Homework for Chapter 14: Matching

1. You want to know whether practicing cursive improves your penmanship (on a 1-10 scale). You find that, among people who don’t practice cursive, average penmanship is 5, 10 people are left-handed, 2 are ambidextrous, and 88 are right-handed. Among people who do practice cursive, 6 are left-handed with average penmanship 7, 4 are ambidextrous with average penmanship 4, and 90 are right-handed with average penmanship 6.
   1. You want to create a set of weights that will *make the treated group match the control group on handedness*. Follow the process in section 4.2, paying attention to *why* certain numbers are going in certain positions. What weights will be given to the left, ambidextrous, and right-handed people *in the control group*?

Answer: I will the weight of 1 for all the left, ambidextrous, and right-handed people in the control group.

* 1. What weights will be given to the left, ambidextrous, and right-handed people *in the treated group*?

Answer: In the treated group:

For left-handed: 10/6=5/3≈1.6667

For ambidextrous: 2/4=0.5

For right-handed: 88/90=44/45≈0.97778

* 1. Use the weights from part b to calculate the *proportion of left-handed people in the treated group*, as well as the proportion of ambidextrous people and the proportion of right-handed people. If you don’t get 10%, 2%, and 88% (or very close with some rounding error), your weights are wrong, try again.

Answer: For left-handed: 6\*1.6667/(6\*1.6667+4\*0.5+90\*0.97778)≈10%

For ambidextrous: 4\*0.5/(6\*1.6667+4\*0.5+90\*0.97778)≈2%

For right-handed: 90\*0.97778/(6\*1.6667+4\*0.5+90\*0.97778)≈88%

* 1. What is the weighted average penmanship score in the treated group?

Answer: The weighted average penmanship score in the treated group is about .

(1.6667\*7+0.5\*4+0.97778\*6)/(1.6667+0.5+0.97778)≈6.21

* 1. What is the effect of practicing cursive that we would estimate using this data?

Answer: We would estimate the average effect on the untreated of practicing cursive because we are trying to match treated observations to control observations. The effect of practicing cursive is about 1.21 (6.21-5).

1. For each of the following descriptions of matching on the variable , determine whether this is describing *one-to-one distance matching, k-nearest-neighbor distance matching, kernel matching,* or *propensity score matching* (hint: it’s one of each)*.*
   1. The treated observation has . For each control observation, is calculated, with the result run through a weighting function. The resulting weight is applied to that observation. (*kernel matching)*
   2. The treated observation has . Among the control observations, the nearest values are and . The observations with and are chosen as a control, since they’re the two closest. (*k-nearest-neighbor distance matching)*
   3. The treated observation has . You estimate a model that suggests that observations with have a .6 chance of being treated. You similarly calculate the chance of treatment for each control observation, and use those calculated probabilities to create a weight for each observation. (*propensity score matching)*
   4. The treated observation has . Among the control observations, the observation with is closest to that, and so is selected as a control. (*one-to-one distance matching)*
2. For each of the following decisions to be made in the process of matching, determine which option produces *more bias* (in each case, the other option will produce *more variance*)
   1. (A) selecting one control match for each treatment vs. (B) selecting multiple control matches for each treatment

Answer: b. selecting multiple control matches for each treatment

* 1. (A) using a relatively wide bandwidth vs. (B) using a narrower bandwidth

Answer: a. using a relatively wide bandwidth

* 1. (A) selecting matches with replacement vs. (B) selecting matches without replacement

Answer: b. selecting matches without replacement

* 1. (A) selecting one control match for each treatment vs. (B) applying a weight that accepts many controls but decays with distance

Answer: b. applying a weight that accepts many controls but decays with distance (because b will be involved with more variables than a)

1. Why should exact matching (or coarsened exact matching) generally be reserved for very large samples or situations where a very small number of matching variables is appropriate?

Answer: Because exact matching will lead to dropping out too many treated observations which might result in a poor representation of the average treatment effect if certain kinds of treated observations are more likely to find matches than others if the sample size is not large enough. Besides, a big number of matching variables will lead to the curse of dimensionality and create too many cells after interacting all the variables, which means that matching becomes harder and treated observation is more likely to be dropped out. So exact matching (or coarsened exact matching) should generally be reserved for very large samples or situations where a very small number of matching variables is appropriate.

1. You are looking at the effect of participating in high school sports on high school grades. You compare students who did and did not participate in sports, using one-to-one matching with a Mahalanobis distance, with replacement and a caliper of .3, to match on high school athleticism, parental income, gender, race, and middle school grades. You find that sports participation reduces grades, but by only .1 grade points. As clearly and precisely as possible, outline the steps that were taken in performing this analysis.

Answer: First, calculate the Mahalanobis distance between the treated variables and the untreated valuables. (Take each matching variable (high school athleticism, parental income, gender, race, and middle school grades) and divide its value by its standard deviation, then take the square root of the sum of the squares of all the differences between A and B, if the matching variables are independent from each other; if they are related, we can instead divide out the whole covariance matrix from the squared values of the variables.) Second, select the untreated/treated variables to match treated/untreated variables with a caliper of 0.3 based on the Mahalanobis distance. If the Mahalanobis distance is bigger than 0.3, we can drop the variables out. For the untreated variables that is smaller than 0.3, we should take the variables with the least Mahalanobis distance into the control group. The variables can be taken into the control group for several times as long as its mahalanobis distance is smaller than 0.3sd and with the least mahalanobis distance. Then we should give a weight to the control variables equal to the number of times it has matched. Third, then use the matched data sets of treated and control groups to the effect of participating in high school sports on high school grades.

