POLS 4368 - Fall 2014 Midterm Solution Set

Professor Donald P. Green
TA Alex Coppock

April 6, 2014

This exam reviews concepts from FEDAI Chapters 1-6. At the conclusion of this exam, please turn in this sheet along with your blue exam booklet.

1 Section I

Briefly define and state the significance of the following terms or phrases. Use formal notation to make your definitions as clear as possible. (9 points each)

1. Randomization inference
   Answer:
   Randomization inference is procedure for constructing 95% confidence intervals and conducting hypothesis tests. Subjects reveal one of two potential outcomes depending on their treatment assignment: treated subjects reveal their $Y_i(1)$ and untreated subjects reveal their $Y_i(0)$. Randomization inference proceeds by hypothesing what the *other* potential outcome is for each subject. When constructing 95% condifence, we fill in the missing potential outcome using the estimated treatment effect – subtracting it off for treated subjects and adding back for untreated subjects. We can then generate the sampling distribution of the experiment under all possible random assignments – that is, under the hypothesis of constant effects (equal to the estimated ATE), all the ways the experiment could have turned out. Each possible random assignment generates a simulated ATE – the 95% confidence interval is the range extending from the 2.5^th percentile to the 97.5^th.

   Hypothesis tests conducted using randomization inference, on the other hand, are carried out by supposing the sharp null hypothesis of no effect, i.e., that there is exactly zero treatment effect for all subjects. The sharp null implies that the missing potential outcome is the same as the observed potential outcome. We can then generate the sampling distribution under the sharp null by considering the estimated ATEs under all possible random assignments. The p-value we assign to the estimated ate is the fraction of simulated ATEs that equal or exceed the observed ATE. For two-tailed tests, we calculated the fraction of simulated ATEs that equal or exceed the observed ATE in absolute value.

2. Covariate adjustment
   Answer:

   Covariate adjustment is a statisical procedure employed to dampen sampling variability. When researchers adjust for pre-treatment covariates that are prognostic of the outcome, the sampling distribution around the estimated ATE is tighter. Adjusting for covariates is method

of rescaling the dependent variable so we can focus on the variation induced by treatment, rather than the variation that is correlated with background attributes. Covariate adjustment can also help to provide greater balance on observable characteriscs that may be slightly imbalanced across treatment and control due to chance. Covariate adjustment does introduce Freeman bias (the average of the covariate-adjusted estimates over all possible randomizations is not exactly equal to the true ATE), the size of this bias is small and diminishes rapidly as N rises above 20. An important point to underscore in the discussion of covariates is that the estimated coefficients on the covariates should never be given any causal interpretation – they included only to reduce sampling variability. Finally, researchers should pre-register which covariates they intend to use so as to reduce the researcher discretion and suspicion of fishing.

3. Clustered random assignment
Answer:
When treatment assignment occurs at the group or cluster level, it is refered to as clustered random assignment. A canonical example is experiments in schools – treatment is often assigned at the classroom level rather than at the school level. Clustered assignment, relative to individual-level treatment assignment, often increases sampling variability. The worst case is when all the individuals within a cluster have similar potential outcomes, but potential outcomes across clusters vary widely. In this case, the effective N of the experiment is closer to the number of clusters rather than the number of individuals. When potential outcomes are more homogenous across clusters, the effective N is higher.

Another trouble with cluster random assignment when clusters are of unequal size is that the difference-in-means estimator of the ATE is biased. This is because the denominator in each group is a random variable (i.e., the number of subjects in the treatment group varies depending on which clusters happen to be treated.). The expectation of a ratio is not, in general, the ratio of the expectations. There are two recommendations for addressing this problem. The first is design-based: perform block random assignment by cluster size – obviating the changing denominator problem. The second is an analytic recommendation to use the Horvitz-Thompson difference-in-totals estimator, which is less efficient than difference-in-means, but is unbiased even in the presence of clusters of unequal size.

