Denis Ostroushko - HW2

Introduction

Imputation

Following instructions from *Homework 1*, we identify Birthweight, N.prev.preg, Race_ethnicity, and Use.Tob as variables that are subject to imputation..

For the purpose of this problem, we need to impute categorical variables with mode values, and continuous variables with medians.

Median Birth weight from observed values is 3270, median number of previous pregnancies is 1.

Most common value of Race/Ethnicity Variable is Hispanic (any race), most common status of tobacoo usage is No

Study Objective

Like in Assignment 1, we consider indicator for pregnancies that ended before 37 weeks, `Preg.ended...37.wk`, and birth weights of newborns as outcomes. We will use IPW2 and Propensity Score Stratification methods to obtain average treatment effects. For pregnancy terms, we wish to test if proportion of pregnancies that ended before 37 weeks reduced for the treated sample. We also wish to test if the average birth weight of newborns was higher in the treatment group.

Study Sample Summary

Outcome variable is treatment assignment. There are 184 patients in the treatment group, which corresponds to the 31.19% of the study population.

Table 1: Summary of Confoudnding Variables Between Treatment Groups

	level	C	T	SMD
n		406	184	
Preg.ended...37.wk (%)	No	353 (86.9)	166 (90.2)	0.103
	Yes	53 (13.1)	18 (9.8)	
Birthweight (mean (SD))		3181.48 (724.83)	3259.16 (574.17)	0.119
Race_ethnicity (%)	Black	127 (31.3)	81 (44.0)	0.299
	Hispanic (any race)	247 (60.8)	85 (46.2)	
	Other	32 (7.9)	18 (9.8)	
Public.Asstce (%)	No	97 (23.9)	48 (26.1)	0.051
	Yes	309 (76.1)	136 (73.9)	
Use.Tob (%)	No	362 (89.2)	167 (90.8)	0.053
	Yes	44 (10.8)	17 (9.2)	
N.prev.preg (mean (SD))		1.86 (1.81)	1.67 (1.73)	0.111
Live.PTB (%)	No	363 (89.4)	175 (95.1)	0.215
	Yes	43 (10.6)	9 (4.9)	
BL.GE (mean (SD))		1.42 (0.39)	1.49 (0.42)	0.173
BL..BOP (mean (SD))		69.12 (17.04)	70.08 (16.99)	0.057
BL..PD.4 (mean (SD))		24.80 (15.94)	25.67 (15.48)	0.055
BL..CAL.3 (mean (SD))		14.03 (16.05)	16.91 (17.80)	0.170

Table 1 shows that there are some covariates that are imbalanced between the two treatment groups. BL..CAL.3, BL.GE, Live.PTB, Race_ethnicity all have SMD greater than 0.1.

Model Specification

In order to control for these confounding variables we will use a logistic regression model with no interactions and linear terms only. We specify the following model in order to obtain propensity score.

Let A_i be a binary random variable, with 1 representing assignment to the treatment group.

Let $\pi_i = P(A_i = 1)$ be a propensity score for i^{th} subject.

We obtain propensity scores π_i by fitting the following model:

```
propensity_score_model <- glm(

  I(data$Group == "T") %>% as.numeric() ~
    Race_ethnicity + Public.Asstce +
    Use.Tob + N.prev.preg + Live.PTB +
    BL.GE + BL..BOP + BL..PD.4 + BL..CAL.3,

  data = data,
  family = "binomial"
```

```
)

data$propensity_scores = propensity_score_model$fitted.values
```

At this point I do not verify that the model fits the data reasonably well, for the purpose of this assignment I assume that this model is appropriate for propensity score estimation (which may or may not be true).

Figure 1 shows distribution of estimated propensity scores. It appears that there are no values beyond score of 0.65.

