Chapter 9
Various Practical Issues in Matching

Abstract Having constructed a matched control group, one must check that it is sat-
isfactory, in the sense of balancing the observed covariates. If some covariates are
not balanced, then adjustments are made to bring them into balance. Three adjust-
ments are almost exact matching, exact matching, and the use of small penalties.
Exact matching has a special role in extremely large problems, where it can be used
to accelerate computation. Matching when some covariates have missing values is
discussed.

9.1 Checking Covariate Balance

Although a table such as Table 7.1 in Chapter 7 may suffice to describe covariate
balance in a scientific paper [15], somewhat more is typically needed when con-
structing a matched sample. Checks on covariate balance in matching are informal
diagnostics, not unlike residuals in a regression. They aid in thinking about whether
the treated and control groups overlap sufficiently to be matched, and whether the
match currently under consideration has achieved reasonable balance or whether
some refinements are required. The study of ovarian cancer [15] in Chapter 7 is
used to illustrate.

One common measure of covariate imbalance is a slightly unusual version of
an absolute standardized difference in means [12]. The numerator of the standard-
dized difference is simply the treated-minus-control difference in covariate means
or proportions, and it is computed before and after matching. The first variable,
surgeon-type-GO, in Table 7.1 is a binary indicator, which is 1 if a GO performed
the surgery and is 0 if someone else performed the surgery. Before matching,
the difference in means is $0.76 - 0.33 = 0.43$, or 43%, and after matching it is
$0.76 - 0.75 = 0.01$ or 1%. Write $\bar{x}_{tk}$, $\bar{x}_{ck}$, $\bar{x}_{cmk}$ for the means of covariate $k$ in, re-
spectively, the treated group, the control group before matching, and the matched
control group, so $\bar{x}_{t1} = 0.76$, $\bar{x}_{c1} = 0.33$, and $\bar{x}_{cmk} = 0.75$ for the first covariate in
Table 7.1. It is the denominator that is just slightly unusual in two ways. First, the
denominator always describes the standard deviation before matching, even when measuring imbalance after matching. We are asking whether the means or proportions are close; we do not want the answer to be hidden by a simultaneous change in the standard deviation. Second, the standard deviation before matching is calculated in a way that gives equal weight to the standard deviation in the treated and control groups before matching. In many problems, the potential control group is much larger than the treated group, but we do not want to give much more weight to the standard deviation in the control group. Write $s_{tk}$ and $s_{ck}$ for the standard deviations of covariate $k$ in the treated group and in the control group before matching. The pooled standard deviation for covariate $k$ is $\sqrt{(s_{tk}^2 + s_{ck}^2)/2}$. The absolute standardized difference before matching is $sd_{bk} = |\bar{x}_{tk} - \bar{x}_{ck}|/\sqrt{(s_{tk}^2 + s_{ck}^2)/2}$ and the absolute standardized difference after matching is $sd_{mk} = |\bar{x}_{tk} - \bar{x}_{cmk}|/\sqrt{(s_{tk}^2 + s_{ck}^2)/2}$; notice that they are identical except that $\bar{x}_{cmk}$ replaces $\bar{x}_{ck}$. For the first covariate in Table 7.1, surgeon-type-GO, the absolute standardized difference is 0.95 before matching and 0.02 after matching, that is, almost a full standard deviation before matching, and about 2% of a standard deviation after matching.

The boxplots in Figure 9.1 display 67 absolute standardized differences before and after matching. The list of 67 covariates is slightly redundant; for instance, all three categories of surgeon type appear as three binary variables, even though the value of one of these variables is determined by the values of the other two. The imbalances before matching are quite large: there are four covariates with differences of more than half a standard deviation. After matching, the median absolute standardized difference is 0.03 or 3% of a standard deviation, and the maximum is 0.14. In fact, because fine balancing was used to construct this matched sample, 18 of the 67 absolute standardized differences equal zero exactly.

The principal advantage of an absolute standardized difference over an unstandardized difference, say $|\bar{x}_{tk} - \bar{x}_{cmk}|$, is that variables on different scales, such as age and hypertension, can be plotted in a single graph for quick inspection. The disadvantage is that a covariate such as age means more in terms of years than in terms of standard deviations. In practice, it is helpful to examine an unstandardized table such as Table 7.1 in addition to graphs of standardized differences.

In matching with variable controls, as in §8.5, or in full matching, as in §8.6, the mean in the matched control group, $\bar{x}_{cmk}$, is a weighted mean, as described in §8.5 and §8.6. Then the absolute standardized difference is computed using this weighted mean, $|\bar{x}_{tk} - \bar{x}_{cmk}|/\sqrt{(s_{tk}^2 + s_{ck}^2)/2}$.

It might seem desirable that all of the absolute standardized differences equal zero, but this would not happen even in a completely randomized experiment. How does the imbalance in the boxplots in Figure 9.1 compare with the imbalance expected in a completely randomized experiment?

Imagine a completely randomized experiment. This means that 688 unmatched patients are randomly divided into two groups, each with 344 patients. If a randomization test were applied to one covariate to compare the distribution of the covariate in these randomly formed groups, it would produce a $P$-value less than or equal to