Green // Experimental Research // 2014

# Final Exam

1. Identification of Terms. Briefly define and state the significance of the following terms or phrases.
2. Complete random assignment

Let *0 < m < N*. Under complete random assignment, exactly *m* of *N* subjects are assigned to the treatment group. The use of complete random assignment ensures researchers that a fixed number of subjects will be placed in treatment and control (which makes budgeting and power calculations more predictable). Complete random assignment also affects how randomization inference is conducted, insofar as the randomization distribution is conditioned on a fixed *m*.

1. Trimming bounds vs. extreme value bounds

Trimming bounds: Given an additional monotonicity assumption, estimate lower and upper bound for the ATE for Always Reporters (see pp. 227-­‐8, FEDAI).  Extreme value bounds: Estimate lower and upper extreme value bounds under assumptions that the missing values in the treatment and control group are imputed using the max (min) and min (max) values of the outcome variable (see pp. 226-­‐7, FEDAI).

1. Within-subjects design

In contrast to a between-subjects design, in which different subjects are randomly allocated to treatment and control groups, a within-subjects design tracks the same subject over time, and assigns certain periods to treatment or control. The fact that the same individual is followed as he/she/it travels from a treated to an untreated state (or vice versa) poses a potential threat to the non-interference assumption.

The non-interference assumption states that a subject’s potential outcomes are unaffected by the treatment assignment of any other subject, or Yi(d) = Yi(d). (pp. 253, FEDAI)

Non-interference may be jeopardized if there is “anticipation” (i.e., my potential outcomes in the present period are affected by my future treatment status) or by “carry-over” (i.e., my current treatment status affects my potential outcomes in future periods.)

In the context of a within-subjects design, random assignment’s role is fairly limited, as it merely ensures that the timing of the treatment is independent of potential outcomes. The main threat to inference is carry-over or anticipation. To some extent, these concerns may be allayed by designs that blind subjects (mitigating anticipation) or allow for a wash-out period (mitigating spillover).

1. Mediation

Mediation refers to the process by which causal effects are transmitted from a treatment to an outcome. For example, the relationship between consumption of limes and the occurrence of scurvy is mediated by the amount of vitamin C in the bloodstream. Mediation poses fundamental challenges to causal inference. Let M(d) be a potential outcome of the mediator when d=0 (control) or d=1 (treatment). When we speak of the effect of the (unmediated) direct effect of a treatment on an outcome (i.e., the effect of D, holding M(d) constant), we are in effect trying to estimate empirically a quantity such as:

E[Y(M(1),1) – Y(M(1),0)] or E[Y(M(0),1) – Y(M(0),0)],

but the second and third potential outcomes in the line above are fundamentally unobservable. We cannot observe, for example, how Y would respond if a subject were untreated and yet M took on the value that follows from the administration of the treatment. The same problem also applies to average mediation effects (i.e., the effect of varying M in response to treatment while holding D constant):

E[Y(M(1),1) – Y(M(0),1)] or E[Y(M(1),0) – Y(M(0),0)].

Again, the second and third terms are complex counterfactuals that are inherently unobservable. An important special case occurs when M(1) = M(0) for all observations: in that case, it follows from the above equation that the average causal mediation effect must be zero. Empirically, this result implies that we can rule out mediators when the sharp null hypothesis of no effect of the treatment on the mediator seems to hold.

An analysis NOT justified by randomized experimental design is a regression of Y on M and the treatment. Such a regression is prone to post-treatment bias because unobserved causes of M may be related to unobserved causes of Y. Not only is such a regression prone to bias, it also apportions the “total” effect of the treatment on the outcome into “direct” and “indirect” components under the special case where the effects of the treatment and the mediator are constant across subjects.

1. Short answer
2. Briefly summarize the implications of clustered random assignment for experimental design and analysis.

