Political Science 239

Problem Set 3

Due date: Wednesday, September 27, 2006

Exercise 1 *Consider the following variation of the Lady Tasting Tea example. The lady tastes six cups, three of which have milk added first and three of which have tea added first. The cups are presented to the lady in random order. The lady knows both that there are exactly three milk-first cups and three tea-first cups, and that the cups will be presented to her randomly.*

(1.1) What is the significance level for a test that rejects the null hypothesis that the lady has no ability to discriminate the order in which milk is added to tea?

(1.2) Now allow the lady to make one mistake, i.e. to classify a milk-first cup as a tea-first cup. What is now the significance level for a test that rejects the null hypothesis of no ability to discriminate?

(1.3) Now allow the lady to make two mistakes, i.e. to classify two milk-first cups as a tea-first cups. What is now the significance level for a test that rejects the null hypothesis of no ability to discriminate?

(1.4) How sensitive is this experimental design?

Exercise 2 *Consider the Lady Tasting Tea example of Exercise 1 under binomial randomization. There are* 6 *cups and each cup has probability* $p = \frac{1}{2}$ *of having milk first and* $1 - p$ *probability of having tea added first. The lady does not know the value of* p*.*

(2.1) What is the significance level for a test that rejects the null hypothesis that the lady has no ability to discriminate the order in which milk is added to tea when the lady is allowed no mistakes?

(2.2) What is the significance level for a test that rejects the null hypothesis that the lady has no ability to discriminate the order in which milk is added to tea when the lady is allowed one mistake?

(2.3) [BONUS] Now suppose that each cup has probability $p = \frac{1}{3}$ of having milk first. You can choose between two null hypotheses: the lady has no ability to identify milk-first cups and the lady has no ability to identify tea-first cups. Which of these null hypotheses would provide you with a more sensitive test to analyze wether the lady has the ability to discriminate? Calculate the significance level of this test if we allow the lady one mistake.

Exercise 3 *In this exercise, we will continue the univariate matching example that we saw in lecture (see code "nn1.R"). Use the dataset "election.raw" to perform the following analysis. The codebook can be downloaded from http://sekhon.berkeley.edu/causalinf/R/codebook_election04.raw.txt. The variable that indicates whether an observation is treatment or control is "etouch", an indicator for electronic voting, and the outcome analyzed is the variable "b04pc", the proportion of votes obtained by Bush in the 2004 election.*

1. *Using the "Match()" function, perform univariate matching on the variable "d96pc" (proportion of votes for Dole in 1996) and calculate the ATT of electronic voting on the proportion of votes obtained by Bush in 2004.*

2. *Using the "MatchBalance()" function, analyze the balance in your matched sample on the variable d96pc and the interaction of d96pc and the hispanic dummy (variable "hispanic"). How good is your balance? Report the means for treatment and control both before and after matching, and the mean, maximum and median difference in the empirical quantiles of the*

treatment and control distribution both before and after matching. Report also the ratio of the treatment and control variances and the T-test p-value.

3. *Using the "qqplot()" function, plot the empirical quantiles of the distribution of d96pc for treatments and controls. Report two qq-plots: one before matching and one after matching. In the qqplot after matching, can you identify any outlier counties for which a good match couldn't be found? Which counties?*

Exercise 4 *[THIS IS A BONUS EXERCISE] Consider the following variation of the Lady Tasting Tea example. The lady tastes seven cups, three of which have milk added first and four of which have tea added first. The cups are presented to the lady in random order. The lady knows both that there are exactly three milk-first cups and four tea-first cups, and that the cups will be presented to her randomly.*

(4.1) Using the number of cups correctly identified as milk-first as the test statistic of the experiment, what is the significance level for a test that rejects the null hypothesis that the lady has no ability to discriminate the order in which milk is added to tea?

(4.2) Using the same test-statistic, now allow the lady to make one mistake, i.e. to classify a milk-first cup as a tea-first cup. What is now the significance level for a test that rejects the null hypothesis of no ability?

(4.3) Using the same test-statistic, now allow the lady to make two mistakes, i.e. to classify two milk-first cups as a tea-first cups. What is now the significance level for a test that rejects the null hypothesis of no ability?

(4.4) Now use the number of cups correctly identified as tea-first as the test statistic of the experiment. Recalculate the significance levels for a test that rejects the null if the lady is allowed (i) no mistakes, (ii) one mistake and (iii) two mistakes. Are these different than the ones you calculated before?

(4.5) How sensitive is this experimental design?