Bayesian Analysis (2013) **8**, Number 2, pp. 361–380

Integral Priors and Constrained Imaginary Training Samples for Nested and Non-nested Bayesian Model Comparison

Juan Antonio Cano * and Diego Salmerón †‡

Abstract. In Bayesian model selection when the prior information on the parameters of the models is vague default priors should be used. Unfortunately, these priors are usually improper yielding indeterminate Bayes factors that preclude the comparison of the models. To calibrate the initial default priors Cano et al. (2008) proposed integral priors as prior distributions for Bayesian model selection. These priors were defined as the solution of a system of two integral equations that under some general assumptions has a unique solution associated with a recurrent Markov chain. Later, in Cano et al. (2012b) integral priors were successfully applied in some situations where they are known and they are unique, being proper or not, and it was pointed out how to deal with other situations. Here, we present some new situations to illustrate how this new methodology works in the cases where we are not able to explicitly find the integral priors but we know they are proper and unique (one-sided testing for the exponential distribution) and in the cases where recurrence of the associated Markov chains is difficult to check. To deal with this latter scenario we impose a technical constraint on the imaginary training samples space that virtually implies the existence and the uniqueness of integral priors which are proper distributions. The improvement over other existing methodologies comes from the fact that this method is more automatic since we only need to simulate from the involved models and their posteriors to compute very well behaved Bayes factors.

Keywords: Bayesian model selection, Bayes factor, intrinsic priors, integral priors.

1 Introduction

For the sake of objectivity default methods have been proposed to obtain prior distributions for Bayesian estimation and model selection problems. Default methods usually yield an improper prior distribution $\pi^N(\theta) \propto h(\theta)$, then $\pi^N(\theta)$ is determined up to a positive multiplicative constant c, that is $\pi^N(\theta) = c\, h(\theta)$. This is not a serious issue in estimation problems where the posterior is defined by the formal Bayes rule and therefore does not depend on c. However, in model selection problems we have two models M_i, $i = 1, 2$, the data $\mathbf{x}$ are related to the parameter $\theta_i \in \Theta_i$ by a density $f_i(\mathbf{x}|\theta_i)$, the

*Department of Statistics and Operational Research, University of Murcia, Espinardo, Spain, jacano@um.es

†CIBER Epidemiología y Salud Pública (CIBERESP), Spain. diego.salmeron@carm.es

‡Department of Epidemiology, Regional Health Authority, Murcia, Spain.

 DOI:10.1214/13-BA812

default priors are $\pi_i^N(\theta_i) = c_i\ h_i(\theta_i)$, $i = 1, 2$, and the Bayes factor,

$$B_{21}^N(\mathbf{x}) = \frac{m_2^N(\mathbf{x})}{m_1^N(\mathbf{x})} = \frac{c_2 \int_{\Theta_2} f_2(\mathbf{x}|\theta_2)h_2(\theta_2)d\theta_2}{c_1 \int_{\Theta_1} f_1(\mathbf{x}|\theta_1)h_1(\theta_1)d\theta_1},$$

depends on the arbitrary ratio c_2/c_1. Therefore we are left with the indetermination of the ratio c_2/c_1 and several attempts for solving this problem using the so called intrinsic priors are in Berger and Pericchi (1996) and Moreno et al. (1998). However, in Cano et al. (2004) it is shown that when comparing the double exponential versus the normal location models any prior is intrinsic and the same prior has to be used for the location parameter of both models. To overcome this difficulty integral priors have been proposed as prior distributions for Bayesian model selection in Cano et al. (2008), where under some assumptions it was stated that integral priors are unique. Furthermore, integral priors have been successfully applied to the one way random effects model in Cano et al. (2007a, 2007b) and they have been illustrated in different situations in Cano et al. (2012b). In this last paper we have dealt with situations where integral priors are unique being proper or not, in particular we developed the toy example of testing the mean of a normal population with known variance but treated as a testing problem (not an estimation problem) and the comparison of the normal and the double exponential location models, where integral priors are unique and improper.

