REGRESSION WITH COVARIATES

In the previous module, we took up the model based approach to the estimation of causal effects. We focused on linear regression in lessons 2 and 3.

We saw that in a completely randomized experiment

$$\hat{\tau}^* = \bar{Y}_1 - \bar{Y}_0 - \underline{\beta}^{*\prime}(\underline{\bar{X}}_1 - \underline{\bar{X}}_0)$$

is unbiased for the ATE. It also has smaller variance than the estimator

$$\bar{Y}_1 - \bar{Y}_0$$

Note, we made no assumption that the regression model was correctly specified nor that ATE(X) $=$ ATE.

REGRESSION BIAS

In observational studies, we know that $\bar{Y}_1 - \bar{Y}_0$ is generally neither unbiased nor consistent for the ATE due to the presence of confounders $\underline{X}$.

But

$$E(Y|Z = z, \underline{X}) = E(Y(z)|\underline{X})$$

This suggests modeling the conditional expectation function for the observed data and using our estimate of this to estimate the ATE. In lesson 3, we did just that, again using linear regression.

Unlike the case of the completely randomized experiment, however, we saw that if the regression function was not specified correctly, the estimate of the ATE was biased, unless it just so happened that $\bar{\underline{X}}_1 = \bar{\underline{X}}_0$. In observational studies, the difference in means in the treatment and control groups can be substantial.

ALLEVIATING REGRESSION BIAS

To remove the bias from estimates of ATE's in observational studies, the previous analysis suggests two remedies.

1. One might attempt to model the regression function nonlinearly. However, the researcher often does not have enough knowledge to specify a functional form. This could lead to a non-parametric model.

2. One might attempt to make the covariate distributions the same in the treatment and control group like a completely randomized study. A traditional way to achieve such "balance" is by matching. For each unit in the treatment group, we find an observation in the control group with the same covariate values $\underline{X}$. The difference in the outcomes between the two units would then be unbiased for the ATE($\underline{X}$).

PROPENSITY SCORE

Both approaches
1. estimating the regression function
2. matching,

may become unwieldy when there are many covariates; the so-called curse of dimensionality. In the past 30 years, great advances have been made in computing that have allowed significant progress to be made in both directions.

We shall discuss some of these advances. But to place these in context and to understand the role of certain key ideas — also as the majority of empirical studies are not conducted using these latest advances — it is useful to start with the propensity score that comes from the seminal article of Rosenbaum and Rubin (1983).

The propensity score also figures prominently in "weighting" approaches. So it is important for us to obtain a good [redacted]

PROPENSITY SCORE

The propensity score is simply the probability of treatment, given covariates **X**

$$e(\underline{X}) = Pr(Z = 1 \mid \underline{X})$$

Rosenbaum and Rubin showed that if treatment assignment is strongly ignorable, given covariate **X**, that is,

1. treatment assignment is unconfounded, given covariates:
   $Y(0), Y(1) \perp\!\!\!\perp Z \mid \underline{X}$
   and

2. : $0 < e(\underline{X}) < 1$,

then treatment assignment is strongly ignorable given $e(\mathbf{X})$.

They also showed that the distribution of the confounding covariates is the same in the treatment group and the control

IGNORABLE ASSIGNMENT GIVEN PROPENSITY SCORE

These results are not hard to prove. First,

$$
\begin{aligned}
Pr(Z = 1 \mid e(\underline{X}), Y(0), Y(1)) &= E(Z \mid e(\underline{X}), Y(0), Y(1)) \\
&= E(E(Z \mid \underline{X}, Y(0), Y(1))|e(\underline{X}), Y(0), Y(\\
&= E(e(\underline{X})|e(\underline{X}), Y(0), Y(1)) = e(\underline{X})
\end{aligned}
$$

The unconfoundedness assumption is used in the second to last equality.

This result implies $E(Y \mid Z, e(\underline{X})) = E(Y(z) \mid e(\underline{X}))$. Thus, even though it might be difficult to non-parametrically model the conditional expectation as a function of many covariates due to the curse of dimensionality, it may be easier to do so as a function of the scalar propensity score. The estimated regression functions may then be used to form an estimate of ATE($e(\underline{X})$), and averaging over the distribution of the propensity score gives an estimate of ATE $= \int$ ATE$(e(\underline{X}))dF(e(\underline{X}))$

BALANCE GIVEN PROPENSITY SCORE

That the distribution of the confounding covariates is the same in the treatment group and the control group for subjects with the same propensity score, follows by a similar argument:

$$
\begin{aligned}
Pr(Z = 1 \mid e(\underline{X})) &= E(Z \mid e(\underline{X})) \\
&= E(E(Z \mid \underline{X}) \mid e(\underline{X})) \\
&= E(e(\underline{X}) \mid e(\underline{X})) \\
&= e(\underline{X}) \\
&= Pr(Z = 1 \mid \underline{X}) \\
&= Pr(Z = 1 \mid \underline{X}, e(\underline{X}))
\end{aligned}
$$

Because the distribution of the covariates is the same in the treatment and control groups, given the propensity score, the propensity score is said to be a balancing score.

BALANCE GIVEN PROPENSITY SCORE

The scores, $e(\underline{X})$ divide the covariate space into components with equal probability of receiving treatment. Any division that is finer will also "balance" the covariates, while any division that is coarser will not.

As is evident from these arguments, but sometimes forgotten, the propensity score is a balancing score whether or not treatment assignment is unconfounded.

However, using the propensity score to create groups that are balanced on $\underline{X}$ does not imply potentially important confounders that the investigator has failed to include in $\underline{X}$ are balanced across the treatment and control groups.

PROPENSITY SCORE SUB-CLASSIFICATION

In a randomized block experiment, the function $S(\underline{X})$ mapping the covariates into strata is a balancing score.

Previously, we saw that a natural way to estimate the ATE from a randomized block experiment is to first estimate ATE's using the stratum proportions as weights. The result thus immediately suggests using the propensity score to form strata and proceeding as in the randomized block experiment.

Rosenbaum and Rubin refer to this approach as sub-classification.

PROPENSITY SCORE MATCHING

More generally, the result suggests comparing subjects in the treatment group and subjects in the control group with the same propensity score. Matching observations from the treatment and control groups on the propensity score balances the distribution of covariates across groups.

Prior to the Rosenbaum and Rubin article, statisticians would attempt to match on the covariates themselves to achieve such balance.

But when the covariate space is large and the sample is not so large, one quickly runs out of matches, requiring the investigator to either throw away cases that cannot be matched or accept matches that may not be that close on the covariate values.

PROPENSITY SCORE MATCHING

In theory, matching on the propensity score achieves the same balance without requiring matching on the covariates themselves. As the propensity score is a many to one function of covariates, in theory an investigator should be able to match more subjects more precisely.

In practice, however, matters may not be as rosy. Further, over the last 25 years improved methods for matching directly on the covariates have been developed.

In the next few lessons, we shall study methods used in the statistical literature to estimate effects such as the ATE and ATT in observational studies. We begin by discussing methods that use the propensity score for regression, subclassification and matching. Throughout, we will assume the unconfoundedness assumption holds.