Confidence intervals around counterfactuals

Confidence intervals around counterfactuals

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

Confidence intervals around counterfactuals

- The counterfactuals produce, e.g., a new predicted distribution of Y
- But we don't know what the confidence intervals are around the prediction

Confidence intervals around counterfactuals

- 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

Confidence intervals around counterfactuals

- 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

Confidence intervals around counterfactuals

- 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

Bootstrapping to get CIs around counterfactuals

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

Bootstrapping to get CIs around counterfactuals

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

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

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

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

- We sample B times from $\mathit{MVN}(\theta, -H^{-1})$ 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

- 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

- 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

- 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
- P.B. can be intensive if it is costly to conduct counterfactuals $\rightarrow$ rarely used