# Pietro Rigo's contribution to the Discussion of 'Martingale Posterior Distributions' by Fong, Holmes and Walker

**Pietro Rigo**

Dipartimento di Scienze Statistiche 'P. Fortunati', University of Bologna, Bologna, Italy

*Address for correspondence:* Pietro Rigo, Dipartimento di Scienze Statistiche 'P. Fortunati', University of Bologna, via delle Belle Arti 41, 40126 Bologna, Italy. Email: pietro.rigo@unibo.it

This paper is really interesting and offers a number of intriguing hints. Here, I just make a few isolated remarks without any claim of being exhaustive.

1. I would give more emphasis to the Ionescu-Tulcea theorem (ITT). This theorem is quoted only in passing, but it is the cornerstone of this paper. By ITT, the distribution of $Y_{1:\infty} = (Y_1, Y_2, \ldots)$ is completely determined by the assignment of $\{P_n : n \geq 0\}$, where

$P_0(\cdot) = P(Y_1 \in \cdot)$ and $P_n(\cdot) = P(Y_{n+1} \in \cdot \mid Y_{1:n})$. Exploiting ITT has at least two advantages: (i) The first part of the paper can be made shorter and clearer; (ii) Relying on ITT makes transparent that, in general, to introduce the problem investigated in this paper, there is no need of any distributional assumption on $Y_{1:\infty}$. In particular, $Y_{1:\infty}$ needs not be exchangeable. Exchangeability should be assumed if (and only if) the inferrer feels that it is reasonable for the specific data at hand, which is true in some problems but false in others. As regards this paper, the only advantage of exchangeability is that the distribution of $Y_{1:\infty}$ can be assigned in two ways: by the usual likelihood/prior scheme (thanks to de Finetti's theorem) or via ITT selecting $\{P_n : n \geq 0\}$. These two routes are equivalent and both determine the distribution of $Y_{1:\infty}$. Hence, it is obvious that predictive resampling is identical to posterior sampling for exchangeable data. I realise that the existence of these two routes is expository useful. But, I do not see any other general reason for assuming exchangeability from the outset. See e.g. Berti et al. (2021, 2023).

2. Suppose the distribution of $Y_{1:\infty}$ is assigned via ITT, but, for some reason, $Y_{1:\infty}$ is requested to satisfy some distributional assumption. For instance, $Y_{1:\infty}$ is asked to be exchangeable, or c.i.d., or stationary, and so on. This puts some constraints on the predictive distributions $P_n$. So, the problem arises: Is it possible to characterise a distributional assumption on $Y_{1:\infty}$ in terms of the $P_n$ ? This issue has been addressed in some cases (exchangeability and c.i.d.) but not in others (stationarity, partial exchangeability). See Berti et al. (2021, 2023) and references therein.

3. The information at time $n$ is usually larger than the observed values $y_{1:n}$. This could be modelled by introducing a filtration $\mathcal{G}_n$ such that $\sigma(Y_{1:n}) \subset \mathcal{G}_n$ and defining $P_n$ as $P_n(\cdot) = P(Y_{n+1} \in \cdot \mid \mathcal{G}_n)$. Such a generalisation should have a little cost, as most results on c.i.d. sequences work for an arbitrary filtration $\mathcal{G}_n$.

4. Most probably I miss something, but I have some doubts on Section 2.4.1. It is obviously tempting to assign $P_n$ as the empirical measure. But, it does not work. In fact, if $P_n$ is the empirical measure for every $n \geq 1$, one obtains the trivial sequence $Y_n = Y_1$ a.s. for each $n$.

*Conflict of interest:* None declared.

## References

Berti P., Dreassi E., Leisen F., Pratelli L., & Rigo P. (2023). Bayesian predictive inference without a prior. *Statistica Sinica*, *33*, 1–25. https://doi.org/10.5705/ss.202021.0238

Berti P., Dreassi E., Pratelli L., & Rigo P. (2021). A class of models for Bayesian predictive inference. *Bernoulli*, *27*(1), 702–726. https://doi.org/10.3150/20-BEJ1255