4. The assumption of monotonicity in the context of two-sided noncompliance Answer:
In experiments facing two-sided non-compliance, subjects can be assigned to treatment or control ($Z = 1$ or $Z = 0$), and they can either take treatment or not ($D = 1$ or $D = 0$). This creates four types: Always-takers ($D(Z = 1) = D(Z = 0) = 1$), Compliers ($D(Z = 1) = 1$; $D(Z = 0) = 0$), Never-takers ($D(Z = 1) = D(Z = 0) = 0$), and Defiers ($D(Z = 1) = 0$; $D(Z = 0) = 1$). Formally, the assumption of monotonicity can be stated in the following manner: $D(Z = 1) \geq D(Z = 0)$. This restriction rules out one of the four types: the defiers. This assumption is crucial for the identication of the Complier Average Causal Effect using experimental data – without it, we can't identify the average effects of treatment for any group.

2 Section II

The following table was presented in Chapter 5. The results refer to the New Haven voter mobilization experiment, in which a random subset of the subject pool was assigned to be canvassed, but only some of those assigned to be canvassed were actually canvassed. The outcome is voter turnout. (8 points each)

Voter turnout by experimental group, New Haven voter mobilization experiment

	Treatment Group	Control Group
Turnout rate among those contacted by canvassers	54.43 (395)	
Turnout rate among those not contacted by canvassers	36.48 (1,050)	37.54 (5,645)
Overall turnout rate	41.38 (1,445)	37.54 (5,645)

Note: Entries are percent voting, with number of observations in parentheses.
Sample restricted to households containing a single registered voter.

1. Define a "Complier."
   Answer:
   A complier is a subject who takes treatment if and only assigned to the treatment group. In this case, a complier is a subject who opens the door to a canvasser if and only if assigned to the treatment group.

2. Estimate the proportion of Compliers in the subject pool.

```
pr.compliers <- 395/1445
pr.compliers
```

```
## [1] 0.2734
```

   The proportion of compliers is 0.273.

3. Show (with algebra) that under the assumption of non-interference and excludability, the CACE is identified in this application.

The CACE is defined as $E(Y_i(1) - Y_i(0)|D_i(1) = 1)$.

- Expected value of voting rate (Y) in the control group $=$
  $E(Y_i(0)|D_i(1) = 1) * ITT_d + E(Y_i(0)|D_i(1) = 0) * (1 - ITT_d)$

- Expected value of voting rate in treatment group $=$
  $E(Y_i(1)|D_i(1) = 1) * ITT_d + E(Y_i(0)|D_i(1) = 0) * (1 - ITT_d)$

- Expected value of rate of successful canvassing $= E(D_i(1)) = ITT_d$

- Expected value of voting rate among treatment group minus voting rate of control group
  $= (E(Y_i(1)|D_i(1) = 1) - E(Y_i(0)|D_i(1) = 1)) * (ITT_d) = CACE * ITT_d.$

- $CACE = (E(Y_i(1)|D_i(1) = 1) - E(Y_i(0)|D_i(1) = 1))/(ITT_d)$.

4. Are non-interference and excludability plausible in this example?
   Answer:
   Non-interference requires that the subjects respond only to their own treatment assignment, and not to the treatment assignment of others. This would be violated if, perhaps, neighbors called each other after being canvassed and talked about voting. The possibility of interference was explored experimentally by Sinclair, McConnell, and Green (2012) – Having treated neighbors does not appear to increase turnout. Excludability requires that nothing about assignment to canvassing itself affected potential outcomes, only the canvassing itself. This is plausible in this experiment.

5. Estimate (by hand) the CACE. Provide a substantive interpretation of your estimate.

```
itt <- .4138 - .3754
ittd <- 395/1445
cace <- itt/ittd
cace


## [1] 0.1405
```

The $\widehat{CACE}$ is 0.14, meaning that compliers are 14 percentage points more likely to vote as a result of canvassing.