**Implications for analysis:**

If cluster sizes are equal, then difference-in-means is an unbiased estimator of the ATE. If cluster sizes are unequal, then the difference-in-means estimator is biased because the denominator is now a random variable, and the ratio of an expectation is not equal to the expectation of a ratio. If the number of clusters is relatively small, use difference-in-totals instead (Eq, 3.24, pp. 83). Better yet, in the design phase, block on cluster size before random assignment.

We cannot use standard methods to estimate uncertainty: since the effective N is smaller (due to randomization at the cluster level), the sampling variability typically increases in the variability of the cluster means (pp. 82).

The true SE of the estimated ATE, assuming fixed cluster means, is described by Equation 3.22 (pp. 82). This quantity cannot be identified because we cannot observe the covariance term, so we estimate standard errors using Equation 3.23 (pp. 83).

We can form confidence intervals by creating a schedule of potential outcomes under the assumption of constant treatment effects (τi = ATEhat); we also apply a degrees of freedom adjustment that expands the width of the interval by the square root of [(k-1)/(k‐2)] where k is the number of clusters (pp. 83, including footnote 20).

**Implications for design (based on Eq 3.22 and extensions to the discussion in pp. 57-59):**We can decrease the SE (i.e. shrink the sampling distribution) by:

(1) increasing the number of clusters;

(2) placing greater number of clusters into the treatment group where cluster-level means of potential outcomes associated with that group have a higher  variance, but if unknown use a balanced design;

(3) examine treatments that minimize (or have negative) covariance between  average treated potential outcomes and average untreated potential  outcomes at the cluster level;

(4) minimize the cluster level mean treated potential outcome variance and the cluster level mean untreated potential outcome variance.

(5) block on cluster size to avoid the issue of bias in difference-in-means

1. Explain (preferably using a bit of algebra) why rejecting the null hypothesis that implies rejection of the null hypothesis of homogeneous treatment effects (i.e., ).

See pp. 293, FEDAI.

Var(Yi(1))

=  Var(Yi(0) + τi)

=  Var(Yi(0)) + Var(τi) + 2 Cov( Yi(0), τi ) and the equality Var(Yi(1)) = Var(Yi(0)) holds when  Var(τi) = -2 Cov ( Yi(0), τi )

Under the null hypothesis that τi is constant across subjects, both sides of Equation (9.3) are zero, since the covariance between a variance and a constant is zero. Thus rejecting the hypothesis Var(Yi(1)) = Var(Yi(0)) means rejecting the null hypothesis that Var(τi) = 0.

1. Modeling and data analysis

The table below shows the results of a recent experiment in which 630,640 subjects were randomly sent a “social pressure” mailing immediately prior to an election in June of 2012. The remaining 33,380 subjects were sent nothing. Turnout in that election is indicated by the variable votedS. This variable equals 1 when a subject voted; 0 otherwise. Later that year, a presidential election occurred, and subjects voted or abstained (see the variable votedG).

Suppose you sought to estimate the “downstream” effect of votedS on votedG.

1. Show algebraically how one can identify the average causal effect among those who vote in the June election if and only if they are encouraged by the mailer.
2. Explain and critically evaluate the excludability assumption required to obtain this identification result.
3. Use the results below to estimate this average causal effect. (Don’t worry about estimating standard errors.)

Subjects assigned to the control group

| votedS

votedG | 0 1 | Total

-----------+----------------------+----------

abstained | 7,990 1,275 | 9,265

| 69.96 5.81 | 27.76

-----------+----------------------+----------

voted | 3,431 20,684 | 24,115

| 30.04 94.19 | 72.24

-----------+----------------------+----------

Total | 11,421 21,959 | 33,380

Subjects assigned to the treatment (mail) group

| votedS

votedG | 0 1 | Total

-----------+----------------------+----------

abstained | 147,147 24,721 | 171,868

| 70.46 5.86 | 27.25

-----------+----------------------+----------

voted | 61,691 397,081 | 458,772

| 29.54 94.14 | 72.75

-----------+----------------------+----------

Total | 208,838 421,802 | 630,640

The downstream effect of **votedS** on **votedG** may be identified by instrumenting for **votedS** using random assignment to the mailer.

