JOURNAL OF COMPUTATIONAL AND GRAPHICAL STATISTICS
2017, VOL. 0, NO. 0, 1
https://doi.org/10.1080/10618600.2017.1377082

Taylor & Francis
Taylor & Francis Group

CORRIGENDUM

Correction to Cook, Gelman, and Rubin (2006)

Andrew Gelman

Department of Statistics and Department of Political Science, Columbia University

We thank Sean Talts and Michael Betancourt for sharing an example that made us realize the incorrectness of the following statement in Cook, Gelman, and Rubin (2006):

> Let $q_i = \frac{1}{L}\sum_{\ell=1}^{L} I_{\theta_i^{(0)} > \theta_i^{(\ell)}}$, the empirical quantile for the ith replication. For any generic function h, we can determine the distribution of $h(q)$ for correctly working software. In particular, if the software works properly and therefore the posterior quantiles are uniformly distributed, then $h(q) = \Phi^{-1}(q)$ should have a standard normal distribution, where Φ represents the standard normal CDF.

The above claim is false, for two reasons. First, the relevant reference distribution is discrete uniform, not continuous uniform, so the normal CDF is at best just an approximation. Second, with Markov chain simulation, the draws $\theta^{(\ell)}$ are dependent, so for any finite L, the distribution of q_i will not even be discrete uniform.

The error wasn't noticed in the original paper because the method happened to work out on the examples. In general, however, any claim of the statistical properties of any statement about the distribution of draws from iterative simulation should account for the dependence of these simulation draws.

References

Cook, S., Gelman, A., and Rubin, D. B. (2006). Validation of software for Bayesian models using posterior quantiles. *Journal of Computational and Graphical Statistics* **15**, 675–692. [1]

CONTACT