3   Section III

A recent experiment tested the effects of sending registered voters in New York City a (nonpartisan) postcard that encouraged them to donate to the candidate of their choice, on the grounds that small donations keep elected officials focused on important policy issues. Randomization was conducted within each of 5 blocks; blocks were created based on voters' expected probability of donating to campaigns in the future. Below is a block-by-block summary of the results. The first table displays the distribution of donations in control and treatment, by block. The next table displays the results of regressions, by block, of donations on treatment. The third table displays the results of inverse-probability weighted regression and the accompanying p-values derived from randomization inference under the sharp null hypothesis of no effect. The last table presents the results of an unweighted regression that controls for blocks. Based on your reading of the tables, answer the following questions: [8 points each]

1. Why would an (unweighted) regression of donations on treatment, ignoring blocking, be a biased estimator of the average treatment effect?
   Answer:
   Because the blocks are of unequal size, the probabilities of assignment to treatment differ by block. If treatment effects vary by block, then an unweighted regression would be biased because of the induced correlation between potential outcomes and treatment assignment.

2. Interpret the third table's estimate of the ATE and the accompanying p-value.
   Answer:
   The treatment postcards raised the amount individuals donated to campaigns by $3.82. This

result is unlikely under the null model that the postcards had zero effect for all units – only 3.17% of simulations under the null returned difference-in-means estimates as large or larger than $3.82 in absolute value.

3. The block-by-block regressions seem to suggest that the postcard has a significantly negative effect on donations in blocks 1 and 4. The p-values based on robust standard errors are below 0.05. Yet the overall estimate of the ATE using inverse-probability weights is positive. What do you think accounts for this apparent discrepancy?
Answer: Treatment effects appear to vary by block – there is nothing about a positive average effect that precludes the treatment from having a negative effect among some individuals.

4. Bonus: The weighted regression presented in table 3 produces an estimated ATE that is different from an unweighted regression (presented in table 4) that controls for blocks by using a dummy variable for each block (except one, which is the intercept). What, specifically, about the two regressions causes them to produce different estimates? (3 points)
Answer:
The two regressions use different weighting schemes. The IPW regression uses inverse probability weighting to stich together the blocks of the experiment. The regression in table 4 weights each block proportionally to the precision with which the treatment effect is estimated. The two answers are similar because, by and large, as N increases, so does the precision of the estimate. The estimands are also different. Table three presents an estimate of the average individual treatment effect. Table four presents an average of the block-level average treatment effects.

Table 1: Distribution of Donations, by Experimental Group

block	control_mean	control_sd	control_n	treatment_mean	treatment_sd	treatment_n
1	0.58	38.78	62315	0.00	0.00	300
2	0.80	41.89	30065	16.67	288.68	300
3	1.12	48.29	14788	0.83	14.43	300
4	2.03	73.16	7727	0.25	4.33	300
5	0.78	14.66	4198	8.42	144.34	300

Table 2: Block by Block regressions

| | Amount Donated | | | | |
	Block 1	Block 2	Block 3	Block 4	Block 5
Treatment	−0.584***	15.860	−0.287	−1.781**	7.632
	(0.155)	(16.640)	(0.922)	(0.869)	(8.324)
Constant	0.584***	0.804***	1.121***	2.031**	0.784***
	(0.155)	(0.242)	(0.397)	(0.832)	(0.226)
N	62,615	30,365	15,088	8,027	4,498
R^2	0.00000	0.001	0.00000	0.00002	0.002

*p < .1; **p < .05; ***p < .01
Robust Standard Errors in Parentheses

Table 3: Randomization Inference on IPW estimate

Estimate from inverse probability weighted regression	3.8211
Two-sided p-value from randomization inference	0.0317

Table 4: Unweighted Regression Controlling for Blocks

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.5612	0.1558	3.6015	0.0003
Z	4.1659	3.7597	1.1080	0.2678
as.factor(block)2	0.3585	0.3209	1.1173	0.2638
as.factor(block)3	0.4708	0.4238	1.1108	0.2666
as.factor(block)4	1.2477	0.8301	1.5030	0.1328
as.factor(block)5	0.4542	0.5670	0.8010	0.4231