Integral priors are a new methodology to deal with Bayesian model selection problems and for the sake of self content they are introduced in this section, where it is also stated how they operate. For the purpose of comparison we also present the expected posterior priors introduced in Pérez and Berger (2002) that are defined as

$$\pi_i^*(\theta_i) = \int \pi_i^N(\theta_i|x)m^*(x)dx,$$

where $m^*(x)$ is an arbitrary predictive density for the imaginary training sample x and $\pi_i^N(\theta_i|x) \propto f_i(x|\theta_i)\pi_i^N(\theta_i)$, $i = 1, 2$. The priors $\pi_i^*(\theta_i)$ will not be proper priors unless $m^*(x)$ itself is proper, but the resulting Bayes factor $B_{ij}^*(\mathbf{x})$ is 'well calibrated' in the sense that if $m^*(x)$ is replaced by $cm^*(x)$ for a constant $c > 0$, the Bayes factor obtained is $B_{ij}^*(\mathbf{x})$ again.

Integral priors are defined as the solutions of the system of integral equations

$$\pi_1(\theta_1) = \int \pi_1^N(\theta_1|x)m_2(x)dx \tag{1}$$

and

$$\pi_2(\theta_2) = \int \pi_2^N(\theta_2|x)m_1(x)dx, \tag{2}$$

where x is an imaginary minimal training sample and for $i = 1, 2$, $m_i(x) = \int_{\Theta_i} f_i(x|\theta_i)$ $\pi_i(\theta_i)d\theta_i$. Note that a minimal training sample is a sample of minimal size for which the marginals under both models are finite. We emphasize that in this system both priors $\pi_i(\theta_i)$, $i = 1, 2$, are unknown. The details on the justification and good properties of the integral priors can be found in Cano et al. (2008), where the main result states that in the continuous case if the Markov chain with transition density $Q(\theta_1'|\theta_1) = \int g(\theta_1, \theta_1', \theta_2, x, x')\, dx dx' d\theta_2$, where

$$g(\theta_1, \theta_1', \theta_2, x, x') = \pi_1^N(\theta_1'|x)\, f_2(x|\theta_2)\, \pi_2^N(\theta_2|x')\, f_1(x'|\theta_1),$$

is recurrent then there exists a solution to the integral equations and it is unique up to a multiplicative constant; in fact the σ-finite invariant measure for the recurrent Markov chain is the integral prior $\pi_1(\theta_1)$. The transition density $Q(\theta_1'|\theta_1)$ can be made into a transition density on the parameter space Θ_2 too. Therefore, there exists a parallel Markov chain with the same properties; in particular, if one is (Harris) recurrent then so is the other. Moreover, in the case where Harris recurrence holds but we are unable to explicitly find the unique pair of integral priors the corresponding Bayes factor can be approximated simulating the associated Markov chain and using the ergodic theorem. Then, we can operate in one of the two following ways, the theoretical one, looking for the invariant measure of the Markov chain with transition density $Q(\theta_1'|\theta_1)$, and the empirical one, obtaining a realization of this Markov chain and using it to approximate the corresponding Bayes factor. The transition $\theta_1 \to \theta_1'$ of this Markov chain is made of the following four steps:

$$\begin{array}{lccl} 1. & x' & \sim & f_1(x'|\theta_1) \\ 2. & \theta_2 & \sim & \pi_2^N(\theta_2|x') \\ 3. & x & \sim & f_2(x|\theta_2) \\ 4. & \theta_1' & \sim & \pi_1^N(\theta_1'|x), \end{array} \tag{3}$$

that is we jump from parameters to samples and between models. Note that to operate in the empirical way we just need to simulate from the models and the posteriors which is likely to be easy. In Cano et al. (2007a, 2007b) using the theoretical way we obtained a couple of integral priors and its corresponding Bayes factor for the nested case of the one way random effects model.

