

• The counterfactuals produce, e.g., a new predicted distribution of Y

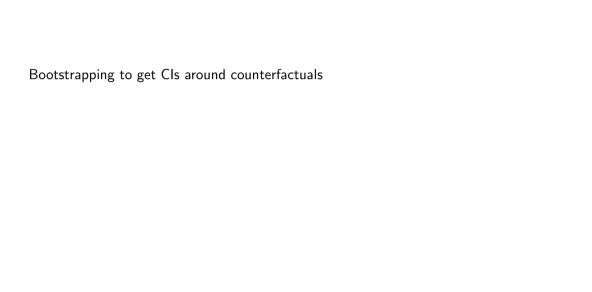
• The counterfactuals produce, e.g., a new predicted distribution of Y

• But we don't know what the confidence intervals are around the prediction

- The counterfactuals produce, e.g., a new predicted distribution of Y
- But we don't know what the confidence intervals are around the prediction
- Our estimates inherently come with sampling variation

- The counterfactuals produce, e.g., a new predicted distribution of Y
- But we don't know what the confidence intervals are around the prediction
- Our estimates inherently come with sampling variation
- We should incorporate that variation into our counterfactual

- The counterfactuals produce, e.g., a new predicted distribution of Y
- But we don't know what the confidence intervals are around the prediction
- Our estimates inherently come with sampling variation
- We should incorporate that variation into our counterfactual
- How do we do this? Bootstrapping



Bootstrapping to get CIs around counterfactuals	
Recall: $-H^{-1}$ is the variance matrix of our estimates, where $H$ is the H	lessian

Bootstrapping to get CIs around counterfactuals

Assume that  $\hat{\theta} \sim MVN(\theta, -H^{-1})$ 

Recall:  $-H^{-1}$  is the variance matrix of our estimates, where H is the Hessian

Bootstrapping to get CIs around counterfactuals

Recall:  $-H^{-1}$  is the variance matrix of our estimates, where H is the Hessian

Assume that  $\hat{\theta} \sim MVN(\theta, -H^{-1})$ 

ullet but remember that we want to minimize  $-\ell$  so we just use  $H^{-1}$ 

$ullet$ We sample $B$ times from $\mathit{MVN}( heta, -H^{-1})$ and compute our new counterfactuals

- We sample B times from  $MVN(\theta, -H^{-1})$  and compute our new counterfactuals
- Then we look at the 2.5th and 97.5th percentiles of cfl statistics

- We sample B times from  $MVN(\theta, -H^{-1})$  and compute our new counterfactuals
- Then we look at the 2.5th and 97.5th percentiles of cfl statistics
- This will represent a 95% confidence interval of that statistic

- ullet We sample B times from  $MVN( heta,-H^{-1})$  and compute our new counterfactuals
- Then we look at the 2.5th and 97.5th percentiles of cfl statistics
- This will represent a 95% confidence interval of that statistic
- This algorithm is known as the parametric bootstrap

- ullet We sample B times from  $MVN( heta,-H^{-1})$  and compute our new counterfactuals
- Then we look at the 2.5th and 97.5th percentiles of cfl statistics
- This will represent a 95% confidence interval of that statistic
- This algorithm is known as the parametric bootstrap
- Contrast with nonparametric bootstrap which is when we randomly re-sample data

- We sample B times from  $MVN(\theta, -H^{-1})$  and compute our new counterfactuals
- Then we look at the 2.5th and 97.5th percentiles of cfl statistics
- This will represent a 95% confidence interval of that statistic
- This algorithm is known as the parametric bootstrap
- Contrast with nonparametric bootstrap which is when we randomly re-sample data
- $\bullet$  P.B. can be intensive if it is costly to conduct counterfactuals  $\to$  rarely used