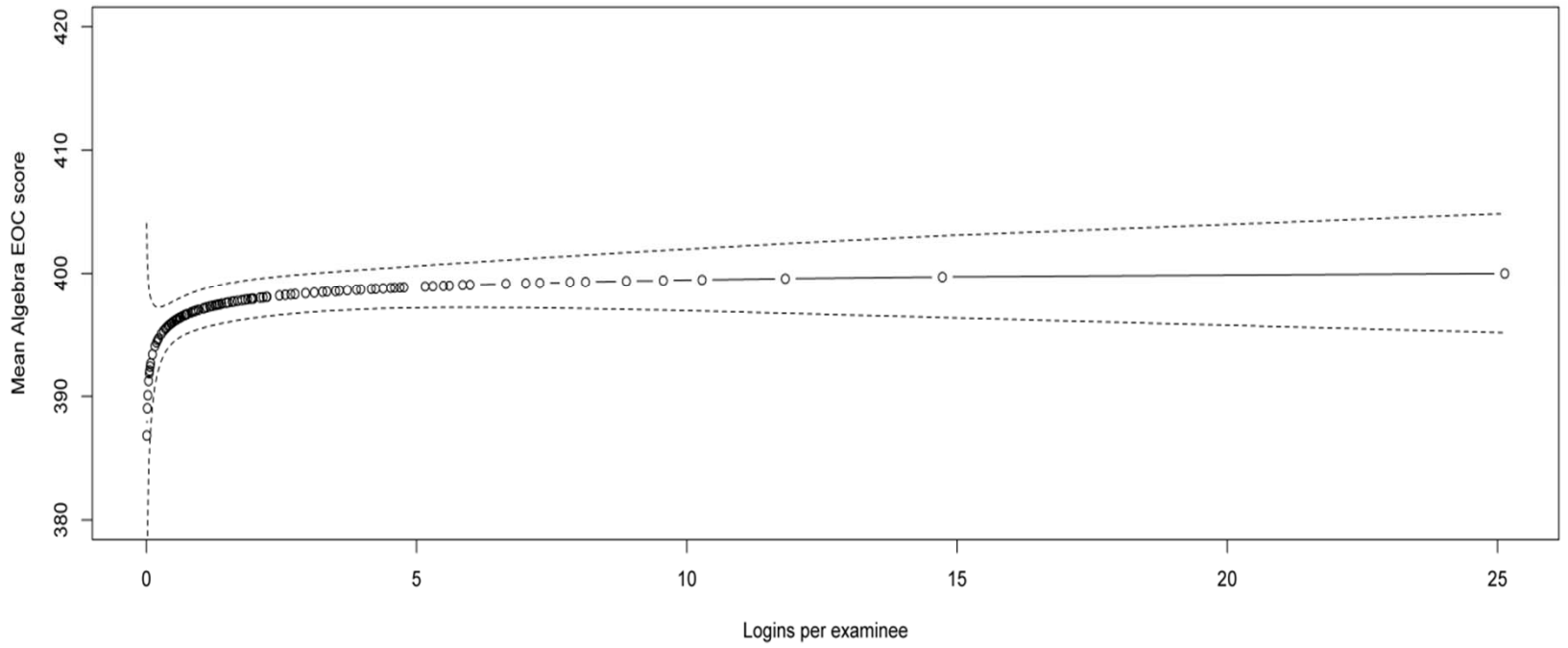
* Individual treatment effects are the average potential outcome at treatment dose *z* given the coefficients of the outcome model

$$E\left[Y_i(z)\right] = \frac{1}{N}\sum_{i=1}^{N}\left(\gamma_0 + \gamma_1 z + \gamma_2 z^2 + \gamma_3 r_i(Z,X) + \gamma_4 r_i(Z,X) + \gamma_4 z r_i(Z,X)\right)$$

Treatment effect at each dose of treatment

Percentile	Logins per examinee	Mean algebra EOC score	Standard error	Lower confidence limit	Upper confidence limit
10%	0.16	394.00	1.68	390.70	397.29
20%	0.46	395.94	0.87	394.24	397.65
30%	0.71	396.63	0.81	395.05	398.21
40%	1.08	397.24	0.80	395.66	398.81
50%	1.52	397.68	0.81	396.09	399.26
60%	2.11	398.07	0.81	396.49	399.64
70%	3.28	398.53	0.81	396.94	400.12
80%	4.52	398.83	0.85	397.16	400.50
90%	6.65	399.14	1.00	397.19	401.09
100%	25.13	399.83	2.48	394.96	404.70

Example of plot of dose response function



Inverse Probability Weighting Robins, Hernán & Brumback (2000)

* Balances distribution with respect to covariates by creating a pseudo-population.

* In the pseudo-population, a set of individuals receiving dose *z* of treatment and another set receiving a different dose has similar distributions of covariates

* It can be used to estimate the average treatment effect of a continuous treatment.

Calculation of inverse probability weights for continuous treatments

$$w_i = \frac{f_Z(\mathrm{z})}{f_{Z|\mathrm{X}}(\mathrm{z} \mid \mathrm{x})}$$

$$f_Z(\mathrm{z}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \left(Z_i - \mu_i \right)^2 \right)$$

$$r(\mathrm{z}, \mathrm{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \left(Z_i - \beta_0 - \beta X_i \right)^2 \right)$$

Covariate Balance Evaluation with IPW

* Estimate bivariate regressions of the treatment dosage on each covariate with and without the IPW
* Compare the regression coefficients obtained with and without the use of the IPW.
* In the pseudo-population generated by the IPW, the treatment dose should be uncorrelated with the covariates
* Coefficients of the regression of the treatment dose on the covariate should be close to zero.

Results of covariate balance

Variable	Baseline	IPW
Charter	0.321	0.200
Magnet	0.459	0.019
Title I school	0.105	0.054
Rural school	0.321	0.007
Location size	0.397	0.071
Number of students	0.000	0.000
Senior high	0.064	0.085
Number of students 2014	0.106	0.030
Mean scale 2012	0.117	0.019
Level 1 percent 2012	0.108	0.027
Level 5 percent 2012	0.084	0.017

Estimation of the average treatment effect with IPW

* Any parametric model can be used that incorporates the IPW into the estimation method. For example:

$$y_i = \beta_0 + \beta_1 Z_i + \varepsilon_i$$

* Covariates can be included in the model to:
  * Provide additional adjustment of selection bias
  * Increase power
  * Achieve double robustness

Results of estimation of treatment effects with IPTW for example

Model in R: meanScale2014 ~ logLoginsPerExaminee + Charter + Charter:logLoginsPerExaminee

```
Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                397.42444    0.52839 752.138   <2e-16 ***
logLoginsPerExaminee         0.66387    0.31777   2.089   0.0373 *
Charter                      4.33191    2.25213   1.923   0.0551 .
logLoginsPerExaminee:Charter 0.00714    1.15741   0.006   0.9951
```