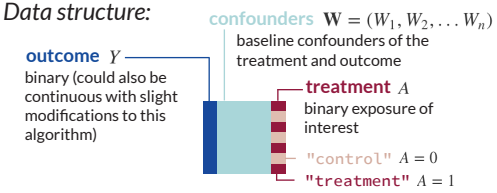
A VISUAL GUIDE TO TMLE

Katherine Hoffman, MS
@kat_hoffman_

Targeted Maximum Likelihood Estimation (TMLE) is a general semiparametric estimation technique. TMLE can incorporate machine learning algorithms while still yielding valid standard errors for statistical inference.

Here we will use TMLE to estimate the mean difference in a binary outcome, adjusting for confounders. Under causal assumptions (not presented here) this is the Average Treatment Effect (ATE), or the difference in outcomes if all observations had received treatment compared to if no observations had received treatment.

Data structure:



Estimand: $ATE = E[E[Y|A = 1, \mathbf{W}] - E[Y|A = 0, \mathbf{W}]]$

1: INITIAL OUTCOMES

Estimate the expected outcome for all observations, using confounders and treatment status as predictors.

$$Q(A, W) = E[Y|A, \mathbf{W}]$$

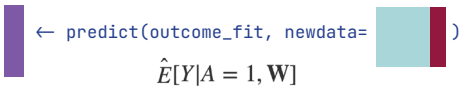
Many flexible machine learning algorithms can be used to fit this equation. See Application.

Then use that model fit to predict every observation's outcome using:

1. The original data set

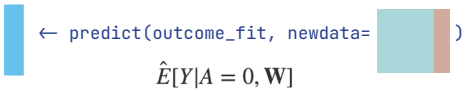
$$\hat{E}[Y|A, \mathbf{W}]$$

2. Every treatment status set to "treatment"



$$\hat{E}[Y|A = 1, \mathbf{W}]$$

3. Every treatment status set to "control"

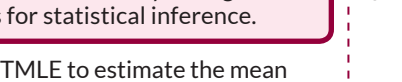


$$\hat{E}[Y|A = 0, \mathbf{W}]$$

These predicted outcomes should be on the same scale as the outcome. Since our outcome is binary, they should be predicted probabilities (rather than the logit of the probability). In Step 3 we will temporarily transform the predicted outcomes to the logit scale to solve an equation.

2: PROBABILITY OF TREATMENT

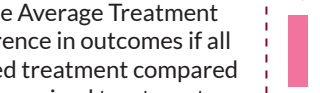
Estimate all observations' probability of receiving the treatment using the confounders as predictors (propensity score).



$$g(W) = P(A = 1|\mathbf{W})$$

Then use that model fit to predict two probabilities:

1. Inverse probability of receiving treatment

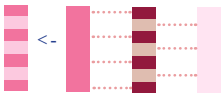


$$H(A = 1, \mathbf{W}) = \frac{1}{\hat{P}(A=1|\mathbf{W})}$$

2. Negative inverse probability of not receiving treatment

$$H(A = 0, \mathbf{W}) = -\frac{1}{\hat{P}(A=0|\mathbf{W})}$$

Finally, use each observation's treatment status to make a "clever covariate." For observations who were treated, the clever covariate is their inverse probability of recieving treatment, and for observations who weren't treated, it's their negative inverse probability of not receiving treatment.



$$H(A, \mathbf{W}) = \frac{I[A=1]}{\hat{P}(A=1|\mathbf{W})} - \frac{I[A=0]}{\hat{P}(A=0|\mathbf{W})}$$

3: FLUCTUATION PARAMETER

The regression fit from Step 1 is optimal to estimate the expected outcome (given treatment and confounders), but not to estimate the ATE. We need to use information about the treatment mechanism in Step 2 to optimize the bias-variance tradeoff for our ATE estimate so that we can obtain valid inference. We will do this by solving an equation to figure out how much to update, or fluctuate, our initial outcome estimates.

$$logit(E[Y|A, \mathbf{W}]) = logit(\hat{E}[Y|A, \mathbf{W}]) + \epsilon H(A, \mathbf{W})$$

To solve this equation, fit a logistic regression using the clever covariate as the only predictor of the observed outcome, and the initially predicted outcome under the observed treatment as a fixed intercept.

The regression's only coefficient is the fluctuation parameter:

$$\hat{\epsilon}$$

Fitting the logistic regression solves an "efficient influence function estimating equation" which yields many useful statistical properties of TMLE, such as: 1) as long as either `outcome_fit` *or* `treatment_fit` *are estimated correctly (consistently), the final estimate is consistent; 2) if both are estimated consistently, the final estimate achieves its smallest possible variance as sample size approaches infinity (efficiency).*

4: UPDATE INITIAL OUTCOMES

The fluctuation parameter, epsilon, from Step 3 is used to update the initial expected outcome estimates:

1. Updated estimate of the expected outcome under treatment

$$\hat{E}^*[Y|A = 1, \mathbf{W}] = expit(logit(\hat{E}[Y|A = 1, \mathbf{W}]) + \hat{\epsilon}H(1, \mathbf{W}))$$

2. Updated estimate of the expected outcome under no treatment

$$\hat{E}^*[Y|A = 0, \mathbf{W}]) = expit(logit(\hat{E}[Y|A = 0, \mathbf{W}]) + \hat{\epsilon}H(0, \mathbf{W}))$$

The logit function, `qlogis`*, and inverse logit function,* `plogis`*, are needed to transform the outcome to the logit scale to fit the logistic regression, and then to transform it back to the original outcome scale.*

5: COMPUTE ATE

Calculate the ATE by taking the average difference between the updated expected outcomes.

$$\hat{ATE} = \hat{E}[\hat{E}^*[Y|A = 1, \mathbf{W}] - \hat{E}^*[Y|A = 0, \mathbf{W}]]$$

6: INFERENCE

We can use the following equation to get standard errors of our TMLE estimate (for confidence intervals and p-values):

See accompanying blog post or references for a brief explanation and formal notation. The equation relies on the functional delta method and empirical process theory.

APPLICATION

Implementation of the TMLE algorithm is straightforward in R using the `tmle`, `tmle3`, and `lmtp` packages:

For best results, estimate `outcome_fit` and `treatment_fit` using superlearning (default in the `tmle` package). Superlearning combines many regressions and greatly improves predictions on complex and/or high-dimensional data.

This guide is based on Chapter 4 of Targeted Learning by Mark van der Laan and Sherri Rose. Additional references and a full tutorial on TMLE can be found at: www.khstats.com/blog/tmle/tutorial.

KEY