Subset Bootstrap

ZHANG Yan

March 30, 2021

Original bootstrap:

1. draw $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ from F, and obtain $\hat{\theta} = \hat{\theta}(\boldsymbol{X})$;
2. resample $\boldsymbol{X}^{*b} = (X_1^{*b}, X_2^{*b}, \ldots, X_n^{*b})$ from $\boldsymbol{X}$, $b = 1, \ldots, B$;
3. obtain $\hat{\boldsymbol{\theta}}^* = (\hat{\theta}^{*1}, \hat{\theta}^{*2}, \ldots, \hat{\theta}^{*B})$, where $\hat{\theta}^{*b} = \hat{\theta}(\boldsymbol{X}^{*b})$;
4. estimate the $1-\alpha$ confidence interval: $[2\hat{\theta} - \hat{\boldsymbol{\theta}}^*_{(1-\alpha/2)}, 2\hat{\theta} - \hat{\boldsymbol{\theta}}^*_{(\alpha/2)}]$.

Subset bootstrap: (assume the bias is asymptotically negligible, and $STD(\hat{\theta}) \propto n^{-\beta}$, normally $\beta = 1/2$)

1. draw $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ from F, and obtain $\hat{\theta} = \hat{\theta}(\boldsymbol{X})$;

2. set $0 < \gamma < 1$ so that γn is an integer, resample (with or without replacement) $\boldsymbol{X}^{*b}_\gamma = (X_1^{*b}, X_2^{*b}, \ldots, X_{\gamma n}^{*b})$ from $\boldsymbol{X}$, $b = 1, \ldots, B$;

3. obtain $\hat{\boldsymbol{\theta}}^*_\gamma = (\hat{\theta}^{*1}_\gamma, \hat{\theta}^{*2}_\gamma, \ldots, \hat{\theta}^{*B}_\gamma)$, where $\hat{\theta}^{*b}_\gamma = \hat{\theta}(\boldsymbol{X}^{*b}_\gamma)$;

4. scale $\hat{\boldsymbol{\theta}}^*_\gamma$ to be $\hat{\boldsymbol{\theta}}^* = \gamma^\beta(\hat{\boldsymbol{\theta}}^*_\gamma - \hat{\theta}) + \hat{\theta}$;

5. estimate the $1-\alpha$ confidence interval: $[2\hat{\theta} - \hat{\boldsymbol{\theta}}^*_{(1-\alpha/2)}, 2\hat{\theta} - \hat{\boldsymbol{\theta}}^*_{(\alpha/2)}]$.

Experiment:

Here I set $n = 2000$, $B = 1000$, $\gamma = 0.1$ and $\alpha = 0.1$ with 100 replicates to estimate the probabilities that the confidence intervals contain the target parameters.

	normal (one mode)	gamma	normal (two mode)
mean	0.87/0.86/0.84	0.87/0.88/0.86	0.91/0.92/0.9
median	0.89/0.88/0.88	0.88/0.89/0.88	**0.81/0.7/0.68**
std	0.93/0.93/0.92	0.93/0.92/0.93	0.86/0.87/0.85
variance	0.91/0.91/0.89	0.91/0.92/0.91	0.9/0.88/0.87

Table 1: Original bootstrap/Subset with replacement/Subset without replacement