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*?
   2. What weights will be given to the left, ambidextrous, and right-handed people *in the treated group*?
   3. 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.
   4. What is the weighted average penmanship score in the treated group?
   5. What is the effect of practicing cursive that we would estimate using this data?
2. 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.
   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.
   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.
   4. The treated observation has . Among the control observations, the observation with is closest to that, and so is selected as a control.
3. 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
   2. (A) using a relatively wide bandwidth vs. (B) using a narrower bandwidth
   3. (A) selecting matches with replacement vs. (B) selecting matches without replacement
   4. (A) selecting one control match for each treatment vs. (B) applying a weight that accepts many controls but decays with distance
4. 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?
5. 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 Mahanobis 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.
6. Which of the following is a downside of propensity score matching compared to other methods of matching?
   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.
7. 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?
  2. 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?
  3. What concern might we have about there only being one untreated Service business with 11-20 employees?
  4. 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?

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?
  2. 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.
  3. 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?

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).
2. Coding