Let Z = random assignment to mailer D = votedS = voted in spring election Y = votedG = voted in general election





Critically evaluate exclusion restriction assumption: The assumption would be violated if the mailer affected voting in the general election via some causal pathway other than voting in the spring. This is likely, for instance, if the mailer were so memorable that people were influenced by it when deciding whether to vote in November.

1. Interpreting results

Suppose that researchers in Kenya, a society where more than 90% of schoolchildren test positive for (communicable) intestinal parasites, conduct a randomized experiment to assess the effect of de-worming medicine on educational attainment. The dependent variable is whether a child later graduates from high school. The treatment is randomly assigned at two levels: first, from a list of 1000 high schools, 100 high schools are randomly selected for an intervention. In the schools chosen to receive the intervention (which vary in size from 50 to 500 pupils), 25 children will be selected to receive de-worming medicine. Assume that the design is implemented as planned, and that outcomes are observed for all children in all schools.

* 1. Define the relevant potential outcomes (using appropriate notation) and the causal estimands of interest. Explain the substantive meaning of each estimand.

The number of potential outcomes that could be defined for this design is in principle very large, as there are many possible types of second-hand treatment. For example, I could be untreated in a school of 500 students, of whom 25 are treated, or I could be untreated in a school of 50 students, of whom 25 are treated, etc. In order to make the estimation problem theoretically tractable, some simplifying assumptions will be necessary. First, we will impose the assumption that spillovers travel within schools but not across them. Next, we will assume that conditional on their own treatment status, children are affected by whether some of their schoolmates receive treatment. That categorization implies the following potential outcomes: Y(s,d), where s (0,1) indexes whether a school receives treatment (0,1) and d indexes whether the subject himself receives treatment. This analysis may be refined by partitioning schools by size, a pre-treatment covariate.

The relevant estimands then are

E[Y(1,0) – Y(0,0)] 🡪 the average effect, conditional on not receiving the treatment directly, of switching from an untreated school to a treated school. This estimand conveys the protective value to oneself of being around others who are treated in a school.

E[Y(1,1) – Y(1,0)] 🡪 the average effect, conditional on some part of the school receiving the treatment directly, of switching from being untreated personally to being treated personally. This estimand conveys the extra protective value of immunization in schools where some fraction of schoolmates are being treated.

* 1. Explain whether these estimands are identified given the design and whatever assumptions you see fit to invoke. Be sure to make your assumptions explicit.

These quantities are each identified given the randomized design and the assumption of no spillovers across schools and no unmodeled potential outcomes beyond what was stipulated above. These non-interference assumptions imply that potential outcomes are unaffected by how the random assignment happens to come out in any school. This assumption would be jeopardized if children who live nearby attend different schools; presumably, children in control schools are less likely to contract disease if they live nearby children in immunized schools, in which case the expected outcome in untreated schools will be higher (a higher graduation rate) than would be the case if there were no spillovers.

* 1. What special complications does this design pose for estimation of average causal effects and hypothesis testing?

This design poses the following estimation challenges:

1. Subjects have different probabilities of being assigned to treatment. Because schools are assigned irrespective of size, children in small schools have a higher probability of being treated directly than children in large schools. Probabilities of exposure to second-hand treatments also vary by school size. Inverse probability weights are required to account for different assignment probabilities; IPW will be asymptotically unbiased (rather than purely unbiased) because the sum of weights is random in this design.

2. The clustered design presents further complications. Schools are assigned as clusters, which implies, for example, that everyone in a pure control school is assigned together to the same condition. These clusters vary in size, which further implies that difference-in-means estimation is biased (if potential outcomes are related to cluster size), but consistent. Hypothesis testing and estimation of sampling variances must take clustering into account; fortunately, doing so is relatively straightforward using randomization inference.

3. As always, care must be exercised when interpreting any apparent interaction between treatments/spillovers and school size, because the latter is not randomly assigned. In effect, we have a distinct treatment/spillover/control experiment within a given level of school size.