CAUSAL INFERENCE–AN INTRODUCTION

I. CAUSATION AND STATISTICS—HISTORY

A. GALTON, PEARSON, YULE——-OBSERVATIONAL STUDIES
B. NEYMAN, FISHER, RUBIN—-EXPERIMENTAL STUDIES AND EXPERIMENTAL DESIGN

II. CAUSATION

A. HUME–MECHANISM?
B. MILL–FULL CAUSE–CAUSES OF EFFECT AND EFFECTS OF CAUSES
C. COLLINGWOOD–MANIPULATION
D. PROBABILISTIC CAUSATION—BACK TO YULE
E. COUNTERFACTUAL CONDITIONALS

III. UNIT AND AVERAGE CAUSAL EFFECTS–DEFINITIONS

A. NOTATION FOR CAUSAL INFERENCE
B. UNIT EFFECTS AND THE FUNDAMENTAL PROBLEM OF CAUSAL INFERENCE
C. AVERAGE CAUSAL EFFECTS

IV. ESTIMATION OF CAUSAL EFFECTS

A. IGNORABILITY (TREATMENT ASSIGNMENT)
B. CONDITIONAL IGNORABILITY–AN IMPORTANT APPROACH TO ANALYZING NON-EXPERIMENTAL STUDIES

IA. GALTON, PEARSON, YULE—OBSERVATIONAL STUDIES

REGRESSION AND CORRELATION SPURIOUS CORRELATION
PARTIAL CORRELATION
PATH ANALYSIS, SIMULTANEOUS EQUATION MODELS, STRUCTURAL
EQUATION MODELS, TIME SERIES—OFTEN USED TO MAKE CAUSAL
INFERENCES IN SOCIOLOGY, ECONOMICS, EPIDEMIOLOGY, PSYCHOL-
OGY

IB. Neyman, Fisher, Rubin—EXPERIMENTAL DESIGN

POTENTIAL OUTCOMES NOTATION
RANDOMIZATION
ANALYSIS OF VARIANCE

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}, \tag{1}$$

Y_{ij} = OUTCOME for SUBJECT i under CONDITION j, $E(\varepsilon_{ij}) = 0$, $\sum_{j=1}^{J} \alpha_j = 0$.

EXAMPLE: Y = EARNINGS (SALARY)
J = 3; j = 1 if no job training program, j = 2 if standard job training program, j = 3 if new job training program.

IIA. HUME–CAUSATION

Z = CAUSE, Y = EFFECT.
1. Z BEFORE Y
2. Z NEAR Y
3. IF Z, ALWAYS Y

    PROBLEM: Z = NIGHT, Y = DAY.
DOES NIGHT CAUSE DAY?

IIB. MILL–CAUSATION

FULL CAUSE—-Y = f(z(1),...,z(L)).
EFFECTS OF CAUSES VS. CAUSES OF EFFECTS.

IIC. COLLINGWOOD–CAUSATION

THEORETICAL CAUSATION—FULL CAUSE–CAUSES OF EFFECT
APPLIED CAUSATION—DO THIS TO Z, SEE WHAT HAPPENS TO Y—
EFFECTS OF CAUSES

IID. PROBABILISTIC CAUSATION

SIMON, REICHENBACH, SUPPES, GRANGER,SIMS
APPLICATIONS—SOME USES OF REGRESSSION, PATH ANALYSIS, STRUC-
TURAL EQUATIONS, SOME TIME SERIES
1) Z AND Y ARE CORRELATED (NOT INDEPENDENT)
$\Pr(Y = y, Z = z) \neq \Pr(Y = y)\Pr(Z = z)$
2) Z BEFORE Y
DOES Z CAUSE Y? MAYBE THERE IS A VARIABLE (OR GROUP OF
VARIABLES) X, X BEFORE Z, X BEFORE Y:
$\Pr(Y = y, Z = Z \mid X = X) = \Pr(Y = y \mid X = x)\Pr(Z = z \mid X = x)$.
IN THIS CASE, WE WOULD SAY THE RELATIONSHIP BETWEEN Z AND
Y IS NOT CAUSAL, DUE TO X.

    THE PROBLEM WITH THIS IS IT NEVER SAYS WHAT CAUSATION
OR PROBABILISTIC CAUSATION IS—-IT JUST IS A WAY TO INFER
CAUSATION (OR THE ABSENCE OF CAUSATION) WITHOUT EVEN
SAYING WHAT CAUSATION IS—
DOES NOT MAKE SENSE.

