Evaluation of bootstrap methods in nonlinear mixed-effects models

Hoai -Thu Thai\(^1,2,\)
France Mentré\(^1\), Nicholas N.G Holford\(^3\), Christine Veyrat-Follet\(^2\), Emmanuelle Comets\(^1\)

\(^1\)INSERM, UMR 738, Univ Paris Diderot, Sorbonne Paris Cité, F-75018 Paris, France
\(^2\)Drug Disposition Department, Sanofi, France
\(^3\)Department of Pharmacology and Clinical Pharmacology, University of Auckland, Auckland, New Zealand

PAGE MEETING
Glasgow, 14 June 2013
Principle of bootstrap (Efron 1979)\(^1\)

- Resample with replacement the observed data to construct the distribution of an estimate or a statistic of interest
- A general bootstrap procedure in regression:

  - Original data
  - Bootstrap sample 1
    - Parameter estimate 1
  - Bootstrap sample 2
    - Parameter estimate 2
  - Bootstrap sample B
    - Parameter estimate B

- “Main” bootstrap methods
  - case bootstrap\(^2\): resample the pairs
  - residual bootstrap\(^1\): resample the residuals after model fitting


«Pull one self up with your own bootstrap»
Bootstrap in mixed-effects models (MEM)

MEM model

\[ y_i = f(\xi_i, \mu, \eta_i) + g(\xi_i, \mu, \eta_i, \sigma)\varepsilon_i \]

- \( \mu \): fixed parameters
- \( \eta_i \sim N(0, \Omega) \)
- \( \varepsilon_i \sim N(0,1) \)

- Take into account the characteristics of MEM\(^1\): repeated measures, heteroscedasticity, nonlinearity
- Respect two levels of variability: interindividual and residual variabilities\(^2,3\)

- Previous simulation study in linear MEM to compare bootstrap methods

(THAI HT et al. Pharm Stat 2013; 12 (3): 129-140)

- poor performance of bootstraps resampling only one level of variability
- good performance of three bootstrap methods resampling two levels of variability
- some differences between bootstrap methods when applied to a real unbalanced dataset

Objective

Evaluate the performance of bootstrap methods for estimating uncertainty of parameters in nonlinear mixed-effects models (NLMEM) using a simulation study
Introduction

Objective

Methods

Results

Conclusions

Bootstrap methods

- Case bootstrap \( (B_{\text{case}}) \)

- Nonparametric random effect and residual bootstrap \( (B_{\eta,\epsilon}^{NP}) \)

- Parametric random effect and residual bootstrap \( (B_{\eta,\epsilon}^{P}) \)

Resample the individuals

\[
(\xi_i, y_i) \rightarrow (\xi_i^*, y_i^*)
\]
Bootstrap methods

- Case bootstrap ($B_{\text{case}}$)
- Nonparametric random effect and residual bootstrap ($B_{\eta,\varepsilon}^{NP}$)
- Parametric random effect and residual bootstrap ($B_{\eta,\varepsilon}^{P}$)

Fit the model to the data and estimate random effects $\hat{\eta}_i$ and standardized residuals $\hat{\varepsilon}_{ij}$

Correct for shrinkage in random effects and residuals

Resample the random effects
Resample the residuals

Generate the bootstrap observations

$$y_i^* = f(\xi_i, \hat{\mu}, \hat{\eta}_i^*) + g(\xi_i, \hat{\mu}, \hat{\eta}_i^*, \hat{\sigma})\varepsilon_i^*$$

Bootstrap methods

- Case bootstrap ($B_{\text{case}}$)
- Nonparametric random effect and residual bootstrap ($B_{\eta,\epsilon}^{\text{NP}}$)
- Parametric random effect and residual bootstrap ($B_{\eta,\epsilon}^{\text{P}}$)

Fit the model to the data

Simulate the random effects from $\mathcal{N}(0, \hat{\Omega})$
Simulate the residuals from $\mathcal{N}(0,1)$