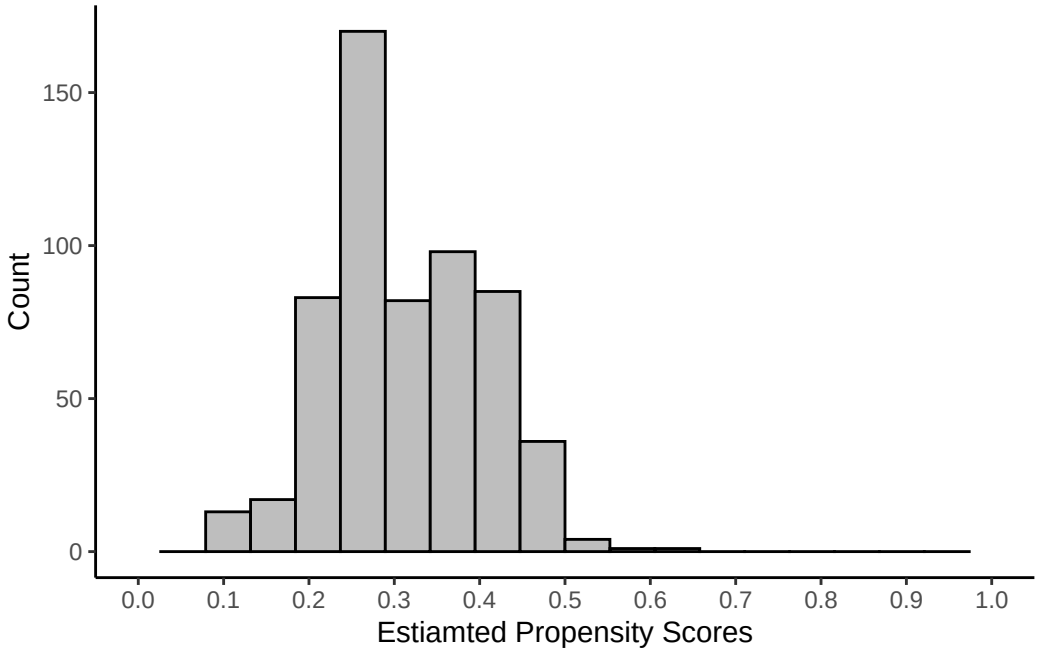

Figure 1: Distribution of Propensity Scores

Problem 1

Propensity Score Stratication

First, we divide obtain propensity scores into 5 quintiles. We observe that the first and last quintiles are quite wide, due to the distribution of scores and shape of said distribution.

Table 2 shows observed number of study subjects in each group after partitioning propensity scores.

Table 2: Distribution of observations across Propensity Score Groups and Treatment Groups

	Controls	Treatment
(0,0.238]	90	28
(0.238,0.295]	139	38
(0.295,0.333]	45	14
(0.333,0.4]	68	50
(0.4,1]	64	54

Table 3: Distribution of classes within each Propensity Score Group

	Controls	Treatment
(0,0.238]	0.76	0.24
(0.238,0.295]	0.79	0.21
(0.295,0.333]	0.76	0.24
(0.333,0.4]	0.58	0.42
(0.4,1]	0.54	0.46

Table 3 shows proportion of each class in a given partition of propensity scores. It appears, for the last two categories the balance of classes within quintiles shifts.

Impact of Treatment on Pre-Term Pregnancy Rate

We estimate ATE using

$$\hat{\delta} = \sum_{j=1}^{5} (\overline{Y}_{1j} - \overline{Y}_{0j}) \frac{n_j}{n} \tag{1}$$

In R, we use the following steps:

1. Obtain average proportion of pre-term pregnancies in each quintile in each treatment group
2. Take the difference within each quintile
3. Take a weighted sum of differences. Weigh differences using distribution of quintiles in the data.

Table 4 shows statistics from the data that we need to estimate $\hat{\delta}$.

Table 4: Statistics for average treatment effect estimation

Quintile	N Treated	% Pre-Term Pregnancies for Treated	N Controls	% Pre-Term Pregnancies for Controls	Quintile Weight
(0,0.238]	28	0.14	90	0.23	0.2
(0.238,0.295]	38	0.03	139	0.10	0.3
(0.295,0.333]	14	0.21	45	0.11	0.1
(0.333,0.4]	50	0.08	68	0.12	0.2
(0.4,1]	54	0.11	64	0.08	0.2

Using these statistics, we obtain the average treatment effect of -0.03, implying that on average this program will reduce pre-term pregnancy rates by 3.1%

Boostrap comments

In order to obtain a standard error of this estimate we follow these steps 500 times:

1. Resample data
2. Fit propensity score model on the resampled data, obtain new propensity scores
3. Obtain statistics required to estimate ATE, as shown in Table 4
4. Store ATE and return to step (1)

Conclusion

- Average (causal) treatment effect of program participation on pre-term pregnancy rates is -0.03
- Bootstrap estimate of standard error for this estimate is 0.03
- 95% normal confidence interval is (-0.09, 0.03)

Impact of Treatment on Average Birth Weight

Using the same steps we can obtain ATE on newborn birth weights. I will skip over the details and will provide a table with statistics required for the estimation process. In order to estimate ATE, we use the same propensity score model and corresponding propensity scores.

