Section 5: More practice with assignment mechanisms (Solutions)

Advanced Quantitative Methods (PLSC 504)

Fall 2017

Example 1: A "broken" RCT?

Suppose you run a randomized experiment on a finite population of $N = 3$ units[1] and $\mathbf{Y}(1) = \mathbf{Y}(0) = (1, 3, 9)$ so that $\tau_i = 0 \; \forall \; i$. Consider the following assignment mechanism,

$$f(\mathbf{W}) = \begin{cases} 1/4 & : \mathbf{W} = (1,0,0) \\ 1/4 & : \mathbf{W} = (0,1,0) \\ 1/4 & : \mathbf{W} = (0,0,1) \\ 1/4 & : \mathbf{W} = (0,1,1) \\ 0 & : \text{otherwise} \end{cases}$$

After randomly assigning units to treatment and control, you observe $\mathbf{Y}$ and $\mathbf{W}$. Let's compare two different estimators for the Average Treatment Effect (ATE).

Estimator A: Difference-in-Means

Let

$$\hat{\tau}_{DM} = \frac{1}{m}\sum_{i=1}^{N} Y_i W_i - \frac{1}{N-m}\sum_{i=1}^{N} Y_i(1 - W_i)$$

denote the Difference-in-Means estimator for the ATE, where m units are assigned treatment and $N - m$ units are assigned control. Is $\hat{\tau}_{DM}$ unbiased? What is $\text{Var}(\hat{\tau}_{DM})$? What is $\text{MSE}(\hat{\tau}_{DM})$?

Solution:

$\hat{\tau}_{DM}$ is unbiased if $\mathbb{E}[\hat{\tau}_{DM}] = 0$. We have four possible realizations of $\mathbf{W}$ here. Each has probability 1/4. For example, if the first unit is selected for treatment we see $Y_1(1) = 1, Y_2(0) = 3, Y_3(0) = 9$ and the DM estimate is $1 - \frac{3+9}{2} = -5$. This logic carries through to the three other possible realizations of $\mathbf{W}$. Therefore,

$$f(\hat{\tau}_{DM}) = \begin{cases} 1/4 & : -5 \\ 1/4 & : -2 \\ 1/4 & : 7 \\ 1/4 & : 5 \\ 0 & : \text{otherwise} \end{cases}$$

So $\mathbb{E}[\hat{\tau}_{DM}] = \frac{1}{4}[-5 - 2 + 7 + 5] = \frac{5}{4}$ is biased.

$$\begin{aligned} \text{Var}\left(\hat{\tau}_{DM}\right) &= \mathbb{E}[\hat{\tau}_{DM}^2] - \mathbb{E}[\hat{\tau}_{DM}]^2 \\ &= \frac{1}{4}\left[(-5)^2 + (-2)^2 + (7)^2 + (5)^2\right] - (5/4)^2 \\ &= \frac{412}{16} - \frac{25}{16} = \frac{387}{16} \approx 24 \end{aligned}$$

[1] This example is based on a homework problem from Peter Aronow's Spring 2015 version of PLSC 503

$$
\begin{aligned}
\mathrm{MSE}(\hat{\tau}_{DM}) &= \mathrm{Var}(\hat{\tau}_{DM}) + \mathrm{Bias}^2(\hat{\tau}_{DM}) \\
&= \frac{387}{16} + \left(\frac{5}{4}\right)^2 = \frac{412}{16} \approx 26
\end{aligned}
$$

Let's check our calculations in R:

```
# Simplification since Y(0)=Y(1)
Y <- c(1, 3, 9)

# Diff-in-Means:
est_dm <- function(Y, W) {
  m <- sum(W)
  N <- length(Y)
  (1/m)*sum(Y*W) - (1/(N-m))*sum(Y*(1-W))
}

dm1 <- est_dm(Y = Y, W = c(1, 0, 0))
dm2 <- est_dm(Y = Y, W = c(0, 1, 0))
dm3 <- est_dm(Y = Y, W = c(0, 0, 1))
dm4 <- est_dm(Y = Y, W = c(0, 1, 1))

# Bias:
dm <- (1/4)*(dm1 + dm2 + dm3 + dm4)
dm - 0
```

```
## [1] 1.25
```

```
# Variance:
var_dm <- (1/4)*(dm1^2 + dm2^2 + dm3^2 + dm4^2) - dm^2
var_dm
```

```
## [1] 24.1875
```

```
# MSE:
mse_dm <- var_dm + (dm-0)^2
mse_dm
```

```
## [1] 25.75
```

Estimator B: Horvitz-Thompson

Let

$$\hat{\tau}_{HT} = \frac{1}{N}\sum_{i=1}^{N}\frac{Y_i W_i}{p_i} - \frac{1}{N}\sum_{i=1}^{N}\frac{Y_i(1-W_i)}{(1-p_i)}$$

denote the Horvitz-Thompson estimator for the ATE, where p_i is the marginal probability of treatment for unit i. Is $\hat{\tau}_{HT}$ unbiased? What is $\mathrm{Var}(\hat{\tau}_{HT})$? What is $\mathrm{MSE}(\hat{\tau}_{HT})$?