Section 2 is devoted to illustrate this methodology with the non-nested case of the one-sided testing for the exponential distribution, where an objective answer is far from being simple since even the use of the encompassing model, see Berger and Pericchi (1996), does not work well because it does not provide true Bayes factors. In Moreno (2005) a good *ad hoc* solution is proposed but it is far from being automatic. In that sense our solution is competitive since it is as good as Moreno's solution and its development is automatic, in fact we obtain in this case the recurrence of the associated Markov chain and for it we only need the usual reference prior for the exponential model. In section 3 we present a way to be followed when recurrence of the associated Markov chain is not present or it is not easy to assess it, in this case we simply impose a technical constraint on the imaginary training samples space that has no loss of generality and ensures recurrence. Next, in section 4 we develop some applications of the theory stated in section 3 by considering testing a normal mean with unknown variance where we obtain similar results to those that are obtained using intrinsic priors (but no more choices than the reference priors are virtually needed) and the one-sided testing for the variance of a normal model with unknown mean, obtaining results that are as good as those that are obtained using expected posterior priors although we do not need to use a predictive density and they are obtained more automatically. We also present an example with heavy tails, testing the location parameter of the Cauchy density with unknown scale, and we finish this section by considering the application of our methodology to the one-way homoscedastic ANOVA problem. Finally, in section 5 we present some relevant conclusions.

2 One-sided testing for the exponential distribution

Now, we suppose that data $\mathbf{x} = (x_1, ..., x_m)$ are independently drawn from the exponential distribution. Consider the problem of testing $H_0 : \theta \in (0, 1)$ versus $H_1 : \theta > 1$, where θ is the mean of the exponential distribution $Exp(\theta)$. As a model selection problem, we are interested in the comparison of the following two models

$$M_1 : f_1(x|\theta_1) = \frac{1}{\theta_1} \exp(-x/\theta_1),\ \pi_1^N(\theta_1) = \frac{c_1}{\theta_1} 1_{(0,1)}(\theta_1)$$

and

$$M_2 : f_2(x|\theta_2) = \frac{1}{\theta_2} \exp(-x/\theta_2),\ \pi_2^N(\theta_2) = \frac{c_2}{\theta_2} 1_{(1,\infty)}(\theta_2).$$

Intrinsic priors do not necessarily exist for this type of non-nested problem and if they do they are not necessarily unique, see Cano et al. (2004). On the other hand, the Bayes factor obtained from the intrinsic priors for the encompassing approach does not correspond to an actual Bayes factor, see Moreno (2005), where an *ad hoc* alternative solution is proposed. Nevertheless, as we show next integral priors are unique, proper and they operate in an automatic way. The minimal training sample is a single replication x and the posterior distributions are

$$\pi_1^N(\theta_1|x) = xe^x \theta_1^{-2} e^{-x/\theta_1} 1_{(0,1)}(\theta_1)$$

and

$$\pi_2^N(\theta_2|x) = \frac{x}{1 - e^{-x}} \theta_2^{-2} e^{-x/\theta_2} 1_{(1,\infty)}(\theta_2).$$

The transition $\theta_1 \to \theta_1'$ of the associated Markov chain is made of the following steps:

1. $x' = -\theta_1 \log u_1$, where $u_1 \sim U(0, 1)$.
2. $\theta_2 = -x'/\log(u_2(1 - e^{-x'}) + e^{-x'})$, where $u_2 \sim U(0, 1)$.
3. $x = -\theta_2 \log u_3$, where $u_3 \sim U(0, 1)$.
4. $\theta_1' = (1 - \frac{1}{x} \log u_4)^{-1}$, where $u_4 \sim U(0, 1)$,

from where it is obtained the following expression for the transition

$$\theta_1' = \left(1 + \frac{\log u_4}{(\theta_1 \log u_1 / \log(u_2(1-u_1^{\theta_1}) + u_1^{\theta_1})) \log u_3}\right)^{-1}.$$

Since $x/(1-e^{-x}) > 1$ we deduce that $\pi_2^N(\theta_2|x) > \theta_2^{-2} e^{-x/\theta_2}$ and therefore

$$Q(\theta_1'|\theta_1) \geq \int \pi_1^N(\theta_1'|x) f_2(x|\theta_2) \left(\int \theta_2^{-2} e^{-x'/\theta_2} f_1(x'|\theta_1) dx'\right) dx d\theta_2$$

$$= \int \frac{\pi_1^N(\theta_1'|x) f_2(x|\theta_2)}{\theta_2(1+\theta_2)} dx d\theta_2 =: q(\theta_1').$$