Table 5 shows intermediate statistics obtained in the estimation process.

Table 5: Statistics for average treatment effect estimation

Quintile	N Treated	Averege Birth-weights for Treated	N Controls	Averege Birth-weights for Controls	Quintile Weight
(0,0.238]	28	3241.43	90	3037.86	0.2
(0.238,0.295]	38	3338.89	139	3261.44	0.3
(0.295,0.333]	14	3112.57	45	3113.13	0.1
(0.333,0.4]	50	3295.56	68	3131.00	0.2
(0.4,1]	54	3216.56	64	3311.50	0.2

Conclusion

- Average (causal) treatment effect of program participation on average birth weights 77.82
- Bootstrap estimate of standard error for this estimate is 57.81
- 95% normal confidence interval is (-35.48, 191.12)

Inverse Probability Weighting

In this section we will use IPW2 weighting, with propensity scores we used for the previous estimation method.

Impact of Treatment on Pre-Term Pregnancy Rate

Using propensity scores, we obtain weights for each subject in the study. Distribution of weights is shown on Figure 2.

```
w_trt =
  ifelse(data$Group == "T", 1, 0)/ # vector of outcomes 0/1
  propensity_score_model$fitted.values # vector of corresponding propensity scores

w_cnt =
  (1 - ifelse(data$Group == "T", 1, 0))/ # vector that identifies '0's as '1's
  (1-propensity_score_model$fitted.values)
```

Using IPW2 weighting approach:

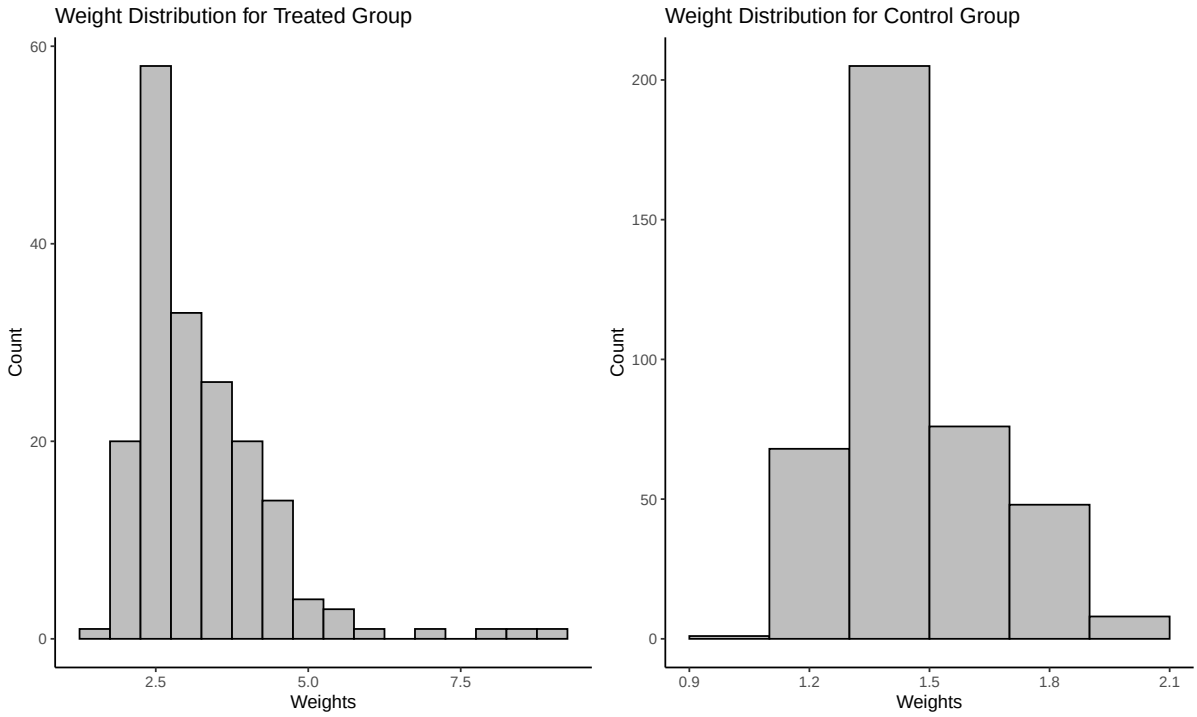

Figure 2: Distribution of Weights for Controls and Treated Using Inverse Propensity Scores

```
weighted.mean(ifelse(data$Preg.ended...37.wk == "Yes", 1, 0), w_trt) -
  weighted.mean(ifelse(data$Preg.ended...37.wk == "Yes", 1, 0), w_cnt) -> preg_ipw2_ATE
```

We employ the same bootstrap approach as described before to obtain standard error for this approach.

