Notes for Wednesday March 3, 2010

Irregular Assignment Mechanism

We consider a special case of irregular assignment mechanism, using instrumental variable to evaluate causal inference in a randomized trial with noncompliance.

Example: Trial - vitamin A on infant mortality

The treatment (vitamin A) is randomly assigned. However not all assigned to receive treatment (Z=1) actually do receive treatment (W).

Assignment Z	Trt received W	Outcome Y	# children	
0	0	0	74	
0	0	1	11514	.9936
1	0	0	34	
1	0	1	2385	.9859
1	1	0	12	
1	1	1	9663	.9988

Some possible estimates that we can construct

- ITT: intention-to-treat analysis looks at the effect of Z on Y

  Compare Z = 1 to Z = 0, on Y

  $$ITT_{z\ on\ y} = \frac{12048}{12094} - \frac{11514}{11588} = 0.026$$

- As treated
  - Compare folks with W = 1 to those with W = 0, on Y

    $$\frac{9663}{9675} - \frac{13899}{14007} = 0.0065$$

- Per protocol – examine folks who got the intended treatment
  - w = z = 1 vs. w = z = 0, on Y

    .9988 - .9936 = 0.052

- ITT effect of Z on W (not directly of interest but useful later
  - ITT_z on w = 9675/12094 – 0/11588 = 0.80

Extending the potential outcome notation for this case

Z_i = assignment $\in \{0, 1\}$

W_i = treatment received

Wi (0) = treatment received if assigned $Z_i = 0$

Wi (1) = treatment received if assigned $Z_i = 1$

$W_i^{obs} = W_i (Z_i)$

Y_i = outcome = $Y_i (Z_i, W_i)$

There are four possible conditions.

Y_i (1, 1) = assign treatment, receive treatment $\rightarrow$ compliance

Y_i (1, 0) = assign treatment, don't receive treatment $\rightarrow$ non - compliance

Y_i (0, 0) = not assign treatment, don't receive treatment $\rightarrow$ compliance

Y_i (0, 1) = not assign treatment, receive treatment $\rightarrow$ non - compliance

We can only observe Y_i (1, Wi (1)) and Y_i (0, W_i (0))

$\Rightarrow Y_i^{obs} = Y_i(Z_i, W_i^{obs}) = Y_i (Z_i, W_i (Z_i))$

So observed data are: Z_i, W_i^{obs}, Y_i^{obs}

Intention – to – Treat Analysis

(1) We can estimate causal effect of Z on W and Z on Y because Z is randomized!

- Z on W

$$ITT_w = \frac{1}{N}\sum_{i=1}^{N} Wi(1) - Wi(0) = \frac{1}{N}\sum_{i=1}^{N} Wi(1)$$ (since Wi (0) = 0 in our example)

So $\hat{ITT}_w = \overline{W}_1^{obs} - \overline{W}_0^{obs}$ (Neyman's approach for randomized exp.)

$= \overline{W}_1^{obs}$ (in our example)

- Z on Y

$$ITT_y = \frac{1}{N}\sum_{i=1}^{N} Y_i(1, W_i(1)) - Y_i(0, W_i(0))$$

So $\widehat{ITT}_y = \overline{Y}_{1.} - \overline{Y}_{0.}$

Some notation for the formulas:

N_{zw} = # people with Z = z and W = w

$N_{z.}$ = # people assigned Z = z

$N_{.w}$ = # treated with W = w

$$\overline{Y}_{z.} = \frac{1}{N_{z.}} \sum_{i=1}^{N} I_{z_i = z} Y_i^{obs}$$

$$\overline{Y}_{.w} = \frac{1}{N_{.w}} \sum_{i=1}^{N} I_{w_i = w} Y_i^{obs}$$

$$\overline{W}_{z.} = \frac{1}{N_{z.}} \sum_{i=1}^{N} I_{z_i = z} W_i^{obs}$$

(2) Effect of W on Y

If we are interested in the effect of W on Y (we often are), then there is a problem in that assignment of W is not randomized or even necessarily un-confounded.

The specific concern is that we don't know about the Z = 1, W = 0, and Z = 0, W = 0 groups, we don't know whether these two groups are the same or not. Indeed we suspect they are not.