Generate the bootstrap observations
$$y_i^* = f(\xi_i, \hat{\mu}, \hat{\eta}_i) + g(\xi_i, \hat{\mu}, \hat{\eta}_i^*, \hat{\sigma})\epsilon_i^*$$
Motivating example

- **Pharmacokinetic data: aflibercept**, an anti-angiogenic agent
  - 2 clinical trials
    - phase I TCD 6120 trial (first dose) \((N_1=53, n_1=\text{median of 9})\)
    - phase III VITAL trial (first two doses) \((N_2=291, n_2=2)\)
  - two-compartment PK model with first-order elimination
    - exponential model for IIV
    - proportional model for residual error

- **Model fit to the data: SAEM (MONOLIX 4.1.2)**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (RSE)</th>
<th>IIV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL (L/hr)</td>
<td>0.04 (2)</td>
<td>28.8 (5)</td>
</tr>
<tr>
<td>V1 (L)</td>
<td>3.62 (2)</td>
<td>19.7 (10)</td>
</tr>
<tr>
<td>Q (L/hr)</td>
<td>0.14 (15)</td>
<td>111(12)</td>
</tr>
<tr>
<td>V2 (L)</td>
<td>2.9 (5)</td>
<td>-</td>
</tr>
<tr>
<td>Corr(CL,Q) (ρ)</td>
<td>0.90 (8)</td>
<td></td>
</tr>
<tr>
<td>σp (%)</td>
<td>24.8 (4)</td>
<td></td>
</tr>
</tbody>
</table>

Simulation settings

**Frequent balanced design**
- First-order
- \(N=30, n=9\)

**Sparse balanced design**
- First-order
- \(N=68, n=4\)

**Frequent balanced design**
- Mixed-order (Michaelis-Menten)
- \(N=30, n=9\)

**Unbalanced design**
- First-order
- \(N_1=15, n_1=9\)
- \(N_2=75, n_2=2\)

\(K=100\) replications for each design
Bootstrap settings

- **K=100 simulations**
  - Simulated dataset

- **B=1000 bootstrap datasets**
  - $B_{\text{case}}$
  - $B_{\eta, \varepsilon}^{\text{NP}}$
  - $B_{\eta, \varepsilon}^{\text{P}}$

- **100x1000 estimations**
  - $1000 \hat{\theta}_{\text{case}}$
  - $1000 \hat{\theta}_{\text{NP}}$
  - $1000 \hat{\theta}_{\text{P}}$

- **Estimation method:** SAEM (Monolix 4.1.2)
  - Fisher information matrix ($M_F$) computed by linearisation

- **Asymptotic $M_F$ method (Asym)**
  - $\hat{\theta}_{\text{obs}}$
  - $\text{SE}_{\text{obs}}$
  - $\text{CI}_{\text{obs}}$

- **Bootstrap**
  - $\hat{\theta}_B$
  - $\text{SE}_B$
  - $\text{CI}_B$

- **Median**
- **Standard deviation**

- **95% CI=2.5 - 97.5th percentile**

---

Hoai Thu Thai

14/06/2013
Bootstrap settings for the unbalanced design

- **Introduction**
- **Objective**
- **Methods**
- **Results**
- **Conclusions**

K=100 simulations

Simulated dataset

B=1000 bootstrap datasets

B\_case

B\_strat

B\^{NP}\_\eta,\epsilon

B\^{NP, strat}\_\eta,\epsilon

B\^{P}\_\eta,\epsilon

100x1000 estimations

1000 \( \hat{\theta}_{\text{case}} \)

1000 \( \hat{\theta}_{\text{strat}} \)

1000 \( \hat{\theta}_{\text{NP}} \)

1000 \( \hat{\theta}_{\text{strat, strat}} \)

1000 \( \hat{\theta}_{\text{P}} \)

\( \hat{\theta}_{\text{obs}} \)

