PS C236A / Stat C239A
Problem Set 2
Due: Oct. 1, 2012

Instructions

This assignment is due **4 pm Monday, Oct. 1.** You may submit your analytical work either electronically or in paper form. Electronic versions must be sent as a .pdf to <`jahenderson[at]berkeley.edu`>. Paper copies should be placed in my mailbox in 210 Barrows. For the computing portion of the assignment, you <u>must</u> submit a fully executable version of all .R code, along with any data used in the code (excepting that provided through the course webpage) to the email above. All files for each assignment sent electronically should be included in one omnibus email, with the subject line containing the course and homework number, and your last name (e.g., PS239A/STAT236A: HW2 - Romney).

You are encouraged to work together in groups to complete the assignments. However, you must hand in your own individual answers. Photocopies and other reproductions of someone else's answers are not acceptable. Please also list the names of everyone with whom you have collaborated on this assignment.

Problem 1: The Lady Tasting Tea Consider the following variation of the Lady Tasting Tea example that we discussed in class. The Lady tastes eight cups of tea, four of which have milk added first and four of which have tea added first. The cups are organized into matched pairs and for each pair, a fair coin is flipped to determine which gets milk first. The Lady knows the design, meaning that she knows there is one milk-first cup and one tea-first cup in each matched pair.

a. Consider the following hypothesis test: The null hypothesis is that the Lady has no ability to discriminate the order in which milk is added to tea. The alternative is that the Lady's ability to discriminate the order is better than random chance. In the case where the Lady makes one mistake (classifies one milk-first cup as a tea-first cup), what is the p-value for this hypothesis test?

b. Consider the same null and alternative as in part (a). Suppose now that the cups are no longer paired; instead milk-first or tea-first assignment is completely randomized, with four cups receiving each assignment. The Lady is told that exactly four cups are milk-first, but is given no additional information. If the Lady makes one mistake (classifies one milk-first cup as a tea-first cup), what is the p-value for the hypothesis test? Is this p-value different from the one calculated in part (a)? Why or why not? If you are trying to discern whether the Lady can correctly identify milk-first and tea-first cups, which design would you prefer, the one in (a) or the one in (b)?

c. You believe that the Lady guesses "milk-first" 2/3 of the time. Suppose you have a coin that lands heads 2/3 of the time. For each of the eight cups, you flip the coin and pour milk first or tea first depending on the outcome of the coin. Which would you prefer: Pour milk first on heads or pour tea first on heads?

d. Consider the same null and alternative hypotheses as in part (a). Suppose that for each of the eight cups, you flip a fair coin, and you pour milk into that cup first if that coin lands heads (otherwise, you pour tea first). By chance, seven of the cups are milk-first, and only one of the cups is tea-first. The Lady, when told about the randomization mechanism, states that she will choose at most six cups to be milk-first and at most six cups to be tea-first, as any more than that "is far too unlikely to happen." Suppose that the Lady makes exactly two mistakes. What is the p-value for this hypothesis test?

Problem 2: Catholic School – I In an observational study of the effects of attending a Catholic school, the central dependent variable of interest is a binary variable, Y_i, which indicates whether or not student i graduated from high school. The treatment variable, T_i, indicates Catholic school attendance. In a very large sample of students, half attended Catholic school and half did not. You observe that the treated students have a graduation rate of .7 and the control students have a graduation rate of .5. You wish to estimate the average treatment effect of attending Catholic school. Assume that your sample is large enough to make sampling variability negligible.

a. Without making any assumptions about the relationship between the students' potential outcomes and treatment assignment, what is the largest possible value of the ATE? What is the smallest possible value of the ATE? What is the difference between these two values? Will this difference between the maximum and minimum possible ATE always be the same, irrespective of the specific observed values of the outcome variable?

b. Again making no assumptions about treatment assignment, assume that Catholic school does not prevent any student from graduating. What is the largest possible value of the ATE? What is the smallest possible value of the ATE?

Problem 3: Catholic School – II To address this question more precisely, researchers randomly sample 6 students, measuring some covariate X for students who attend Catholic ($T = 1$) and non-Catholic school ($T = 0$). Assume unconfoundedness holds conditionally on X *and* U. Moreover, assume the conditionality in T_i follows the logit distribution, $\pi_i/(1 - \pi_i) = \exp(X_i\beta + \gamma U_i)$, where π_i is student i's probability of attending Catholic school, and γ and β are additive parameters. In this study, Y, T, and X are observed, while U is not. Though unobserved for the researchers, however, the "true" values of U are given to us below in U_{true}.