Then $\beta = \int_0^1 q(\theta_1') d\theta_1' = \log 2 \simeq 0.69$ and therefore $Q(\theta_1'|\theta_1)$ satisfies the Doeblin condition and integral priors are unique and proper priors. Note that the Doeblin condition, see Athreya et al. (1996), states that when for a Markov chain with transition density $Q(\theta_1'|\theta_1)$ there exists a density $f(\theta_1')$, an integer k and an $\varepsilon > 0$ such that

$$Q^k(\theta_1'|\theta_1) \geq \varepsilon f(\theta_1') \quad \forall \theta_1 \text{ and } \forall \theta_1',$$

where $Q^k(\theta_1'|\theta_1)$ denotes the k-step transition density, then there exists a unique invariant probability measure to which the Markov chain converges at a geometric rate from any starting point.

Although in this case we have not been able to explicitly obtain the integral priors we have simulated their associated Markov chains and we present their histograms in figures (1) and (2).

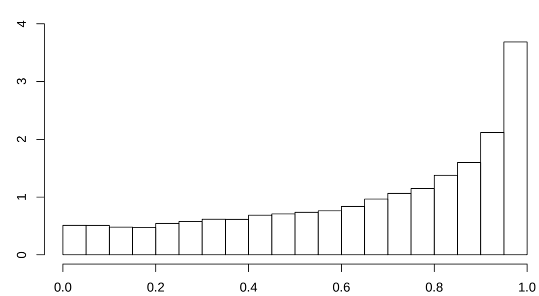

Figure 1: Histogram of the integral prior for model M_1 in the one-sided testing example of the Exponential Distribution.

Table 1 shows the Bayes factor of model M_2 to model M_1 for different values of $(m, \bar{\mathbf{x}})$ using the intrinsic priors proposed in Moreno (2005) and the Bayes factor obtained running the Markov chains associated with the integral priors. Numbers in this table

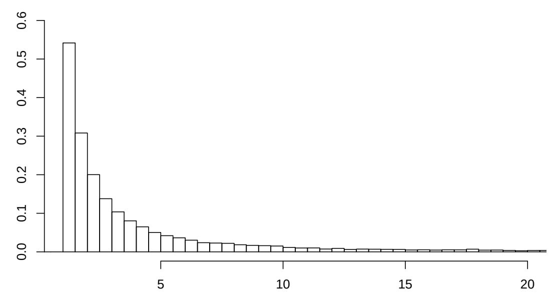

Figure 2: Histogram of the integral prior for model M_2 in the one-sided testing example of the Exponential Distribution.

m	$\bar{\mathbf{x}}$=0.1	$\bar{\mathbf{x}}$=0.6	$\bar{\mathbf{x}}$=1.4	$\bar{\mathbf{x}}$=1.9
5	0.002, 0.000	0.205, 0.241	1.596, 4.370	5.64, 20.22
10	0.000, 0.000	0.089, 0.098	2.320, 7.676	21.74, 98.06
15	0.000, 0.000	0.045, 0.046	3.326, 12.360	81.91, 422.28
20	0.000, 0.000	0.024, 0.023	4.720, 19.08	306.8, 1735.1
50	0.000, 0.000	0.001, 0.001	35.34, 187.09	813537, 6040420

Table 1: Bayes factor B_{21} of model M_2 to model M_1 using the integral priors (left) and the intrinsic priors proposed in Moreno (2005) (right).

show that both types of priors are well behaved. Note that the integral prior for model M_1 is increasing despite the restriction of the initial default prior is decreasing in the constrained parameter space specified in model M_1 while the integral prior for model M_2 is decreasing as it is the restriction of the initial default prior in the constrained parameter space specified in model M_2; on the other hand, intrinsic priors are decreasing for both models.

