Bayesian Robustness
IMS Lecture Notes - Monograph Series (1996) Volume 29

COMPARING HIERARCHICAL MODELS USING BAYES FACTOR AND FRACTIONAL BAYES FACTOR: A ROBUST ANALYSIS

BY FULVIO DE SANTIS AND FULVIO SPEZZAFERRI
Carnegie Mellon University and Universita' di Roma "La Sapienza"

In the recent years several alternative Bayes Factors have been introduced in order to handle the problem of the extreme sensitivity of the Bayes Factor (BF) to the priors of the models under comparison in model selection or hypothesis testing problems. In particular, the impossibility of using the Bayes Factor with standard noninformative priors has led to introduce new automatic criteria as the Intrinsic Bayes Factors (IBFs) and the Fractional Bayes Factor (FBF). As pointed out by De Santis-Spezzaferri (1995), the use of IBFs and of the FBF seems to be appealing also in robust Bayesian analyses, when the priors of the parameters of the models vary in large classes of distributions, containing, in the limiting case, improper priors. In this paper we study the behaviour of the BF and of the FBF in a problem of comparing two hierarchical models. We assume the exchangeability of the parameters and introduce a class of distributions at the third stage of the hierarchy of the "biggest" model. In this context, the use of the FBF seems to avoid the problems of lack of robustness of the BF, providing an alternative to the use of the BF itself.

1. Introduction. Suppose that we want to compare two models M_1 and M_2 given some data y. Let $f_i(y|\eta_i)$ and $\pi_i(\eta_i)$ be respectively the likelihood and the prior distribution of model M_i, $i = 1, 2$. A measure of the evidence given by the data y in favour of model M_2 versus M_1 is represented by the *Bayes factor* (BF) that is defined as

$$B_{21}(y) = \frac{m_2(y)}{m_1(y)}$$

where $m_i(y) = \int f_i(y|\eta_i)\pi_i(\eta_i)d\eta_i$, is the marginal distribution of the data y under model M_i, $i = 1, 2$.

The BF is extremely sensitive to prior assumptions since it depends on the absolute values of the priors of the parameters. Specifically, several problems arise when prior information is weak (see for example Aitkin, 1991, O'Hagan, 1995 and De Santis-Spezzaferri, 1995, for a discussion on this topic). In fact the use of reference priors is not possible, since they are typically improper and hence defined only up to arbitrary constants that do not cancel out in the resulting BF. A possible solution to this problem is to split the sample into two parts $y(l)$ and $y(n-l)$ and then to use $y(l)$ as a *training sample* to convert improper priors into proper ones, and the rest of the data to compute the BF, called *Partial Bayes factor* (PBF):

$$B_{21}(l) = \frac{\int f_2(y(n-l)|\eta_2)\pi(\eta_2|y(l))d\eta_2}{\int f_1(y(n-l)|\eta_1)\pi(\eta_1|y(l))d\eta_1}.$$