Bootstrapped ATEs were calculated in the same loop as Propensity Score Stratification ATEs, which means bootstrapped models and propensity scores were reused for the two estimators within each bootstrap iteration.

Conclusion

- IPW2 ATE estimate for reduction in pre-term pregnancy rates is -0.0294
- Bootstrapped standard error is 0.014
- 95% Normal Confidence Interval is (-0.06, 0)

Impact of Treatment on Average Birthweight

In this section we reuse inverse weight shown on Figure 2. Using the same steps, we obtain IPW2 estiamte for the ATE on average birth weights.

IPW2 estimation approach:

```
weighted.mean(data$Birthweight, w_trt) -
  weighted.mean(data$Birthweight, w_cnt) -> bw_ipw2_ATE
```

Conclusion

- IPW2 ATE estimate for increase in average birth weight of newborns 79.6336
- Bootstrapped standard error is 56.64
- 95% Normal Confidence Interval is (-31.37, 190.64)

Final Results Problem 1

Putting all work together:

- 1. We have fit a logistic regression to estimate each participants probability of being in the treatment group
  - We assumed the model is correct, fits the data reasonably well (which I did not verify)
  - We assume positivity, and verify that obtained propensity scores are greater than 0
- 2. Using obtained propensity scores, we identified 5 strata of propensity scores based on quintiles
- 3. For each outcome, newborn birth weights and events of pre-term pregnancies, we found ATE using propensity score stratification
- 4. For each outcome we found ATE using IPW2
- 5. Using bootstrap we found distribution of ATEs for each method for each outcome, using these distributions we found standard errors for each ATE estimate, and corresponding 95% normal confidence intervals

Table 6: Final Estimates for Pre-Term Pregnancy Reduction ATE

Method	ATE	Estimate SE	95% C.I. Lower Bound	95% C.I. Upper Bound
PSS	-0.03	0.03	-0.09	0.03
IPW2	-0.03	0.01	-0.06	0.00

Table Table 6 shows estimated quantities for treatment effect on pre-term pregnancy rates.

Table Table 7 shows estimated quantities for treatment effect on average birth weights.

Table 7: Final Estimates for Pre-Term Pregnancy Reduction ATE

Method	ATE	Estimate SE	95% C.I. Lower Bound	95% C.I. Upper Bound
PSS	77.82	57.81	-35.48	191.12
IPW2	79.63	56.64	-31.37	190.64

Problem 2

Recall that in HW1 we obtained ATE for the two outcomes using regression adjustment methods and comparing average outcomes when everyone in the population would be treated vs scenarios where no one would be treated.

I will refer to these ATEs as Regression ATE when adding them to the plot.

Figure 3 compares three methods of obtaining ATE in terms of point estimate and 95% confidence intervals.

Figure 4 compares three methods of obtaining ATE in terms of point estimate and 95% confidence intervals.

Pre-Term pregnancy results

Figure 3 shows that IPW2 method of estimation of causal effect produces an estimate with the lowest variance. Regression Adjustment and Propensity Score Stratification methods produce very similar estimates between the two methods.

Point estimates are pretty similar between the three methods.

We know that the common perception is that modeling treatment allocation can be 'easier' than modeling the outcome.

Therefore, it may be the case that a simple linear model with no interactions or higher order terms is more appropriate to model the treatment allocation.

Also, it is likely that IPW2 method can handle outliers or extreme values better, and the effect of these values is not as strong within the bootsrap replications.

Figure 4 shows that IPW2 method of estimation of causal effect produces an estimate with the lowest variance. Regression Adjustment and Propensity Score Stratification methods produce very similar estimates.

Birth weight results

Figure 4 shows that all three estimates have the same point estimates and same variance, as expressed by their confidence intervals. It appears that the data we have is not sensitive to estimation method.

