Causal Inference Methods and Case Studies

STAT24630

Jingshu Wang

Lecture 15

Topic: Doubly robust estimator

- Doubly robust estimator

Doubly robust estimator

- Outcome regression relies on a correctly specified model for the (potential) outcomes depending on $\boldsymbol{X}_i$
- IPW / Matching relies on a correctly specified model for the propensity score

- Doubly robust estimator: provide a good estimate of the propensity score when either the outcome or the propensity score model is correct

- Define

$$f(1, \boldsymbol{X}_i, Y_i^{\text{obs}}) = \frac{Y_i^{\text{obs}} 1_{W_i=1}}{\tilde{e}(\boldsymbol{X}_i)} - \frac{1_{W_i=1} - \tilde{e}(\boldsymbol{X}_i)}{\tilde{e}(\boldsymbol{X}_i)} \tilde{\mu}_1(\boldsymbol{X}_i)$$

$$f(0, \boldsymbol{X}_i, Y_i^{\text{obs}}) = \frac{Y_i^{\text{obs}} 1_{W_i=0}}{1 - \tilde{e}(\boldsymbol{X}_i)} - \frac{1_{W_i=0} - (1 - \tilde{e}(\boldsymbol{X}_i))}{1 - \tilde{e}(\boldsymbol{X}_i)} \tilde{\mu}_0(\boldsymbol{X}_i)$$

  - If we correctly specify the propensity score model, then $\tilde{e}(\boldsymbol{X}_i) = e(\boldsymbol{X}_i)$
  - If we correctly specify the outcome model, then $\tilde{\mu}_w(\boldsymbol{X}_i) = \mu_w(\boldsymbol{X}_i)$

Doubly robust estimator

$$f(1, \boldsymbol{X}_i, Y_i^{\text{obs}}) = \frac{Y_i^{\text{obs}} 1_{W_i=1}}{\tilde{e}(\boldsymbol{X}_i)} - \frac{1_{W_i=1} - \tilde{e}(\boldsymbol{X}_i)}{\tilde{e}(\boldsymbol{X}_i)} \tilde{\mu}_1(\boldsymbol{X}_i)$$

$$f(0, \boldsymbol{X}_i, Y_i^{\text{obs}}) = \frac{Y_i^{\text{obs}} 1_{W_i=0}}{1 - \tilde{e}(\boldsymbol{X}_i)} - \frac{1_{W_i=0} - (1 - \tilde{e}(\boldsymbol{X}_i))}{1 - \tilde{e}(\boldsymbol{X}_i)} \tilde{\mu}_0(\boldsymbol{X}_i)$$

- $\tilde{e}(\boldsymbol{X}_i)$, $\tilde{\mu}_w(\boldsymbol{X}_i)$: our working models (model under our model assumption)
- $e(\boldsymbol{X}_i)$, $\mu_w(\boldsymbol{X}_i)$: true model that we don't know

- Double robust property

$$\mathbb{E}\left[f(1, \boldsymbol{X}_i, Y_i^{\text{obs}}) \mid \boldsymbol{X}_i\right] = \frac{\left(\mu_1(\boldsymbol{X}_i) - \tilde{\mu}_1(\boldsymbol{X}_i)\right)\left(e(\boldsymbol{X}_i) - \tilde{e}(\boldsymbol{X}_i)\right)}{\tilde{e}(\boldsymbol{X}_i)} + \mu_1(\boldsymbol{X}_i)$$

$$\mathbb{E}\left[f(0, \boldsymbol{X}_i, Y_i^{\text{obs}}) \mid \boldsymbol{X}_i\right] = \frac{\left(\mu_0(\boldsymbol{X}_i) - \tilde{\mu}_0(\boldsymbol{X}_i)\right)\left(\tilde{e}(\boldsymbol{X}_i) - e(\boldsymbol{X}_i)\right)}{1 - \tilde{e}(\boldsymbol{X}_i)} + \mu_0(\boldsymbol{X}_i)$$

  - If either the outcome or propensity score model is correct, we have

$$\mathbb{E}\left(f\left(w, \boldsymbol{X}_i, Y_i^{\text{obs}}\right)\right) = \mathbb{E}(Y_i(w) \mid \boldsymbol{X}_i)$$

Doubly robust estimator

$$f(1, \boldsymbol{X}_i, Y_i^{\text{obs}}) = \underbrace{\frac{Y_i^{\text{obs}} 1_{W_i=1}}{\tilde{e}(\boldsymbol{X}_i)}} - \underbrace{\frac{1_{W_i=1} - \tilde{e}(\boldsymbol{X}_i)}{\tilde{e}(\boldsymbol{X}_i)} \tilde{\mu}_1(\boldsymbol{X}_i)}$$

IPW estimate of $\mathbb{E}(Y_i(1) | \boldsymbol{X}_i)$

Adjust for bias if the propensity score model is incorrect (if PS model is correct, then this part has expectation 0)