Solution:

As before, we have 4 possible realizations of $\mathbf{W}$, each with probability 1/4. The marginal probabilities of assignment are $p_1 = \frac{1}{4}$, $p_2 = p_3 = \frac{1}{2}$. For example, under $\mathbf{W} = (0, 1, 1)$ the HT estimate is

$$\frac{1}{3}\left[\frac{Y_2(1)}{0.5} + \frac{Y_3(1)}{0.5} - \frac{Y_1(0)}{(1-0.25)}\right] = \frac{1}{3}\left[\frac{3}{0.5} + \frac{9}{0.5} - \frac{1}{0.75}\right] = \frac{1}{3}\left[\frac{72}{3} - \frac{4}{3}\right] = \frac{68}{9}$$

Carrying through this logic to the other scenarios yields,

$$f(\hat{\tau}_{HT}) = \begin{cases} 1/4 & : -\frac{60}{9} \\ 1/4 & : -\frac{40}{9} \\ 1/4 & : \frac{32}{9} \\ 1/4 & : \frac{68}{9} \\ 0 & : \text{otherwise} \end{cases}$$

So $\mathbb{E}[\hat{\tau}_{HT}] = \frac{1}{4}\left[-\frac{60}{9} - \frac{40}{9} + \frac{32}{9} + \frac{68}{9}\right] = 0$ is unbiased.

$$\begin{aligned}
\mathrm{Var}\left(\hat{\tau}_{HT}\right) &= \mathbb{E}[\hat{\tau}_{HT}^2] - \mathbb{E}[\hat{\tau}_{HT}]^2 \\
&= \mathbb{E}[\hat{\tau}_{HT}^2] \text{ (since } \mathbb{E}[\hat{\tau}_{HT}] = 0) \\
&= \frac{1}{4}\left[(-60/9)^2 + (-40/9)^2 + (32/9)^2 + (68/9)^2\right] \\
&\approx 33
\end{aligned}$$

$$\begin{aligned}
\mathrm{MSE}(\hat{\tau}_{HT}) &= \mathrm{Var}(\hat{\tau}_{HT}) + \mathrm{Bias}^2(\hat{\tau}_{HT}) \\
&= \mathrm{Var}(\hat{\tau}_{HT}) \quad \text{(since } \mathbb{E}[\hat{\tau}_{HT}] = 0) \\
&\approx 33
\end{aligned}$$

Let's check our calculations in R:

```r
# Simplification since Y(0)=Y(1)
Y <- c(1, 3, 9)

# Assignment probabilities
p <- c(1/4, 1/2, 1/2)

# Horvitz-Thompson
est_ht <- function(Y, W, p) {
  N <- length(Y)
  (1/N)*(sum(Y*W/p) - sum(Y*(1-W)/(1-p)))
}

ht1 <- est_ht(Y = Y, W = c(1, 0, 0), p = p)
ht2 <- est_ht(Y = Y, W = c(0, 1, 0), p = p)
ht3 <- est_ht(Y = Y, W = c(0, 0, 1), p = p)
ht4 <- est_ht(Y = Y, W = c(0, 1, 1), p = p)

# Bias:
ht <- (1/4)*(ht1 + ht2 + ht3 + ht4)
ht - 0
```

```
## [1] 0
```

```r
# Variance:
var_ht <- (1/4)*(ht1^2 + ht2^2 + ht3^2 + ht4^2) - ht^2
var_ht
```

```
## [1] 33.48148
```

```r
# MSE:
mse_ht <- var_ht + (ht-0)^2
mse_ht
```

```
## [1] 33.48148
```

Example 2: an experiment with $N = 1$?

In lecture Fredrik talked about a scenario where we had a single experimental unit who was assigned to treatment or control with equal probability. That is, we only get to see one of the unit's potential outcomes. Formalize the assignment mechanism and show that $\hat{\tau}_{HT}$ is unbiased.

Solution:

Treatment is randomly assigned to a single unit *once*, so the assignment mechanism is,

$$f(W) = \begin{cases} 1/2 & : \text{W} = 0 \\ 1/2 & : \text{W} = 1 \\ 0 & : \text{otherwise} \end{cases}$$

There are two possible realizations. Therefore,

$$f(\hat{\tau}_{HT}) = \begin{cases} 1/2 & : \frac{Y_1(1)}{0.5} \\ 1/2 & : -\frac{Y_1(0)}{0.5} \\ 0 & : \text{otherwise} \end{cases}$$

So $\mathbb{E}[\hat{\tau}_{HT}] = \frac{1}{2}\left[2Y_1(1) - 2Y_1(0)\right] = \left[Y_1(1) - Y_1(0)\right]$ is unbiased.