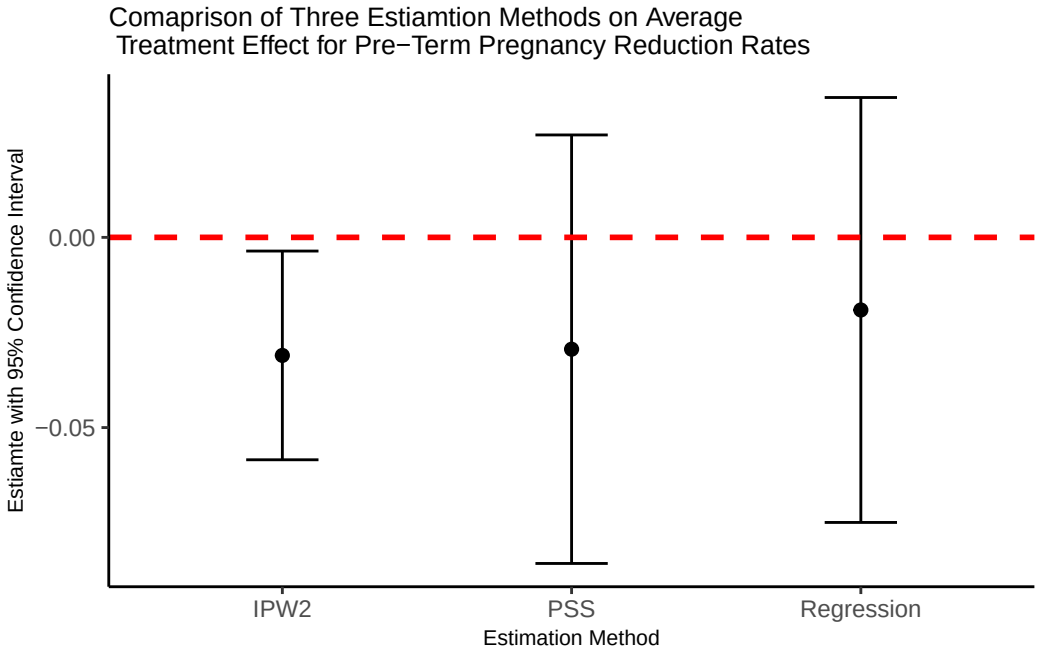

Figure 3: 95% Intervals that capture a red line imply that the estimate is not statistically different from zero

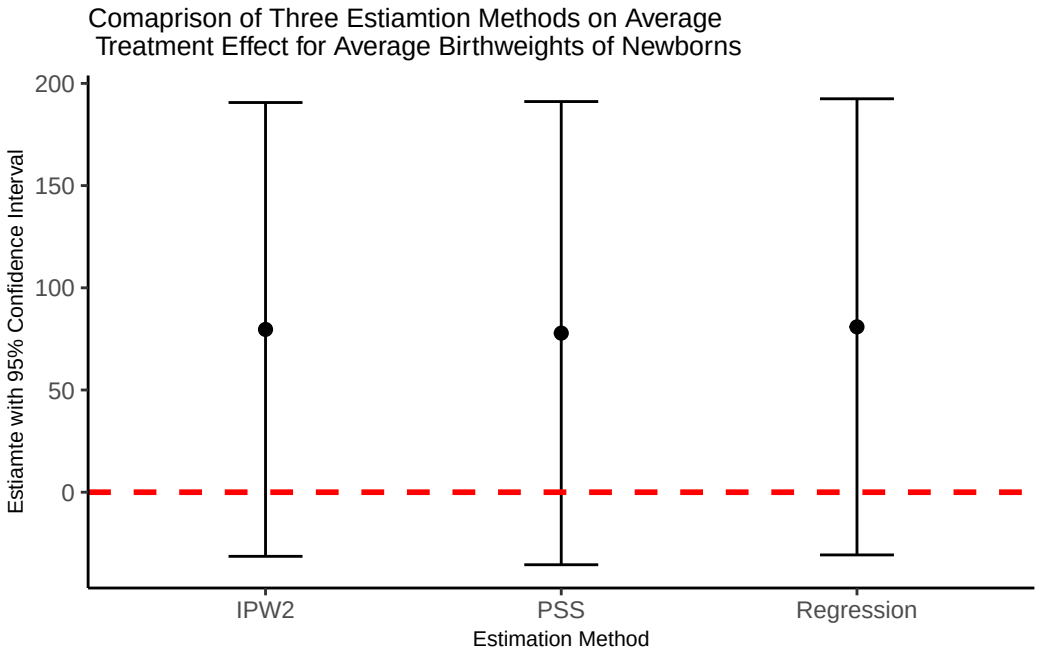

Figure 4: 95% Intervals that capture a red line imply that the estimate is not statistically different from zero