3 Existence of a unique and proper solution when using constrained imaginary training samples

In Cano et al. (2012b) the normal and the double exponential location models were compared using integral priors and it was found that integral priors were unique although improper and exact Bayes factors were obtained. On the other hand, a lack of stability was detected where obtaining approximated Bayes factors running their associated Markov chains. Although small differences were observed between the exact and the approximated posterior probabilities of the normal model for different sample sizes, as we were concerned with this problem we satisfactorily explored how to control these Markov chains by imposing some restrictions in their evolution. The exploration of this problem leads us to investigate how recurrence could be assessed for complex problems where even we initially did not know the existence or uniqueness of the integral priors. Here we propose using a constraint on the imaginary training samples space to ensure recurrence of the associated Markov chain. The constraint is applied in steps 1 and 3 of (3) and now the transition $\theta_1 \to \theta_1'$ is made of the following four steps

$$\begin{array}{lll} 1. & x & \sim & f_1^A(x|\theta_1) \propto f_1(x|\theta_1)\mathbb{I}_A(x) \\ 2. & \theta_2 & \sim & \pi_2^N(\theta_2|x) \\ 3. & x' & \sim & f_2^A(x'|\theta_2) \propto f_2(x'|\theta_2)\mathbb{I}_A(x') \\ 4. & \theta_1' & \sim & \pi_1^N(\theta_1'|x'), \end{array} \tag{4}$$

where $\mathbb{I}_A(x)$ is the indicator function (1 if $x \in A$, 0 otherwise) and A is a subset of the imaginary training samples space to be chosen later.

The idea behind this is that the constraint on the imaginary training samples prevents the Markov chain from escaping to infinity and therefore guarantees the existence and the uniqueness of an invariant probability measure. The result is stated in the following proposition where x, x', x^* and $\tilde{x}$ denote imaginary training samples.

Proposition 1. *If the set A is chosen such that the function*

$$K_A(x|x^*) = \mathbb{I}_A(x^*) \int f_1^A(x|\tilde{\theta}_1)\pi_1^N(\tilde{\theta}_1|\tilde{x})f_2^A(\tilde{x}|\theta_2')\pi_2^N(\theta_2'|x^*)d\theta_2' d\tilde{x} d\tilde{\theta}_1$$

satisfies the minorizing condition $K_A(x|x^*) \geq g_A(x)$, for some function $g_A(x)$ with $\beta = \int g_A(x)dx > 0$, then there exists a unique invariant probability for the Markov chain defined by (4).

Proof. The density of the transition $\theta_1 \to \theta_1'$ defined by (4) is given by

$$Q_A(\theta_1'|\theta_1) = \int \pi_1^N(\theta_1'|x')f_2^A(x'|\theta_2)\pi_2^N(\theta_2|x)f_1^A(x|\theta_1)dx d\theta_2 dx'$$

and therefore the density for the two-step transition is

$$Q_A^2(\theta_1'|\theta_1) = \int \pi_1^N(\theta_1'|x')f_2^A(x'|\theta_2)\pi_2^N(\theta_2|x)K_A(x|x^*)f_1(x^*|\theta_1)dx^* dx d\theta_2 dx'.$$

It follows that $Q_A^2(\theta_1'|\theta_1) \geq p(\theta_1')\beta$, where $p(\theta_1')$ is the density

$$p(\theta_1') = \beta^{-1} \int \pi_1^N(\theta_1'|x')f_2^A(x'|\theta_2)\pi_2^N(\theta_2|x)g_A(x)dx d\theta_2 dx'$$

and hence the Doeblin condition that was previously stated is satisfied and the existence and uniqueness of an invariant probability for the Markov chain with transition given by (4) is proved.

Definition 1. *If $\pi_1^A(\theta_1)$ is the invariant density of (4), then $\pi_1^A(\theta_1)$ and*

$$\pi_2^A(\theta_2) = \int \pi_2^N(\theta_2|x) f_1^A(x|\theta_1)\pi_1^A(\theta_1) d\theta_1 dx,$$

are defined as the integral prior distributions for the comparison of the models $\{M_1, M_2\}$ with imaginary training samples space A. The priors $\{\pi_1^A(\theta_1), \pi_2^A(\theta_2)\}$ can be seen as the solutions of the following system of integral equations

$$\pi_1^A(\theta_1) = \int \pi_1^N(\theta_1|x) m_2^A(x) dx$$

and

$$\pi_2^A(\theta_2) = \int \pi_2^N(\theta_2|x) m_1^A(x) dx,$$

where $m_i^A(x) = \int f_i^A(x|\theta_i)\pi_i^A(\theta_i) d\theta_i$, $i = 1, 2$.