1. Which of the following is a downside of propensity score matching compared to other methods of matching?

Answer: d.

* 1. It can’t be combined with exact matching in cases where one variable must be exactly matched
  2. It focuses the matching adjustment on differences that close back doors, rather than all differences
  3. It requires the selection of matches instead of the use of weights, which increases variance.
  4. It requires that the model used to estimate the propensity score is properly specified.

1. You are planning to evaluate the effect of a tax-rebate plan for small businesses. Some businesses were eligible based on their tax returns and others weren’t. You would like to match on industry and number of employees. A table showing the number of businesses for each combination of industry and number of employees for the treated and untreated groups are in the following table:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Number of small businesses by industry, number of employees, and treatment status | | | | | |
|  | Treated |  |  | Untreated |  |
| N. Employees | Retail | Service |  | Retail | Service |
| 1-5 | 3 | 4 |  | 0 | 4 |
| 6-10 | 3 | 2 |  | 4 | 3 |
| 11-20 | 0 | 5 |  | 5 | 1 |

* 1. For what group of treated businesses would we say that the common-support assumption definitely fails?

Answer: The group of treated business with 1-5 employees in retail industry.

* 1. There are no treated retail businesses with 11-20 employees. Is this a concern for the common support assumption if we are trying to estimate an average treatment on the treated?

Answer: No, it is not a concern for the common support assumption if we are trygin to estimate an average treatment on the treated.

* 1. What concern might we have about there only being one untreated Service business with 11-20 employees?

Answer: If we do the matching without replacement, then about 80% of the treated service business with 11-20 employees cannot find a match, which might lead to the failure of common-support assumption; if we do the matching with replacement, then it might lead to greater noise in our estimates, and our estimates will be less precise.

* 1. If we resolved the common support problem for the group from problem (a) by dropping members of that group from the data, what problem would that create for our analysis?

Answer: Dropping many members of that group from the data might lead to a poor representation of the average treatment effect and we can only get the estimate of the treatment effect among the businesses that match well. What is worse, we can also say that the matching has failed.

1. You perform a matching analysis on a schooling reform to create a set of matching weights, matching on the per-capita income and expenditures of the school. You then produce the below weighted balance table comparing the weighted means for treatment and control.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Treated | |  | Untreated |  |  |
|  | N | Mean | SD | N | Mean | SD | Test |
| Expenditure | 29 | 389 | 106.677 | 21 | 351.524 | 71.529 | F=1.951 |
| Income | 30 | 7749.7 | 1127.359 | 21 | 7406.381 | 888.136 | F=1.356 |
| Matching weights applied. Statistical significance markers: \* p<0.1; \*\* p<0.05; \*\*\* p<0.01 | | | | | | |  |

* 1. This particular balance table reports F-statistics of differences in means, with statistical significance markers. Are there statistically significant differences in either of the variables between the treated and untreated group at the 95% level?

Answer: There are no statistically significant differences in either of the variables between the treated and untreated group at the95% level because there are no stars for the two variables.

* 1. You don’t have enough information to actually evaluate this, but make a list of two things you’d think about when deciding whether it looks like there’s a balance problem based on the difference in means regardless of whether the difference is statistically significant. As an example, answer while thinking of the difference of 7749.7 – 7406.4 = 342.3 between treated and untreated in Income.

Answer: We can also focus on the distribution of the matching variables. So we could make an overlaid density plot of the income of treated and untreated group to see whether they are almost identical. If they are not identical, then there is a balance problem. If they are identical, then there is no balance problem.

We can also use a propensity score to test whether there is a balance problem. We can control for the propensity score based on the income then see whether there is relationship between income and treatment status. If there is a relationship between income and treatment status, then there is a balance problem. If there is no relationship between income and treatment status, then then there is no balance problem.

* 1. Imagine you *did* find lots of significant differences here after constructing matching weights using propensity score matching, even though these variables were included as matching variables. What would your next step be?

Answer: I will first use the overlaid density plots and propensity score to check the balance problems again. If there are balance problems, I will change something about the matching procedure, such as using more variables and taking a tighter caliper, and finally repeat the process of checking balance problems.

1. Explain why selecting untreated observations to match the treated observations produces an average treatment effect on the treated (ATT), while selecting treated observations to match the untreated observations produces an average treatment effect on the untreated (ATUT).

Answer: By matching, we are trying to build up a group of variables with similar conditions but without/with treatment. So it seems that we are trying to estimate the counterfactuals by matching. Therefore, selecting untreated observations to match the treated observations is trying to build up a treated group that would have gotten without treatment, so we can get the average treatment effect on the treated (ATT); selecting treated observations to match the untreated observations is trying to build up a untreated group that would have gotten with treatment, so we can get the average treatment effect on the untreated (ATUT).

1. Coding