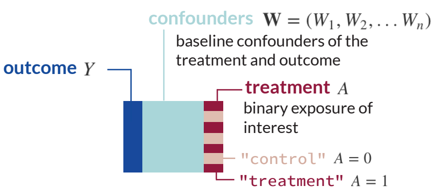
A VISUAL GUIDE TO IPW

Inverse Probability Weighting (IPW) is an **estimation technique for causal inference**. Here we use IPW to estimate the **mean difference in an 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:

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

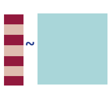
We can rewrite this estimand as:

$$ATE = \mathrm{E}\left[\left(\frac{\mathrm{I}[A=1]}{\mathrm{P}[A=1|\mathbf{W}]} - \frac{\mathrm{I}[A=0]}{\mathrm{P}[A=0|\mathbf{W}]}\right)Y\right]$$

Algorithm:

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

```
treatment_fit ← glm(   ~   )
```



$$\mathrm{P}[A=1|\mathbf{W}]$$

Then, use that model fit to predict two probabilities:

1. Inverse probability of receiving treatment

```
← 1/predict(treatment_fit)
```

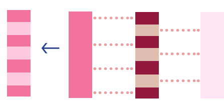
$$\frac{1}{\hat{\mathrm{P}}[A=1|\mathbf{W}]}$$

2. Negative inverse probability of not receiving treatment


```
← -1/(1-predict(treatment_fit))
```

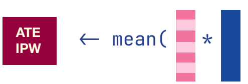
$$\frac{-1}{\hat{\mathrm{P}}[A=0|\mathbf{W}]}$$

Create a new variable: the inverse probability of recieving treatment for observations who received treatment, and the negative inverse probability of not receiving treatment for observations who did not receive treatment.



$$\frac{\mathrm{I}[A=1]}{\hat{\mathrm{P}}[A=1|\mathbf{W}]} - \frac{\mathrm{I}[A=0]}{\hat{\mathrm{P}}[A=0|\mathbf{W}]}$$

Finally, calculate the ATE by multiplying this inverse weight quantity by the outcome.



$$\hat{ATE} = \hat{\mathrm{E}}\left[\left(\frac{\mathrm{I}[A=1]}{\hat{\mathrm{P}}[A=1|\mathbf{W}]} - \frac{\mathrm{I}[A=0]}{\hat{\mathrm{P}}[A=0|\mathbf{W}]}\right)Y\right]$$

The IPW estimator relies on correct specification of the generalized linear model (GLM) for the final estimate to be consistent. Although flexible machine learning models can be used for the treatment regression instead of a GLM, there is generally not statistical theory to support a normal distribution of the estimator for computing confidence intervals, p-values, etc.

kathoffman/causal-inference-visual-guides
@kat_hoffman_