

These exercises are based on The Effect chapter 14. As you will see, some of the questions refer to sections of that chapter.

Problem Set Part 1

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.
 - (a) 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?
 - (b) What weights will be given to the left, ambidextrous, and right-handed people in the treated group?
 - (c) 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.
 - (d) What is the weighted average penmanship score in the treated group?
2. For each of the following decisions to be made in the process of matching, determine which option produces more bias and try to explain why (in each case, the other option will produce more variance):
 - (a) (A) selecting one control match for each treatment vs. (B) selecting multiple control matches for each treatment
 - (b) (A) using a relatively wide bandwidth vs. (B) using a narrower bandwidth
 - (c) (A) selecting matches with replacement vs. (B) selecting matches without replacement
 - (d) (A) selecting one control match for each treatment vs. (B) applying a weight that accepts many controls but decays with distance
3. 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?

4. Which of the following is a downside of propensity score matching compared to other methods of matching?
 - (a) It can't be combined with exact matching in cases where one variable must be exactly matched
 - (b) It focuses the matching adjustment on differences that close back doors, rather than all differences
 - (c) It requires the selection of matches instead of the use of weights, which increases variance
 - (d) It requires that the model used to estimate the propensity score is properly specified
5. 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 is below:

	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

- (a) For what group of treated businesses would we say that the common-support assumption definitely fails?
 - (b) 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?
 - (c) What concern might we have about there only being one untreated service business with 11–20 employees?
 - (d) If we resolved the common support problem for the group from part (a) by dropping members of that group from the data, what problem would that create for our analysis?
6. 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 control (ATC).

Problem Set Part 2: Coding

1. Load the `nsw_mixture` data. Documentation on the variables is available [here](#).

2. Let's see where we're at before we do any matching at all. `nsw_mixture` is from an experiment (read that documentation!) so that should already put us in a pretty good place.
 - (a) First, create a variable called `weight` in your data equal to 1 for all observations (weights that are all 1 will have no effect, but this will give us a clue as to how we could incorporate matching weights easily).
 - (b) Second, write code that uses a set of given weights to estimate the effect of `treat` on `re78`, using `weight` as weights, and prints out a summary of the regression results.

Note: Keep in mind the standard errors on the estimate won't be quite right, since they won't account for the uncertainty in creating the weights.
 - (c) Third, write code yourself (as we did in the lab) that creates and prints out a weighted balance table for all variables across values of `treat`, using `weight` as weights (hint: you can use the `weighted.mean()` function for this).
 - (d) Is there anything potentially concerning about the balance table, given that this is a randomized experiment where `treat` was randomly assigned?
3. Mahalanobis distance matching:
 - (a) Using all of the variables in the data except `treat` and `re78` as matching variables, perform 3-nearest-neighbor Mahalanobis distance matching with replacement and no caliper (The Effect 14.4.1).
 - (b) Create a post-matching balance table and a love plot showing balance for all the matching variables (you can use functions from the Matchit and Cobalt packages to do this or you can write your own). Explain whether the balance looks good.
 - (c) Calculate the post-matching average treatment on the treated effect of `treat` on `re78`.
4. Switching over to propensity score matching:
 - (a) Use the same matching variables as in Question 3 to estimate the propensity to be treated (with a logit regression), and then add the treatment propensity to the data set as a new variable called `propensity`.
 - (b) Create a new variable in the data called `ipw` with the inverse probability weight.
 - (c) Make a common support graph, overlaying the `propensity` for treated observations on top of the `propensity` for untreated observations. Write a line commenting on how the common support looks.
 - (d) Estimate the treatment effect using the `ipw` weights in a linear regression (keeping in mind the standard errors won't be quite right).