Predictive Inference Based on Markov Chain Monte Carlo Output

Fabian Krüger[1], **Sebastian Lerch**[2,*], **Thordis L. Thorarinsdottir**[3], **Tilmann Gneiting**[2]

[1]*Alfred-Weber-Institute for Economics, Heidelberg University, Heidelberg, Germany*
[2]*Institute for Stochastics, Karlsruhe Institute of Technology, Karlsruhe, Germany*
[3]*Norwegian Computing Center, Oslo, Norway*
[*]Email: `Sebastian.Lerch@kit.edu`

In Bayesian inference, predictive distributions are typically in the form of samples generated via Markov chain Monte Carlo or related algorithms. We conduct a systematic analysis of how to make and evaluate probabilistic forecasts from such simulation output. Based on proper scoring rules, we develop a notion of consistency that allows to assess the adequacy of methods for estimating the stationary distribution underlying the simulation output. We then provide asymptotic results that account for the salient features of Bayesian posterior simulators, and derive conditions under which choices from the literature satisfy our notion of consistency. Importantly, these conditions depend on the scoring rule being used, such that the choices of approximation method and scoring rule are intertwined. While the logarithmic rule requires fairly stringent conditions, the continuous ranked probability score yields consistent approximations under minimal assumptions. These results are illustrated in a simulation study and an economic data example. Overall, mixture-of-parameters approximations which exploit the parametric structure of Bayesian models perform particularly well.