SE\_\text{obs}

CI\_\text{obs}

100 estimations

Asymptotic M\_F method (Asym)

* Stratification on rich/sparse data

SE\_\text{empirical}
Evaluation

- For each method (k=100 replicated datasets)
  - Relative bias of parameter (%)
    \[
    \text{RBias}(\hat{\theta}^{(l)}_B) = \frac{1}{K} \sum_{k=1}^{K} \left( \frac{\hat{\theta}^{(l)}_{B:k} - \hat{\theta}^{(l)}_{k}}{\hat{\theta}^{(l)}_{k}} \right) \times 100
    \]
  - Relative bias of SE (%)
    \[
    \text{RBias}(\hat{SE}^{(l)}_B) = \frac{1}{K} \sum_{k=1}^{K} \left( \frac{\hat{SE}^{(l)}_{B:k} - \hat{SE}^{(l)}_{\text{empirical}}}{\hat{SE}^{(l)}_{\text{empirical}}} \right) \times 100
    \]
  - Coverage rate of 95% CI: % bootstrap CI contain the true value of parameter

- “Good” bootstrap
  - Relative bias of parameters and SE: unbiased (±10%)
  - Coverage rate of 95% CI: good (90-100%)
Frequent balanced design with first-order elimination

- **Introduction**
- **Objective**
- **Methods**
- **Results**
- **Conclusions**

- Parametric bootstrap had the best performance
- Asymptotic approach performed slightly less well than the bootstrap methods
Three balanced designs

**Frequent design**
- first order
- (N=30, n=9)

**Sparse design**
- first order
- (N=68, n=4)

**Frequent design**
- mixed-order
- (N=30, n=9)
Unbalanced design with first-order elimination

Asymptotic approach and parametric bootstrap performed reasonably well and better than other bootstrap methods.
Unbalanced design with first-order elimination

Stratification improved slightly the performance of case bootstrap and but degraded that of nonparametric random effect and residual bootstrap in term of coverage rate.
Conclusions

- Better estimation of uncertainty by the bootstrap methods than the asymptotic method in NLMEM with high nonlinearity
- Caution with bootstrap methods in presence of large fluctuations in parameter estimates between bootstrap samples
- The choice of bootstrap methods in NLMEM
  - **parametric bootstrap**: best description of uncertainty
    - study robustness in case of model misspecification
  - **case bootstrap**: fast, simple and robust (e.g. heteroscedasticity, missing data)
    - evaluate stratification for complex designs
  - **nonparametric random effect and residual bootstrap**: maintain the same design as the original dataset
    - improve correction for shrinkage in unbalanced designs
Acknowledgements

- Sanofi for financial support
Annexes
Correction for variance underestimation

- Correction for random effects\(^1,2\)
  - center: \( \tilde{\eta}_i = \hat{\eta}_i - \bar{\eta}_i \)
  - transform: by the ratio between empirical vs estimated variance-covariance (\(A_\eta\))
    - Cholesky decomposition: matrice \(\Omega_R\) positive
    - EVD (Eigen Value Decomposition): matrice \(\Omega_R\) semi-positive
      \[
      \hat{\eta}'_i = \tilde{\eta}_i \times A_\eta
      \]

- Correction for residuals\(^1,2\)
  - center: \( \tilde{\varepsilon}_{ij} = \hat{\varepsilon}_{ij} - \bar{\varepsilon}_{ij} \)
  - transform: by the ratio between empirical vs estimated variance-covariance (\(A_\sigma\))
    - homoscedastic error: \( A_\sigma = \hat{\sigma} / sd(\tilde{\varepsilon}_{ij}) \)
    - heteroscedastic error: \( A_\sigma = 1 / SD(standardized \ residuals) \)
      \[
      \hat{\varepsilon}'_{ij} = \tilde{\varepsilon}_{ij} \times A_\sigma
      \]

Application to aflibercept data
Asym

B_{case,none}

B_{\eta,GR}

B_{p,PR}