A discussion of “Martingale posterior distributions” by Fong, Holmes and Walker

Published version:

DOI:10.1093/jrsssb/qkad095

(Article begins on next page)

DISCUSSION

Journal Section

Contributed Discussion of Martingale Posterior Distributions by E. Fong, C. Holmes and S. Walker

Marta Catalano[1] | Augusto Fasano[2] | Giovanni Rebaudo[3]

[1]University of Warwick, Coventry, UK

[2]Collegio Carlo Alberto, Torino, IT

[3]University of Texas at Austin, TX, USA

Correspondence
Email: rebaudo.giovanni@gmail.com

We congratulate the authors for an insightful foundational contribution to the statistical literature. This work builds on well-established methodologies, such as predictive inference, Doob's theorem, and conditional independent sequences, providing a unifying framework and a novel understanding of Bayesian uncertainty. A central role is played by the predictive characterization of the random parameter in terms of observable (yet unobserved) quantities, which are regarded as the root of all statistical uncertainty. This brings to a new approach to inference, termed martingale posterior, which sheds light on interesting connections between Bayesian and frequentist statistics. Indeed, both leverage on an empirical distribution: the former builds it through the predictive distributions, the latter through independent and identically distributed samples. Importantly, martingale posteriors go beyond some common homogeneity assumptions in the data, such as infinite exchangeability. Moreover, this approach may offer practical advantages both in terms of prior elicitation and computations, as illustrated by the authors in some relevant scenarios.

We believe that the breadth of this contribution will inspire several future research questions. Here, we restrict our attention to two aspects that we found particularly interesting. First, the authors underline that inference and prediction under a martingale posterior might depend on the order of the data, even when a natural order does not exist. Such dependence will vanish as the sample size increases, but can still be relevant for finite samples. To overcome such an issue, the authors suggest using the average of the predictive distribution over M random permutations of the sample (e.g., $M = 10$). We believe that, in such a scenario, it could be useful to define and study predictive rules that go beyond infinite exchangeability yet preserve finite exchangeability for any fixed sample size.

Second, from a theoretical and modeling perspective, it is often relevant to establish frameworks that weaken the homogeneity assumption of infinite exchangeability, while still preserving well-defined limits as the sample size increases and tractable learning updates. The authors rely on an additional principle: the martingale predictive coherence. We believe that it could be of interest to relax this principle to study other classes of converging predictive rules, e.g., those that preserve exchangeability but are not Kolmogorov consistent.

To conclude, we believe the work by E. Fong, C. Holmes, and S. Walker will spur several new theoretical, modeling, and computational research directions. We commend the authors one more time for an outstanding paper.