Vehtari et al 2020
Rank-normalization, folding, and localization

Daniele Zago

March 21, 2021

Summary

The authors show that traditional $\hat{R}$ and split-$\hat{R}$ statistics of Gelman and Rubin, 1992 and Gelman, Carlin, et al., 2013 has serious flaws when

1. The chain *tails are thick* which produce a variance that is either too large or infinite.
2. The variance *varies across chains*, but they have the same mean.

Therefore, they introduce an improved version of the statistic, the *folded-split-*$\hat{R}$, which alleviates some of the problems of the previous methods. This method is based on the computation of multiple chains, and returning the maximum of two split-$\hat{R}$ statistics, in order to monitor the within and between variances with and without the effect of different scales.

1 Background

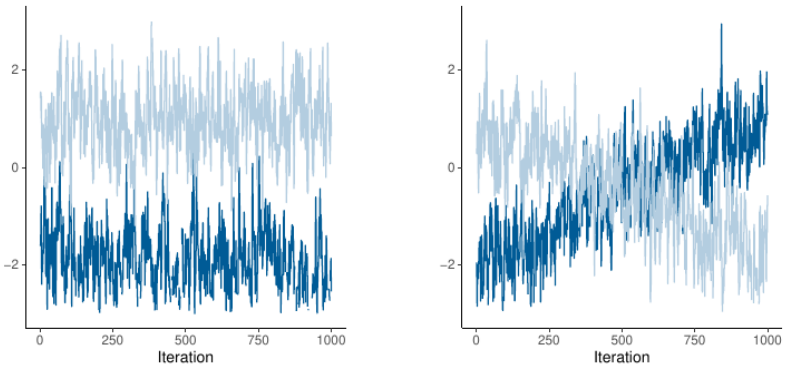

Figure 1: Information should be compared *within* and *between* simulated chains.

The usual split-$\hat{R}$ for M chains of size N is defined as

$$\hat{R} = \sqrt{\frac{\frac{N-1}{N}W + \frac{1}{N}B}{W}},$$

where B is the between-chain variance and W is the within-chain variance:

$$
\begin{aligned}
B &= \frac{N}{M-1} \sum_{m=1}^{M} \left(\bar{\vartheta}_{\cdot m} - \bar{\vartheta}_{\cdot\cdot} \right)^2, \\
W &= \frac{1}{M} \sum_{m=1}^{M} \frac{1}{N-1} \sum_{n=1}^{N} \left(\vartheta_{nm} - \bar{\vartheta}_{\cdot m} \right)^2.
\end{aligned}
\tag{1}
$$

2 Contribution

The improved $\hat{R}$ is computed by:

1. Calculating a rank normalized split-$\hat{R}$, by applying Equation 1 to

$$z_{nm} = \Phi^{-1} \left(\frac{\operatorname{rank} \vartheta_{nm} - 3/8}{NM - 1/4} \right),$$

   where ranks are pooled from all chains.

2. Calculating the rank normalized split-$\hat{R}$ not only for ϑ_{nm} but also for

$$\zeta_{mn} = |\vartheta_{nm} - \operatorname{median}(\vartheta)|$$

   in order to be robust to different scales.

3. Returning the maximum of the quantities obtained in steps 1) and 2).

Equation 1 to the folded draws

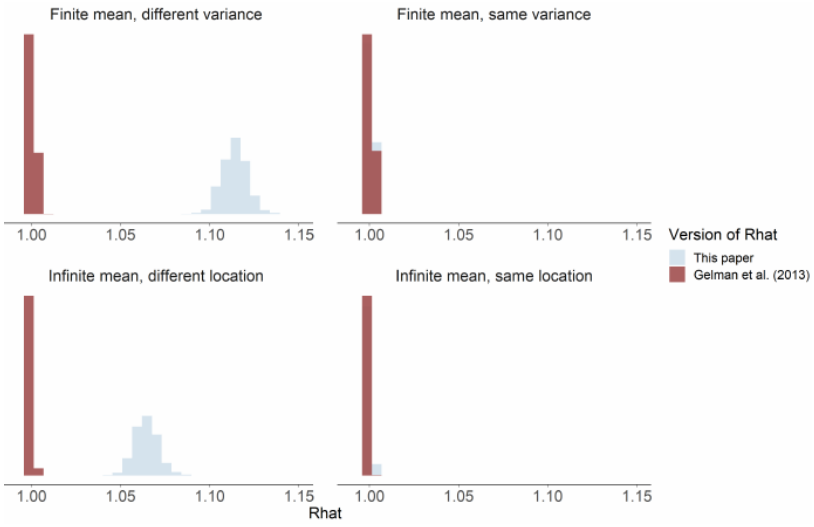

Figure 2: Usual $\hat{R}$ (*red*) and the improved $\hat{R}$ introduced in this paper (*blue*); the two scenarios on the left are constructed not to be mixed, and the usual split-$\hat{R}$ fails to diagnose the lack of convergence.

References

Gelman, A., Carlin, J. B., et al. (2013). *Bayesian data analysis*. 3 edizione. Chapman and Hall/CRC.

Gelman, A. and Rubin, D. B. (1992). "Inference from Iterative Simulation Using Multiple Sequences". In: *Statistical science* 7.4, pp. 457–472.