Bootstrap

True distribution $P_{\theta^*}(x)$ $\xrightarrow{\text{i.i.d samples}}$ $S_n = \{X_1, \ldots, X_n\}$ $\xrightarrow{\text{estimator}}$ $\hat{\theta}_n = \hat{\theta}_n(S_n)$.

Goal: Estimate θ^*.

Q. How close is $\hat{\theta}_n$ to θ^*?

Suppose we had an oracle that told us $P_{\theta^*}(x)$, then we also know the distribution of $\hat{\theta}_n$.
In particular, we know $\mathrm{Var}_{P_{\theta^*}(x)}(\hat{\theta}_n)$.

<u>Idea 1</u>: Replace P_{θ^*} with $\hat{P}_n$ = uniform distribution on $\{X_1, \ldots, X_n\}$, i.e. the <u>empirical distribution</u>.

If $\hat{P}_n$ is "close" to P_{θ^*}, then

$$\mathrm{Var}_{P_{\theta^*}}(\hat{\theta}_n) \approx \mathrm{Var}_{\hat{P}_n}(\hat{\theta}_n)$$

$\hat{\theta}_n(X_1^*, \ldots, X_n^*)$.
$X_i^* \overset{\text{i.i.d.}}{\sim} \hat{P}_n$.

<u>Idea 2</u>: Sample from $\hat{P}_n$ in order to get an approximate of the functional / statistic that we want to compute.

In our case, $\widehat{\mathrm{Var}}_{\hat{P}_n}(\hat{\theta}_n) = \frac{1}{B-1}\sum_{j=1}^{B}\left(\hat{\theta}_n^{*(j)} - \bar{\hat{\theta}}_n^{*}\right)^2$

If B is large enough, then

$$\mathrm{Var}_{\hat{P}_n}(\hat{\theta}_n) \approx \widehat{\mathrm{Var}}_{\hat{P}_n}(\hat{\theta}_n)$$

Hence, $\mathrm{Var}_{P_{\theta^*}}(\hat{\theta}_n) \approx \mathrm{Var}_{\hat{P}_n}(\hat{\theta}_n) \approx \underbrace{\widehat{\mathrm{Var}}_{\hat{P}_n}(\hat{\theta}_n)}_{\text{bootstrap estimate}}$.

<u>Rem</u>: ① Idea applies to other functionals of distributions of $\hat{\theta}_n$, like quantile.

② If we use bootstrap to get CIs, then these are only asymptotically valid.

<u>Bootstrap algorithm:</u>

Input: Sample $X_1, \ldots, X_n$, estimator $\hat{\theta}_n$.

Repeat B times:

- Draw dataset $S_n^{(b)*} = \{X_1^{(b)*}, \ldots, X_n^{(b)*}\}$ i.i.d from $\hat{P}_n$.
- Compute $\hat{\theta}_n^{*(b)} = \hat{\theta}_n(S_n^{(b)*})$.

Use the distribution of $\{\hat{\theta}_n^{*(1)}, \ldots, \hat{\theta}_n^{*(B)}\}$ to estimate some statistic that we want.