A sensible choice of the set A should contain the observed data. Also note that if A is a compact set and the model is regular enough to satisfy

$$\inf\{\pi_2^N(\theta_2'|x_1') : x_1' \in A\} > 0 \ \forall \theta_2' \in \Theta_2,$$

the above minimizing condition is easily obtained. Hence, from a practical point of view we recommend that A be compact, simple to simulate the Markov chain as easily as possible, and large enough to ensure that virtually all possible data will be contained in A. For example, a practical choice is the cartesian product of compact intervals of equal length that can be chosen in a subjective way using the data. Concretely, our recommendation is keeping the imaginary training samples within an interval of $\pm 5s$ about the sample mean, where s is the sample standard deviation.

The resulting Bayes factor

$$B_{12}^A(\mathbf{x}) = \frac{\int f_1(\mathbf{x}|\theta_1)\pi_1^A(\theta_1) d\theta_1}{\int f_2(\mathbf{x}|\theta_2)\pi_2^A(\theta_2) d\theta_2},$$

can be approximated using the simulation of the associated Markov chain.

4 Applying the use of constrained imaginary training samples

4.1 Testing a normal mean with unknown variance

Suppose the data $\mathbf{x}$ are i.i.d. $N(\mu, \sigma^2)$ and we consider testing $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$. A Bayesian setting for this problem is that of choosing between the models

$$M_1 : N(\mathbf{x}|\mathbf{0}, \sigma_1^2 \mathbf{I})$$

and

$$M_2 : N(\mathbf{x}|\mu_2\mathbf{1}, \sigma_2^2\mathbf{I}),$$

where $\mathbf{1} = (1, ..., 1)'$ and $\mathbf{I}$ is the identity matrix. The initial conventional priors are $\pi_1^N(\sigma_1) \propto 1/\sigma_1$ and $\pi_2^N(\mu_2, \sigma_2) \propto 1/\sigma_2$ and the imaginary training samples considered are of size two, $x = (x_1, x_2)$. Here a reasonable choice for the compact set is $A = \{(x_1, x_2) \in \mathbb{R}^2 : |x_1| \leq b, |x_2| \leq b\}$ with $b > 0$.

The posterior distributions are

$$\pi_1^N(\sigma_1|x_1, x_2) \propto \sigma_1^{-3} \exp\left(-\frac{x_1^2 + x_2^2}{2\sigma_1^2}\right)$$

and

$$\pi_2^N(\mu_2, \sigma_2|x_1, x_2) \propto N(\mu_2|\overline{x}, \sigma_2^2/2)\sigma_2^{-2} \exp\left(-\frac{\overline{x^2} - \overline{x}^2}{\sigma_2^2}\right)$$

and therefore the transition $\sigma_1 \to \sigma_1'$ of the associated Markov chain is made of the following steps

1. x_i is simulated from the density proportional to $N(x_i|0, \sigma_1^2)\mathbb{I}_{[-b,b]}(x_i)$, $i = 1, 2$, that is, a truncated normal density.

2. $$\sigma_2^2 = \frac{\overline{x^2} - \overline{x}^2}{v} \text{ and } \mu_2 \sim N(\overline{x}, \sigma_2^2/2),$$

   with v simulated from a gamma density with shape $1/2$ and scale 1.

3. x_i' is simulated from the density proportional to $N(x_i'|\mu_2, \sigma_2^2)\mathbb{I}_{[-b,b]}(x_i')$, $i = 1, 2$.

4. $\sigma_1' = \sqrt{\frac{x_1'^2 + x_2'^2}{2w}}$, where $w \sim Exp(1)$.

For a sample size of $n = 10$ we approximate the Bayes factor $B_{12}^A(\overline{\mathbf{x}}, \overline{\mathbf{x}^2})$ for different values of $\overline{\mathbf{x}}$ and $\overline{\mathbf{x}^2}$. The imaginary training samples spaces A we used are the ones defined for $b = 10, 25, 50$ and 100, respectively. The results are in table 2 and they are based on 100000 transitions of the associated Markov chain. Furthermore, we have compared our results with the ones obtained using intrinsic priors. The intrinsic priors for this problem are given by $\pi_1^*(\sigma_1) = \pi_1^N(\sigma_1)$ and

