

STATISTICS 265 – Winter 2010

EXAM: Wednesday February 24, 2010 (11am-12:20pm)

- The exam is closed book, closed notes. You may use a calculator.
- Please do not write your solutions on the exam paper.
- Good luck!

1. The table below lists two potential outcomes ($Y(0), Y(1)$), where 0 denotes the placebo and 1 denotes the active treatment, for a set of 6 individuals along with an indication of the treatment assignment W for each. Assume that we are interested in an additive effect for the treatment.

Unit	W	Y(0)	Y(1)
1	1	81	93
2	0	76	87
3	1	66	71
4	0	72	80
5	0	65	74
6	1	75	82

- (a) Explain the meaning of $Y(0)$ and $Y(1)$.
 - (b) What is the causal effect of receiving treatment for the first unit?
 - (c) The Stable Unit Treatment Value Assumption (SUTVA) plays a central role in the potential outcomes approach to causal inference. What does this assumption say?
 - (d) What is the average treatment effect in this population of 6 individuals?
 - (e) In a randomized study W is independent of $Y(0)$ and W is independent of $Y(1)$ but W is not independent of Y^{obs} . Explain.
2. Consider a study in which a single covariate X is measured along with the treatment indicator W and outcome variable Y . We propose to analyze the resulting data via a linear regression analysis $Y_i = \beta_0 + \beta_1 X_i + \tau W_i + \epsilon_i$.
- (a) If the study is a randomized experiment, then we don't need to include X in the regression to get a valid estimate of the average treatment effect τ . Briefly explain why this is true.
 - (b) Suppose instead that the study is an observational study with "regular assignment mechanism". Can we still use the above linear regression to provide a valid estimate of the causal effect of treatment? Briefly justify your answer.

NOTE: Questions 3 and 4 are on the reverse side of this page.

3. The table below is a subset of a table in Imbens and Rubin. It gives data for 8 fast food restaurants that were part of a study of the effect of raising the minimum wage in NJ. The “treatment” group is the 2 restaurants in NJ and the “control” group is a set of 6 restaurants in PA (where the minimum wage was not raised). The outcome Y_i^{obs} is the number of people employed (including part time employees) at the end of the year. There are two covariates – X_{i1} , the identify of the fast food chain (Burger King or Kentucky Fried Chicken) and X_{i2} , employment at the end of the year prior to the increase in the minimum wage.

Observation	Treatment	Rest.Chain	Init.Empl.	Final.Empl.
i	W_i	X_{i1}	X_{i2}	Y_i^{obs}
1	NJ	BK	22.5	30.0
2	NJ	KFC	14.0	12.5
3	PA	KFC	13.8	17.0
4	PA	BK	26.5	18.5
5	PA	BK	20.0	19.5
6	PA	BK	13.5	21.0
7	PA	BK	32.5	26.5
8	PA	KFC	21.0	23.0

- (a) We want to use matching to estimate the effect of raising the minimum wage assuming that unconfoundness holds. We will match a single control unit with each treatment unit (without replacement). Our distance measure is $D(i, j) = 100 * I(X_{i1} \neq X_{j1}) + |X_{i2} - X_{j2}|$ where the indicator is 1 if the two units are different chains and 0 if they are the same chain. Identify the matches for the 2 treatment units.
- (b) What is the estimate of the average treatment effect on the treated units (ATT)?
- (c) An alternative estimand is the average treatment effect (ATE). Which estimand makes more sense here? Briefly justify your answer.
4. The table below lists the results of a hypothetical experiment on 200 people. Each row identifies a category of people with the same values of X, W, Y(0), Y(1), and Y^{obs} . Thus for example there are 30 people in category 1 and each of these people has $X = 0$, $W = 0$, $Y(0) = 4$, $Y(1) = 6$ and because they are in the control group we have $Y^{obs} = 4$.

Category	# people	X	Y(0)	Y(1)	W	Y^{obs}
1	30	0	4	6	0	4
2	30	0	4	6	1	6
3	10	1	4	6	0	4
4	30	1	4	6	1	6
5	20	0	10	12	0	10
6	20	0	10	12	1	12
7	15	1	10	12	0	10
8	45	1	10	12	1	12

- (a) Do you believe these data came from a randomized experiment? Justify your answer.
- (b) Describe what it means for the treatment assignment mechanism to be unconfounded given X?
- (c) Do you believe that treatment assignment is unconfounded given X for these data? Justify your answer.
- (d) Give the definition of a propensity score $e(X)$.
- (e) Assuming the table above reflects the population of interest, what is the propensity score for $X = 0$ and $X = 1$?
- (f) The naive estimate of the average treatment effect compares the observed mean value of Y for those assigned to treatment ($W = 1$) and those assigned to control ($W = 0$). In this case the naive estimate is $9.12 - 6.80 = 2.32$. Carry out a propensity score analysis (any suitable propensity score analysis is fine) to estimate the average treatment effect and show that it yields the correct estimate. Be specific about your approach.