Bootstrap Inference for Multiple Imputation Under Uncongeniality and Misspecification

Jonathan W. Bartlett [1]   Rachael A. Hughes [2]

[1]Department of Mathematical Sciences
University of Bath, UK
www.thestatsgeek.com

[2]Bristol Medical School, University of Bristol

ISCB 25th August 2020

Acknowledgement

This research made use of the Balena High Performance Computing (HPC) Service at the University of Bath.

Paper

The work I will describe has just been published in SMMR:

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Article

Statistical Methods in Medical Research
0(0) 1–14

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0962280220932189
journals.sagepub.com/home/smm
SAGE

Bootstrap inference for multiple imputation under uncongeniality and misspecification

Jonathan W Bartlett[1] and **Rachael A Hughes**[2,3]

Abstract
Multiple imputation has become one of the most popular approaches for handling missing data in statistical analyses. Part of this success is due to Rubin's simple combination rules. These give frequentist valid inferences when the imputation

Outline

Motivation

- MI is very popular, for many reasons, part of which are the simplicity of Rubin's rules.
- If imputation and analysis models are 'congenial', and the models are correctly specified, Rubin's rules give valid frequentist inferences (asymptotically):
  - Rubin's variance estimator is unbiased
  - Confidence intervals attain nominal coverage

Motivation

- Uncongeniality can arise in different ways, for example:
  - the analysis model is only fitted to a subgroup
  - one of the models includes an interaction but the other does not
  - reference based imputation in clinical trials
- Uncongeniality and model misspecification often lead to the MI point estimator being **biased**.
- But there are situations where MI point estimator is **unbiased** despite uncongeniality and/or misspecification.
- Under uncongeniality or misspecification, Rubin's variance estimator can be biased upwards or downwards, even when the point estimator is unbiased [3, 4].

Bootstrapping and MI

- In recent years a number of papers have investigated different ways of combining bootstrapping with MI to produce confidence intervals (CI):
  - Schomaker and Heumann 2018 [5]
  - Brand *et al* 2019 [1]
  - von Hippel and Bartlett 2019 [6]
- We investigated CI length and coverage of the methods recommended in these papers under uncongeniality or misspecification in situations when despite these, the MI point estimator is unbiased.

Outline

Motivation

Rubin's rules and congeniality

Impute then bootstrap approaches

Bootstrap then impute approaches

Simulations

Conclusions

MI and Rubin's rules

- The imputer creates M imputations of the missing data, using some model
- The analyst applies some complete data procedure to each, obtaining estimates of some quantity of interest θ, which we denote $\hat{\theta}_m$, $m = 1, \ldots, M$, and variance estimates $\widehat{Var}(\hat{\theta}_m)$
- MI point estimator is $\overline{\theta}_M = \frac{1}{M}\sum_{m=1}^{M} \hat{\theta}_m$
- Rubin's variance estimator is $T_M = W_M + \left(1 + \frac{1}{M}\right) B_M$ where

$$W_M = \frac{1}{M}\sum_{m=1}^{M} \widehat{Var}(\hat{\theta}_m)$$

$$B_M = \frac{1}{M-1}\sum_{m=1}^{M} (\hat{\theta}_m - \overline{\theta}_M)^2$$

Congeniality - Meng

- Meng 1994 [3] defined the notion of 'congeniality' between imputation model and the analyst's incomplete and complete data 'procedures'.
- Imputation model and analyst's complete data procedure are congenial if:
  - there exists a Bayesian model for which given complete data, the posterior mean and variance for θ match the point and variance estimate from the analyst's complete data procedure
  - the predictive distribution for the missing data given observed from this Bayesian model is identical to that of the imputation model being used

Implications of congeniality

- Under congeniality, you can show that
  - the posterior mean of θ given observed data under the congenial Bayesian model equals $\lim_{M\to\infty} \overline{\theta}_M = \overline{\theta}_\infty$
  - the posterior variance of θ given observed data under the Bayesian model equals $\lim_{M\to\infty} T_M = T_\infty$
- Therefore $(\overline{\theta}_\infty, T_\infty)$ gives Bayesian posterior mean and variance under this Bayesian model.
- Therefore for infinite M, Rubin's rules = Bayesian inference, and if the model is correct, asymptotically point estimator and variance estimator are consistent.
- Rubin's rules then make adjustments for the fact M is finite.