$$\pi_2^*(\mu_2, \sigma_2) = \int \pi_2^*(\mu_2, \sigma_2|\sigma_1)\pi_1^*(\sigma_1)d\sigma_1,$$

where $\pi_2^*(\mu_2, \sigma_2|\sigma_1) = \int \pi_2^N(\mu_2, \sigma_2|x)N(x_1|0, \sigma_1^2)N(x_2|0, \sigma_1^2)dx$. The posterior for model M_2 can be written as

$$\pi_2^N(\mu_2, \sigma_2|x) = 2|x_1 - x_2|N(x_1|\mu_2, \sigma_2^2)N(x_2|\mu_2, \sigma_2^2)/\sigma_2,$$

then, using the change of variables $x_1 = u + v$, $x_2 = u - v$, the density $\pi_2^*(\mu_2, \sigma_2|\sigma_1)$ can be expressed as

$$\pi_2^*(\mu_2, \sigma_2|\sigma_1) = \int \frac{8|v|}{\sigma_2}(4\pi^2\sigma_1^2\sigma_2^2)^{-1}\exp\left(-h(u, v, \mu_2, \sigma_1, \sigma_2)\right) du dv,$$

where $h(u, v, \mu_2, \sigma_1, \sigma_2)$ is given by

$$\frac{1}{2\sigma_2^2}\left[(u + v - \mu_2)^2 + (u - v - \mu_2)^2\right] + \frac{1}{2\sigma_1^2}\left[(u + v)^2 + (u - v)^2\right]$$

and therefore $\pi_2^*(\mu_2, \sigma_2|\sigma_1)$ is given by

$$\int \frac{2|v|}{\sigma_2\pi^2\sigma_1^2\sigma_2^2}\left(\frac{\pi}{1/\sigma_1^2 + 1/\sigma_2^2}\right)^{1/2}\exp\left(-\frac{\mu_2^2}{\sigma_1^2 + \sigma_2^2} - (1/\sigma_1^2 + 1/\sigma_2^2)v^2\right) dv =$$

$$\frac{2\exp\left(-\mu_2^2/(\sigma_1^2 + \sigma_2^2)\right)}{\pi^{3/2}\sigma_1^2\sigma_2^3(1/\sigma_1^2 + 1/\sigma_2^2)^{3/2}} = N(\mu_2|0, (\sigma_1^2 + \sigma_2^2)/2)\frac{2\sigma_1}{\pi(\sigma_1^2 + \sigma_2^2)}.$$

The resulting Bayes factor for the intrinsic priors is $B_{12}^*(\mathbf{x}) = m_1(\mathbf{x})/m_2(\mathbf{x})$ with

$$m_1(\mathbf{x}) = \frac{\Gamma\left(\frac{n}{2}\right)}{2}(\pi n\overline{\mathbf{x}^2})^{-n/2}$$

and

$$m_2(\mathbf{x}) = \int H(\sigma_1, \sigma_2, \mathbf{x})N(\mu_2|\overline{\mathbf{x}}, \sigma_2^2/n)N\left(\mu_2|0, (\sigma_1^2 + \sigma_2^2)/2\right) d\sigma_1 d\sigma_2 d\mu_2, \tag{5}$$

where

$$H(\sigma_1, \sigma_2, \mathbf{x}) = \frac{2^{1-n/2}}{\pi^{1+n/2}}\frac{\left(2\pi\frac{\sigma_2^2}{n}\right)^{1/2}\sigma_2^{-n}\exp\left(\frac{-\nu s^2}{2\sigma_2^2}\right)}{\sigma_1^2 + \sigma_2^2}$$

and $\nu s^2 = \sum(x_i - \overline{\mathbf{x}})^2$. We can integrate out μ_2 in (5) and the resulting integral can be solved numerically. The results for intrinsic priors are also given in table 2 for comparison. From table 2 we can see that posterior probabilities for the simpler model are similar for both integral and intrinsic priors and they are well behaved, that is they decrease as $\overline{\mathbf{x}}$ goes away from zero, irrespectively of the value of b. We note that at the time we faced this problem for the first time running the Markov chains without restrictions we obtained completely unstable posterior probabilities for the simpler model, sometimes ranging from zero to one for the same data set; the idea of using constrained imaginary training samples was first explored in Cano et al. (2012b).

