# COMPSTAT 2010 TUTORIAL
# Bayesian discrimination between embedded models

Jean-Michel Marin
Institut de Mathématiques et Modélisation de Montpellier
Université Montpellier 2

We aim at presenting the most standard approaches to the approximation of Bayes factors. The Bayes factor is a fundamental procedure that stands at the core of the Bayesian theory of testing hypotheses, at least in the approach advocated by both Jeffreys (1939) and by Jaynes (2003). Given an hypothesis $H_0 : \theta \in \Theta_0$ on the parameter $\theta \in \Theta$ of a statistical model, with density $f(y|\theta)$, under a compatible prior of the form

$$\pi(\Theta_0)\pi_0(\theta) + \pi(\Theta_0^c)\pi_1(\theta)\,,$$

the *Bayes factor* is defined as the posterior odds to prior odds ratio, namely

$$B_{01}(y) = \frac{\pi(\Theta_0|y)}{\pi(\Theta_0^c|y)} \bigg/ \frac{\pi(\Theta_0)}{\pi(\Theta_0^c)} = \int_{\Theta_0} f(y|\theta)\pi_0(\theta)\mathrm{d}\theta \bigg/ \int_{\Theta_0^c} f(y|\theta)\pi_1(\theta)\mathrm{d}\theta\,.$$

Model choice can be considered from a similar perspective, since, under the Bayesian paradigm (see, e.g., Robert 2001), the comparison of models

$$\mathfrak{M}_i : y \sim f_i(y|\theta_i), \quad \theta_i \sim \pi_i(\theta_i), \quad \theta_i \in \Theta_i, \quad i \in \mathfrak{I}\,,$$

where the family $\mathfrak{I}$ can be finite or infinite, leads to posterior probabilities of the models under comparison such that

$$\mathbb{P}\left(\mathfrak{M} = \mathfrak{M}_i|y\right) \propto p_i \int_{\Theta_i} f_i(y|\theta_i)\pi_i(\theta_i)\mathrm{d}\theta_i\,,$$

where $p_i = \mathbb{P}(\mathfrak{M} = \mathfrak{M}_i)$ is the prior probability of model $\mathfrak{M}_i$.

We consider some of the most common Monte Carlo solutions used to approximate a generic Bayes factor or its fundamental component, the *evidence*

$$m_i = \int_{\Theta_i} \pi_i(\theta_i) f_i(y|\theta_i)\,\mathrm{d}\theta_i\,,$$

aka the marginal likelihood. Longer entries can be found in Carlin and Chib (1995), Chen et al. (2000), Robert and Casella (2004), or Friel and Pettitt (2008). We only briefly mention trans-dimensional methods issued from the revolutionary paper of Green (1995), since our goal is to demonstrate that within-model simulation methods allow for the computation of Bayes factors and thus avoids the additional complexity involved in trans-dimensional methods.

Our focus is on methods that are based on importance sampling strategies, including: crude Monte Carlo, MLE based importance sampling, bridge and harmonic mean sampling (Gelman and Meng 1998), as well as Chib's method based on the exploitation of a functional equality (Chib 1995). We demonstrate how all these methods can be efficiently implemented for testing the significance of a predictive variable in a probit model (Albert and Chib 1993). We compare their performances on a real dataset.

**Keywords:** Bayesian inference; model choice; Bayes factor; Monte Carlo; Importance Sampling; bridge sampling; Chib's functional identity; supervised learning; probit model

**Plan of the tutorial:**

- Introduction on Bayesian model choice
- The Pima Indian benchmark model
- The basic Monte Carlo solution
- Usual importance sampling approximations
- Bridge sampling methodology
- Harmonic mean approximations
- Exploiting functional equalities

## References

Albert, J. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *J. American Statist. Assoc.*, **88** 669–679.

Carlin, B. and Chib, S. (1995). Bayesian model choice through Markov chain Monte Carlo. *J. Royal Statist. Society Series B*, **57** 473–484.

Chen, M., Shao, Q. and Ibrahim, J. (2000). *Monte Carlo Methods in Bayesian Computation*. Springer-Verlag, New York.

Chib, S. (1995). Marginal likelihood from the Gibbs output. *J. American Statist. Assoc.*, **90** 1313–1321.

Friel, N. and Pettitt, A. (2008). Marginal likelihood estimation via power posteriors. *J. Royal Statist. Society Series B*, **70** 589–607.

Gelman, A. and Meng, X. (1998). Simulating normalizing constants: From importance sampling to bridge sampling to path sampling. *Statist. Science*, **13** 163–185.

Green, A. (1995). Reversible jump MCMC computation and Bayesian model determination. *Biometrika*, **82** 711–732.

Jaynes, E. (2003). *Probability Theory*. Cambridge University Press, Cambridge.

Jeffreys, H. (1939). *Theory of Probability*. 1st ed. The Clarendon Press, Oxford.

Robert, C. (2001). *The Bayesian Choice*. 2nd ed. Springer-Verlag, New York.

Robert, C. and Casella, G. (2004). *Monte Carlo Statistical Methods*. 2nd ed. Springer-Verlag, New York.