Hybrid parametric bootstrap for design-based variance

Alana McGovern

2025-10-13

1. Using the real data, obtain design-based mean estimates at desired level and denote $\hat{\boldsymbol{\mu}}$.
2. Compute the residuals, $\delta_{jck} := y_{jck} - \hat{\mu}_j$ where y_{jck} is the response of individual, k, in cluster, c, in area, j.
3. Fit the following model

$$\delta_{jck} = \epsilon_c + \gamma_k$$

where $\epsilon_c \sim \mathcal{N}(0, \sigma_\epsilon^2)$ and $\gamma_c \sim \mathcal{N}(0, \sigma_\gamma^2)$. Record the hyperparameters estimates, $\hat{\sigma}_\epsilon^2$ and $\hat{\sigma}_\gamma^2$.

Now for each bootstrap sample, b,

4. Draw $\epsilon_c^{(b)} \sim \mathcal{N}(0, \hat{\sigma}_\epsilon^2)$ and $\gamma_k^{(b)} \sim \mathcal{N}(0, \hat{\sigma}_\gamma^2)$ for each cluster c and individual k and use to construct new set of samples by constructing $y_{jck}^{(b)} := \hat{\mu}_j + \epsilon_c^{(b)} + \gamma_k^{(b)}$.
5. Use $\mathbf{y}^{(b)}$ to obtain design-based mean estimates, $\hat{\boldsymbol{\mu}}^{(b)}$

After this has been done for all samples,

6. Obtain bootstrap estimate of the design-based variance for each area j, by

$$V_j^B = \frac{1}{B-1} \sum_b (\hat{\mu}_j^{(b)} - \frac{1}{B} \sum_{b'} \hat{\mu}_j^{(b')})^2$$