An equivalent expression:

$$f(1, \boldsymbol{X}_i, Y_i^{\text{obs}}) = \underbrace{\tilde{\mu}_1(\boldsymbol{X}_i)} + \underbrace{\frac{1_{W_i=1}}{\tilde{e}(\boldsymbol{X}_i)} \left(Y_i(1) - \tilde{\mu}_1(\boldsymbol{X}_i) \right)}$$

Outcome regression estimate of $\mathbb{E}(Y_i(1) | \boldsymbol{X}_i)$

Adjust for bias if the outcome regression model is incorrect (if PS model is correct, then this part has expectation 0)

The DR estimator:

$$\hat{\tau} = \frac{1}{N} \sum_i \left[\frac{Y_i^{\text{obs}} 1_{W_i=1}}{\hat{e}(\boldsymbol{X}_i)} - \frac{1_{W_i=1} - \hat{e}(\boldsymbol{X}_i)}{\hat{e}(\boldsymbol{X}_i)} \hat{\mu}_1(\boldsymbol{X}_i) \right] - \frac{1}{N} \sum_i \left[\frac{Y_i^{\text{obs}} 1_{W_i=0}}{1 - \hat{e}(\boldsymbol{X}_i)} - \frac{1_{W_i=0} - (1 - \hat{e}(\boldsymbol{X}_i))}{1 - \hat{e}(\boldsymbol{X}_i)} \hat{\mu}_0(\boldsymbol{X}_i) \right]$$

A simulation study (Kang and Schafer. 2007. Statistical Science)

- The deteriorating performance of propensity score weighting methods when the model is mis-specified
- Setup:
  - 4 covariates X_i^*: all are i.i.d. standard normal
  - Outcome model: linear model
  - Propensity score model: logistic model with linear predictors
  - Misspecification induced by measurement error:
    - $X_{i1} = \exp(X_{i1}^*/2)$
    - $X_{i2} = X_{i2}^*/(1 + \exp(X_{1i}^*) + 10)$
    - $X_{i3} = (X_{i1}^* X_{i3}^*/25 + 0.6)^3$
    - $X_{i4} = (X_{i1}^* + X_{i4}^* + 20)^2$
- Weighting estimators to be evaluated:
  - HT: IPW in the original form
  - IPW: IPW with normalized weights
  - Weighted least squares regression with covariates
  - Doubly-robust estimator

Results: if the propensity score model is correct

		Bias		RMSE	
Sample size	Estimator	logit	True	logit	True
(1) Both models correct					
$n = 200$	HT	0.33	1.19	12.61	23.93
	IPW	−0.13	−0.13	3.98	5.03
	WLS	−0.04	−0.04	2.58	2.58
	DR	−0.04	−0.04	2.58	2.58
$n = 1000$	HT	0.01	−0.18	4.92	10.47
	IPW	0.01	−0.05	1.75	2.22
	WLS	0.01	0.01	1.14	1.14
	DR	0.01	0.01	1.14	1.14
(2) Propensity score model correct					
$n = 200$	HT	−0.05	−0.14	14.39	24.28
	IPW	−0.13	−0.18	4.08	4.97
	WLS	0.04	0.04	2.51	2.51
	DR	0.04	0.04	2.51	2.51
$n = 1000$	HT	−0.02	0.29	4.85	10.62
	IPW	0.02	−0.03	1.75	2.27
	WLS	0.04	0.04	1.14	1.14
	DR	0.04	0.04	1.14	1.14

- Use the true propensity score is worse than using the estimated propensity score when the propensity score model is correct
- Normalizing weights can help a lot in reducing the variance

Results: if the propensity score model is incorrect

		Bias		RMSE	
Sample size	Estimator	logit	True	logit	True
(3) Outcome model correct					
$n = 200$	HT	24.25	−0.18	194.58	23.24
	IPW	1.70	−0.26	9.75	4.93
	WLS	−2.29	0.41	4.03	3.31
	DR	−0.08	−0.10	2.67	2.58
$n = 1000$	HT	41.14	−0.23	238.14	10.42
	IPW	4.93	−0.02	11.44	2.21
	WLS	−2.94	0.20	3.29	1.47
	DR	0.02	0.01	1.89	1.13
(4) Both models incorrect					
$n = 200$	HT	30.32	−0.38	266.30	23.86
	IPW	1.93	−0.09	10.50	5.08
	WLS	−2.13	0.55	3.87	3.29
	DR	−7.46	0.37	50.30	3.74
$n = 1000$	HT	101.47	0.01	2371.18	10.53
	IPW	5.16	0.02	12.71	2.25
	WLS	−2.95	0.37	3.30	1.47
	DR	−48.66	0.08	1370.91	1.81

- Double robust estimator perform better when outcome model is correct but propensity score model is wrong
- Double robust estimator can perform worse when both models are wrong (maybe we should also normalize the weights in DR)