Table 1: Catholic School Graduation Data

Unit	Y	T	X	U_{true}	U_{obs}
1	1	1	1.37	1	1
2	0	1	0.16	1	-
3	0	0	0.51	0	0
4	1	1	0.99	1	-
5	1	0	1.53	1	1
6	1	0	-0.46	0	-

a. Match each treated unit to one control unit on X without replacement, to minimize: $\sum_{s=1}^{S}(X_{si} - X_{sj})^2$, where $T_{si} = 1$ and $T_{sj} = 0$. List which units are matched together in each s. What is the resulting McNemar test statistic after matching? What is the estimated ATT after matching?

b. Assume that after matching, $Pr(T_{si} = 1) = Pr(T_{sj} = 1) = \pi_s$, but that π_s varies across each strata. Using the McNemar test statistic from part (a), what is the p-value that this statistic occurred by chance, under the null hypothesis of no effect of attending Catholic school on graduation (in the stratified design)?

c. Now assume that $\gamma = 0.62$ and $\beta = 1.49$. Given your matches in (a), does $\pi_{si} = \pi_{sj}$ for each of the s strata (where again $T_{si} = 1$ and $T_{sj} = 0$)? If these are different, which strata has the largest difference in the probability of attending Catholic school for treated and control units? Now what is the p-value that the McNemar statistic, estimated from your matches in part (a), occurred by chance under the null of no effect of attending Catholic school on graduation, given this new information about X, U_{true}, γ, and β in the stratified design? Is this p-value the same as you calculated in part (b)? Why or why not?

d. The researchers find additional money to collect some information about U, but can only do so for units 1, 3, and 5. (This data is presented as U_{obs} in Table 1 above; in the table, a "dash" means missing data.) Create a new stratification, this time matching with replacement on X, again to minimize $\sum_{s=1}^{S}(X_{si} - X_{sj})^2$. Calculate the ATT as your test statistic after matching. Again assume that $\gamma = 0.62$ and $\beta = 1.49$. Under the null hypothesis of no effect in this stratification design, what are the *largest* and the *smallest* p-values possible, associated with the probability that this statistic occurred by chance, given our remaining ignorance in U_{obs} about U_{true}?

Problem 4: In this problem, you will analyze a famous experiment conducted by Leonard Wantchekon in Benin in 2001. Wantchekon wanted to examine the effectiveness of different types of campaign messages on voting behavior in a presidential election. For details, see:

`http://www.princeton.edu/~lwantche/Clientelism_and_Voting_Behavior_Wantchekon.pdf`

Wantchekon convinced the campaigns of the major presidential candidates to randomize the messages they employed in 24 villages. The two treatment conditions were as follows:

1. *Public Policy:* Wantchekon describes this treatment condition as: "It was decided that any public policy platform would raise issues pertaining to national unity and peace, eradicating corruption, alleviating poverty, developing agriculture and industry, protecting the rights of women and children, developing rural credit, providing access to the judicial system, protecting the environment, and/or fostering educational reforms."

2. *Clientelist*: Wantchekon describes this treatment as: "A clientelist message, by contrast, would take the form of a specific promise to the village, for example, for government patronage jobs or local public goods, such as establishing a new local university or providing financial support for local fishermen or cotton producers."

The data has been modified for the assignment, but the basic structure of the experiment was *block* randomization. For the purposes of the assignment, villages were divided into groups of 2 based on geography and treatment status was randomized within the 8 groups of 2. The outcome variable is the vote share of the candidate participating in the experiment. The only covariate is the number of registered voters. In the dataset, `block` indicates block group, `reg.voters` is the registered voters covariate,`vote.pop`is the outcome variable, `treat` is a variable indicating treatment status.The data can be found here:

`http://sekhon.berkeley.edu/causalinf/data/hw2data.RData`

In this problem, we are interested in the difference between the clientelist and public policy conditions.

a. Estimate the effect the clientelist message compared to the public policy message, using the ITT estimator and the regression estimator. For the regression estimate, include block level dummy variables in your regression equation.

a. Now test the sharp null of no treatment effect using randomization inference. Use two test statistics: Wilcoxon's signed rank test (Rosenbaum 2002, pg. 32) and the difference in means. What are the two sided p-values under these two tests?

a. Under the assumption of a constant, additive, treatment effect, use randomization inference to find a 95% confidence interval of the treatment effect. Use the signed rank as your test statistic. See pages 44-46 in Rosenbaum (2002).

d. What can you conclude about the effectiveness of clientelistic appeals in Benin?

e. Bonus: Perform randomization inference with covariance adjustment. How does this effect your results? For a very good article on covariance adjustment with randomization inference, see:

   Rosenbaum, Paul. 2002. "Covariance Adjustment in Randomized Experiments and Observational Studies." *Statistical Science* 17(3): 286-327.