Out next approach is to hope that we can make it unconfounded by conditioning on covariates but we have none to condition on. In this case an analogous idea is to try and control for compliance.

- **Define compliance status**

$$C_i = \begin{cases} \text{c: compliance} \\ \text{n: non-compliance} \end{cases}$$

N_c = # compliers

N_n = # non-compliers (We assume never-takers)

$\pi_c = N_c/N$ = proportions of compliers

$\pi_n = 1 - \pi_c$ = proportions of non-compliers

C_i doesn't depend on Z_i, but "observed compliance" does depend on Z_i.

If Z_i = 0 we don't know whether you would have complied or not.

- **Principal Stratification -** stratification on a post-treatment variable (usually a no – no) by considering values under all treatments

*** Create strata**

Ci = c, $W_i(1) = 1$

$W_i(0) = 0$

Ci = n, $W_i(1) = 0$

$W_i(0) = 0$

It is not the same as looking at observed compliance

Non-compliance: $Z_i = 1, W_i = 0$

Compliance: $Z_i = 1, W_i = 1$

In the $Z_i = 0$, $W_i = 0$ group we have compliers and non-compliance (can't tell who is who)

Table from Little, Long, Lin

		Compliance		
		C	N	
Treatment assignment	Z = 0 Z = 1 Total	$(1-\alpha)\pi_c$ $\alpha\pi_c$ π_c	$(1-\alpha)\pi_n$ $\alpha\pi_n$ π_n	$1-\alpha$ A
Pop. Mean of Y	Z = 0 Z = 1 Total	μ_{0c} μ_{1c} μ_{+c}	μ_{0n} μ_{1n} μ_{+n}	μ_{0+} μ_{1+}
Observed data	Z = 0 Z = 1	$\bar{y}_{oc}$ $\bar{y}_{1c}$	$\bar{y}_{on}$ $\bar{y}_{1n}$	

Note: We can't estimate μ_{0c} and μ_{0n} directly. For observed data, we can only observe $\bar{y}_o$ which is combined value of $\bar{y}_{oc}$ and $\bar{y}_{on}$.

*** Estimate causal effect**

Our approach is to examine the causal effect for subgroups defined by compliance

Examine ITT effect separately for compliance categories.

Z on W

$$ITT_{w.n} = \frac{1}{N_n}\sum_{c_i=n} W_i(1) - W_i(0) = 0$$

$$ITT_{w.c} = \frac{1}{N_c}\sum_{c_i=c} W_i(1) - W_i(0) = 1$$

$$\rightarrow ITT_w = \pi_c ITT_{w.c} + \pi_n ITT_{w.n} = \pi_c$$

Z on Y

$$ITT_{y.n} = \frac{1}{N_n}\sum_{c_i=n} Y_i(1,0) - Y_i(0,0)$$

$$ITT_{y.c} = \frac{1}{N_c}\sum_{c_i=c} Y_i(1,1) - Y_i(0,0)$$

$$\rightarrow ITT_y = \pi_c ITT_{y.c} + \pi_n ITT_{y.n}$$

*** Key assumption - exclusion restriction for non-compliers**

$Y_i(0, 0) = Y_i(1, 0)$, for $C_i = n$

- We assume Z has no direct effect on non-compliers; if you don't get the treatment you have the same outcome whether you weren't assigned or were assigned
- It makes sense here, but not always, as shown in an example.

Example: Z = treatment assignment to exercise for weight loss

Even if $W_i = 0$, you don't do exercise, but you might change some other behavior.

Under exclusion restriction:

$$ITT_{y,n} = 0 \rightarrow ITT_y = ITT_{y,c}\pi_c$$

$$\Rightarrow ITT_{y,c} = \tau_{iv} = \frac{ITT_y}{\pi_c} = \frac{ITT_y}{ITT_w}$$

Instrumental variable estimand $\hat{\tau}_{iv} = \frac{I\hat{T}T_y}{I\hat{T}T_w} = \frac{0.0026}{0.8} = 0.0032$

One way to interpret this result is that the benefit of treatment goes to the group who comply. The true effect 0.0032, but we only observe 0.0026 because not everybody complies.