Double Bootstrap
Improve Accuracy of Bootstrap Coverage

Doug Steigerwald

UC Santa Barbara

January 2010
Observe: random sample \( X = (x_1, \ldots, x_n) \)
Estimate \( \mu \) via \( \hat{\mu} \) (may also have \( \hat{\sigma}^2 \))

**Question:**

- How do we obtain a 95% confidence interval with good coverage (and short distance)?

**Coverage** - probability that interval does contain true parameter value

- For many problems, 95% confidence interval has coverage of far less than 95%
random sample, with replacement, from \((x_1, \ldots, x_n)\), generates

- \(X^* = (x_1^*, \ldots, x_n^*)\)
- \(\hat{\mu}^*, \hat{\sigma}^2, t^* = \frac{\hat{\mu}^* - \hat{\mu}}{\hat{\sigma}^*}\)

construct \(R\) (single) bootstrap samples and order outcomes

- \(\left\{\hat{\mu}^*_{(r)}\right\}_{r=1}^R \quad \left\{t^*_{(r)}\right\}_{r=1}^R\)
Confidence Intervals

R=999

Percentile CI: \[ P \left( \hat{\mu}(25) \leq \mu \leq \hat{\mu}(975) \right) = .95 \]

\[ \left( \hat{\mu}(25), \hat{\mu}(975) \right) \]

Basic CI: \[ P \left( (\hat{\mu}^* - \hat{\mu})(25) \leq \hat{\mu} - \mu \leq (\hat{\mu}^* - \hat{\mu})(975) \right) = .95 \]

\[ \left( \hat{\mu} - (\hat{\mu}^* - \hat{\mu})(975), \hat{\mu} - (\hat{\mu}^* - \hat{\mu})(25) \right) \]

Studentized CI: \[ P \left( t^*(25) \leq \frac{\hat{\mu} - \mu}{\hat{\sigma}} \leq t^*(975) \right) = .95 \]

\[ \left( \hat{\mu} - t^*(975) \hat{\sigma}, \hat{\mu} - t^*(25) \hat{\sigma} \right) \]
Accuracy of Coverage Probabilities

Example: Percentile CI

Theory

\[ P \left( \mu > \hat{\mu} - t_{0.025}^{(25)} \hat{\sigma} \right) = 0.025 \]

often, empirical probability is higher (CI too small)

Finite-Sample Correction

\[ P \left( \mu > \hat{\mu} - t_{q}^{*} \hat{\sigma} | X \right) = 0.025 \]

technically \( q (.025) \) as the correction depends on nominal size for the sample, \( X \), \( q (.025) \) delivers the correct critical value
Finite-Sample Correction
Estimation via Double Bootstrap

Correction $q$ satisfies

$$P \left( \mu > \hat{\mu} - t_q \hat{\sigma} | X \right) = .025$$

Estimate $\hat{q}$ satisfies

$$P \left( \hat{\mu} > \hat{\mu}^* - t_{\hat{q}} \hat{\sigma}^* | X, X^* \right) = .025$$

t** from double bootstrap
random sample, with replacement, from $(x_1^*, \ldots, x_n^*)$, generates

- $X^{**} = (x_1^{**}, \ldots, x_n^{**})$
- $\hat{\mu}^{**}, \hat{\sigma}^{**2}, t^{**} = \frac{\hat{\mu}^{**} - \hat{\mu}^*}{\hat{\sigma}^{**}}$
Finite-Sample Correction
Implementation

Estimate \( \hat{q} \) satisfies

\[
P \left( \hat{\mu} > \hat{\mu}^* - t_{\hat{q}}^* \hat{\sigma}^* | X, X^* \right) = .025
\]

Rewriting

\[
P \left( t^* < t_{\hat{q}}^* | X, X^* \right) = .025
\]

For \( F_{**} \) the CDF of \( t^{**} \)

\[
P \left( t^* < F_{**}^{-1} (\hat{q}) | X, X^* \right) = .025
\]

Inverting yields

\[
P \left( F_{**} (t^*) < \hat{q} | X, X^* \right) = .025
\]
Estimate $\hat{q}$ satisfies

$$P (F_{**} (t^*) < \hat{q} | X, X^*) = .025$$

- $F_{**} (t^*)$ returns a number on $[0, 1]$ that indicates how many of the double bootstrap $t$ values lie below the single bootstrap value $t^*$
- For $R = 999$, we have 999 different values of $F_{**} (t^*)$, only 25 of which should be $\leq \hat{q}$
- Order the values of $F_{**} (t^*)$, $\hat{q}$ is order statistic 25
Finite-Sample Correction
Implementation Details

For each single bootstrap \( r = 1, \ldots, R \)

- randomly draw \( M \) double bootstrap samples and construct \( \{ t_{m}^{**} \}_{m=1}^{M} \)
- construct \( u_{r}^{*} = \frac{1}{M} \sum_{m=1}^{M} 1(t_{rm}^{**} \leq t_{r}^{*}) \)
- order \( \{ u_{r}^{*} \}_{r=1}^{R} \) yielding \( \{ u_{(r)}^{*} \} \)

\[ \hat{q} = u_{(25)}^{*} \]

Nankervis (2005) has efficient algorithm