IIE. WHAT IS MISSING IN HUME? COUNTERFACTUAL CONDITIONAL

Example: John is sick. He takes medicine and he gets better. Did the medicine CAUSE John to get better?

A. IF John DID NOT TAKE THE MEDICINE, he WOULD NOT have gotten better.

Or

B. IF John DID NOT TAKE THE MEDICINE, he WOULD have gotten better.
    CONDITIONAL —if, then

COUNTERFACTUAL—because John takes medicine.

In case A) we say taking the medicine causes John to get better

In case B), John would get better under either condition—so taking the medicine does not cause John to get better.

SOCIAL SCIENCE EXAMPLES
1. The effect of marriage on happiness.
2. The effect of job training on employment status.
3. The effect of education on earnings (here maybe college versus not–or maybe single years of schooling).
4. The effect of time spent watching tv on obesity (weighing too much)
5. The effect of amount of time studied on exam scores.

IIIA. NOTATION FOR CAUSAL INFERENCE

Denote John by i.

Let $Y_i(1) = 1$ if John gets better when he takes medicine, 0 if he does not get better when he takes the medicine.

Let $Y_i(0) = 1$ if John gets better when he does not take medicine, 0 if he does not get better when he does not take the medicine.

The unit causal effect—$Y_i(1) - Y_i(0)$.

Now let $Z_i = 1$ is John takes the medicine, $Z_i = 0$ if John does not take the medicine.

Of course, the BIG PROBLEM is that we see the data $(Z_i, Y_i(Z_i)) = (Z_i, Z_i Y_i(1) + (1 - Z_i)Y_i(0))$,

BUT we do not see the hypothetical data $(Z_i, Y_i(0), Y_i(1))$. IF we could see these data, we could take a random sample of size n and use:

$$\frac{\sum_{i=1}^{n}(Y_i(1) - Y_i(0)))}{n} \tag{2}$$

to estimate the AVERAGE CAUSAL EFFECT:

$$E(Y(1) - Y(0)) \tag{3}$$

5

BUT we cannot do this. So what usually happens is that the difference between the treatment and control groups is used to estimate (3):

$$\bar{Y}_1 - \bar{Y}_0 = \frac{\sum_{i=1}^{n} Z_i Y_i}{\sum_{i=1}^{n} Z_i} - \frac{\sum_{i=1}^{n} (1 - Z_i) Y_i}{\sum_{i=1}^{n} (1 - Z_i)} \tag{4}$$

But is this a good estimate or not? Let's look at an example that will help us fix ideas and from which we can get some intuition.

Suppose you have a population of size 12. You are going to conduct a randomized experiment to see if treatment with a drug improves subsequent performance. Let $Y_i(0)$ and $Y_i(1)$ denote, respectively, the value of the performance variable for observation i, where $i = 1, ..., 12$. Let X denote the covariate sex, where $X_i = 1$ if unit i is male, 0 otherwise. You now take a random sample from the population, obtaining units 1-8; these are the experimental units– **units 9-12 are not part of the experiment**. Suppose the (hypothetical) data are as follows:

Unit	Y_0	Y_1	Z	X	Z^*
01	0	6	1	0	1
02	2	2	0	0	0
03	3	2	1	0	1
04	2	8	1	0	0
05	2	4	0	1	0
06	4	4	0	1	1
07	5	4	1	1	0
08	2	3	0	1	1
09	4	4	d	0	d
10	6	9	d	0	d
11	9	7	d	1	d
12	6	8	d	1	d

1. What are the unit treatment effects? Does the drug always improve performance? Recall that we cannot actually observe these, however. Why is this, and what do we actually observe?

2. Define the average treatment effect. That is, using potential outcomes notation, write out the expected value of interest and show how we can calculate it using the hypothetical data in the table. What value does it have here?

3. Suppose that you have conducted a completely randomized experiment with half the units to be assigned to each group and that Z (above) is the vector of assignments, with $Z_i = 1$ if unit i is assigned to the treatment group, $Z_i = 0$ if unit i is assigned to the control group; if the unit is not part of the experiment, we denote this by "d". How would you **estimate** the average treatment effect? Give both the general formula and the actual value. Compare this estimate with the actual average treatment effect.

4. Define (just as we did in problem 2) the average treatment effect for men and the average treatment effect for women. What values do these have here? What do we learn here that we did not know from 2)?

5. How would you **estimate** the average treatment effect for men and the average treatment effect for women? What values do these estimates have here?

IGNORABILTIY CONDITIONS:

$Y(0), Y(1) \perp Z$

$Y(0), Y(1) \perp Z \mid X$