Outline

Motivation

Rubin's rules and congeniality

Impute then bootstrap approaches

Bootstrap then impute approaches

Simulations

Conclusions

MI boot Rubin

1. Impute M times
2. For $m = 1, .., M$, generate B nonparametric bootstraps
3. $\hat{\theta}_{m,b}$ estimate from imputation m, bootstrap b
4. For imputation m, complete data variance estimated by

$$\widehat{\text{Var}}_{bs}(\hat{\theta}_m) = (B-1)^{-1} \sum_{s=1}^{B} (\hat{\theta}_{m,b} - \tilde{\theta}_m)^2$$

   where $\tilde{\theta}_m = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_{m,b}$

5. Rubin's rules applied to $\hat{\theta}_m$ and $\widehat{\text{Var}}_{bs}(\hat{\theta}_m)$, $m = 1, .., M$

Inference is based on Rubin's rules, so we don't in general expect valid inferences under uncongeniality or misspecification.

MI boot pooled

Percentile interval based on values $\hat{\theta}_{m,b}$, $m = 1, \ldots, M$, $b = 1, \ldots, B$.

Under congeniality, M imputations are M independent draws from posterior of missing data given observed under the Bayesian model.

Assuming analysis model is MLE, B bootstraps on each imputed dataset is equivalent to B draws from posterior of θ conditional on each imputed complete dataset.

Thus under congeniality the $M \times B$ values of θ are draws from the posterior of θ given observed data.

MI boot pooled

So under congeniality, MI boot pooled percentile interval = posterior credible interval for Bayesian model.

B is normally chosen large. If you choose M small, you can show the variance across $\hat{\theta}_{m,b}$ is too small.

With $M = 1$, your posterior variance is for the one time imputed complete data, not the observed data posterior variance!

Under uncongeniality or misspecification, no reason why the CI should have correct coverage.

Outline

Motivation

Rubin's rules and congeniality

Impute then bootstrap approaches

Bootstrap then impute approaches

Simulations

Conclusions

Boot MI percentile

1. Bootstrap B times
2. For $b = 1, .., B$, impute M times, and obtain estimates $\hat{\theta}_{b,m}$
3. Let $\overline{\theta}_b = \frac{1}{M} \sum_m \hat{\theta}_{b,m}$
4. Form percentile intervals based on $\overline{\theta}_b$

This is just application of bootstrap to the MI estimator, so we should expect correct coverage even under uncongeniality or misspecification.

Boot MI with $M = 1$

Brand *et al* 2019 [1] investigated many different combination methods (including a number not mentioned here) as point estimator.

They recommended Boot MI percentile confidence intervals with $M = 1$.

The problem with this is that the point estimator and interval are inefficient with $M = 1$.

Thus the intervals are wider than they need to be.

In fact, we discovered an additional curious issue with this approach when you use percentile rather than Wald type intervals (see sim results and paper).

Boot MI for inference under uncongeniality

Boot MI is the only approach we expect to give CIs with correct coverage under uncongeniality/misspecification (assuming point estimator is unbiased).

We need relatively large B for reliable estimates of variance.

If we choose M small, point estimator is inefficient and intervals are wider than necessary.

If we choose M large, $B \times M$ is large, and Boot MI is computationally costly!

von Hippel's Boot MI proposal

von Hippel [6] proposed using Boot MI, with $\overline{\theta}_{BM} = B^{-1} \sum_{b=1}^{B} \overline{\theta}_b$.

