# Skewed Bernstein-von Mises theorem and online skew-modal approximations

Francesco Pozza
*Università Bocconi, Italy*

Gaussian deterministic approximations are routinely employed in Bayesian statistics to ease inference when the target posterior of direct interest is intractable. Although these approximations are justified, in asymptotic regimes, by Bernstein–von Mises type results, in practice the expected Gaussian behavior may poorly represent the actual shape of the target posterior, thereby affecting approximation accuracy. Motivated by these considerations, we derive an improved class of closed–form and valid deterministic approximations of posterior distributions which arise from a novel treatment of a third–order version of the Laplace method yielding approximations within a tractable family of skew–symmetric distributions. Under general assumptions which also account for misspecified models and non–i.i.d. settings, this novel family of approximations is shown to have a total variation distance from the target posterior whose rate of convergence improves by at least one order of magnitude the one achieved by the Gaussian from the classical Bernstein–von Mises theorem. Specializing such a general result to the case of regular parametric models shows that the same improvement in approximation accuracy can be also established for polynomially bounded posterior functionals, including moments. Unlike other higher–order approximations based on, e.g., Edgeworth expansions, our results prove that it is possible to derive closed–form and valid densities which are expected to provide, in practice, a more accurate, yet similarly–tractable, alternative to Gaussian approximations of the target posterior of direct interest. We strengthen these arguments by developing a practical skew–modal approximation for both joint and marginal posteriors which achieves the same theoretical guarantees of its theoretical counterpart by replacing the unknown model parameters with the corresponding maximum a posteriori estimate. Finally, we demonstrate that our theoretical findings can be also extended to online settings, that is, when data are collected sequentially and the posterior must be updated at each new data point. Simulation studies and real-data applications confirm that our theoretical results closely match the remarkable empirical performance observed in practice, even in finite, possibly small, sample regimes.