ABSTRACT
The prior distribution on parameters of a likelihood is the usual starting point for Bayesian uncertainty quantification. In this paper, we present a different perspective. Given a finite data sample $Y_{1:n}$ of size $n$ from an infinite population, we focus on the missing $Y_{n+1:\infty}$ as the source of statistical uncertainty, with the parameter of interest being known precisely given $Y_{1:\infty}$. We argue that the foundation of Bayesian inference is to assign a predictive distribution on $Y_{n+1:\infty}$ conditional on $Y_{1:n}$, which then induces a distribution on the parameter of interest. Demonstrating an application of martingales, Doob shows that choosing the Bayesian predictive distribution returns the conventional posterior as the distribution of the parameter. Taking this as our cue, we relax the predictive machine, avoiding the need for the predictive to be derived solely from the usual prior to posterior to predictive density formula. We introduce the martingale posterior distribution, which returns Bayesian uncertainty directly on any statistic of interest without the need for the likelihood and prior, and this distribution can be sampled through a computational scheme we name predictive resampling. To that end, we introduce new predictive methodologies for multivariate density estimation, regression and classification that build upon recent work on bivariate copulas.