We can express

$$\hat{\theta}_{b,m} = \overline{\theta}_\infty + c_b + d_{bm}$$

where $Var(c_b) = Var(\overline{\theta}_\infty) = \sigma^2_\infty$ and $Var(d_{bm}) = \sigma^2_{\text{btw}}$

Then

$$\overline{\theta}_{BM} = \overline{\theta}_\infty + \frac{1}{B} \sum_{b=1}^{B} c_b + \frac{1}{BM} \sum_{b=1}^{B} \sum_{m=1}^{M} d_{bm}$$

and so

$$\text{Var}(\overline{\theta}_{BM}) = \left(1 + \frac{1}{B}\right) \sigma^2_\infty + \frac{1}{BM} \sigma^2_{\text{btw}}$$

von Hippel's boot MI proposal

We can fit a one way random intercepts model to the estimates $\hat{\theta}_{b,m}$ to estimate σ^2_{∞} and σ^2_{btw}, and insert into the preceding expression.

Since large B is required for reliable variance estimates, von Hippel suggested using $M = 2$. With $M = 2$, the approach becomes computationally much less costly.

Derivations do not rely on assumptions of congeniality or correct specification, so variance estimator should be consistent even under uncongeniality or misspecification.

Outline

Simulation setup

I will briefly talk about simulations on reference based imputation for trials. See the paper for more simulation scenarios.

Sample size $n = 500$.

Binary ‘treatment’ randomly assigned.

Y_1, Y_2 (baseline,follow-up) generated from correlated bivariate normal, with mean of Y_2 dependent on ‘treatment’.

50% of Y_2 values made missing completely at random.

Analysis model is linear regression of Y_2 on treatment and Y_1, and interest focuses on the treatment coefficient.

10,000 simulations

Imputation methods

First we imputed Y_2 using normal linear regression under MAR (see paper).

Next we impute Y_2 using the jump to reference approach, proposed by Carpenter *et al* [2].

This imputes active arm patients Y_2 assuming they switched to control treatment (in a particular way).

The imputation model is uncongenial with the analysis model.

Jump to reference results

	M	B	Time (s)	Median CI width	CI cov.
MI Rubin	10		0.05	0.251	99.78
MI boot Rubin	10	1000	13.6	0.251	99.78
MI boot pooled	10	1000	13.7	0.237	99.63
Boot MI %	10	1000	36.8	0.157	96.06
Boot MI %	1	1000	3.9	0.211	99.40
von Hippel	2	1000	7.6	0.151	95.26

Outline

Conclusions

- Uncongeniality and misspecification often imply the MI estimator is biased.
- But in some cases the MI estimator can be unbiased even under uncongeniality and/or misspecification. In these cases we may want to obtain sharp valid inferences.
- MI then bootstrap is not valid generally under uncongeniality or misspecification.
- In contrast, certain types of bootstrap then MI are valid.
- von Hippel's boot MI is attractive on computational efficiency grounds.
- It is implemented in the R package **bootImpute**, including parallel core functionality.

References I

[1] J Brand, S van Buuren, S le Cessie, and W van den Hout.
Combining multiple imputation and bootstrap in the analysis of cost-effectiveness trial data.
Statistics in Medicine, 38(2):210–220, 2019.

[2] J R Carpenter, J H Roger, and M G Kenward.
Analysis of longitudinal trials with protocol deviations: a framework for relevant, accessible assumptions and inference via multiple imputation.
Journal of Biopharmaceutical Statistics, 23:1352–1371, 2013.

[3] X L Meng.
Multiple-imputation inferences with uncongenial sources of input (with discussion).
Statistical Science, 10:538–573, 1994.

[4] J M Robins and N Wang.
Inference for imputation estimators.
Biometrika, 85:113–124, 2000.

[5] M Schomaker and C Heumann.
Bootstrap inference when using multiple imputation.
Statistics in Medicine, 37(14):2252–2266, 2018.

[6] Paul T. von Hippel and Jonathan W. Bartlett.
Maximum likelihood multiple imputation: Faster imputations and consistent standard errors without posterior draws.
ArXiv e-prints, 2019.
1210.0870v10.