4.2 One-sided testing for the variance of a normal model with unknown mean

Suppose that $\mathbf{x}$ are i.i.d. $N(\mu, \sigma^2)$ and consider the problem of testing $H_0 : \sigma \leq \sigma_0$ versus $H_1 : \sigma > \sigma_0$, where $\sigma_0 > 0$ is a specified value and μ is unknown. This is equivalent to comparing the non-nested models

$\overline{\mathbf{x}^2}$	$\overline{\mathbf{x}}$	b=10	b=25	b=50	b=100	Intrinsic
1	0	0.814	0.809	0.817	0.812	0.789
	0.2	0.786	0.782	0.788	0.785	0.757
	0.4	0.675	0.672	0.677	0.676	0.635
	0.6	0.395	0.398	0.397	0.401	0.351
	0.8	0.058	0.058	0.056	0.058	0.049
	1	0.000	0.000	0.000	0.000	0.000
10	0	0.828	0.820	0.810	0.809	0.789
	0.2	0.826	0.816	0.807	0.806	0.786
	0.4	0.818	0.808	0.798	0.798	0.777
	0.6	0.804	0.793	0.783	0.784	0.761
	0.8	0.783	0.770	0.761	0.762	0.736
	1	0.752	0.737	0.728	0.731	0.701
50	0	0.806	0.826	0.814	0.807	0.789
	0.2	0.806	0.826	0.813	0.806	0.788
	0.4	0.804	0.825	0.811	0.804	0.786
	0.6	0.802	0.822	0.809	0.801	0.783
	0.8	0.798	0.819	0.805	0.797	0.779
	1	0.793	0.814	0.799	0.792	0.774

Table 2: Posterior probabilities of the simpler model, M_1, for different values of $\overline{\mathbf{x}}$, $\overline{\mathbf{x}^2}$ and b and for the intrinsic priors.

$$M_1 : N(\mathbf{x}|\mu_1\mathbf{1}, \sigma_1^2\mathbf{I}),\ \sigma_1 \leq \sigma_0$$

and

$$M_2 : N(\mathbf{x}|\mu_2\mathbf{1}, \sigma_2^2\mathbf{I}),\ \sigma_2 > \sigma_0$$

where the simplest model is difficult to pinpoint. The initial default priors are $\pi_i^N(\mu_i, \sigma_i) \propto \mathbb{I}_{R_i}(\sigma_i)/\sigma_i$, $i = 1, 2$, with $R_1 = (0, \sigma_0]$ and $R_2 = (\sigma_0, +\infty)$. Here we use imaginary training samples of size three, $x = (x_1, x_2, x_3)$, to make computations easier although the minimal size is two.

The posterior distributions are

$$\pi_i^N(\mu_i, \sigma_i|x) \propto \sigma_i^{-4} \exp\left\{-\frac{1}{2\sigma_i^2}\left[\nu s^2 + 3(\mu_i - \overline{x})^2\right]\right\} \mathbb{I}_{R_i}(\sigma_i),\ i = 1, 2,$$

where $\nu s^2 = \sum_{i=1}^3 (x_i - \overline{x})^2$ and therefore $\pi_i^N(\mu_i|\sigma_i, x) = N(\mu_i|\overline{x}, \sigma_i^2/3)$ and

$$\pi_i^N(\sigma_i|x) \propto \sigma_i^{-3} \exp\left\{-\frac{\nu s^2}{2\sigma_i^2}\right\} \mathbb{I}_{R_i}(\sigma_i),\ i = 1, 2.$$

The simulation from $\pi_i^N(\sigma_i|x)$ can be carried out using the probability integral transform and the integrals

$$\int_0^\sigma t^{-3}\exp\left\{-\frac{\nu s^2}{2t^2}\right\}dt = \frac{\exp\left(-\nu s^2/2\sigma^2\right)}{\nu s^2}$$

and

$$\int_{\sigma_0}^\sigma t^{-3}\exp\left\{-\frac{\nu s^2}{2t^2}\right\}dt = \frac{\exp\left(-\nu s^2/2\sigma^2\right)-\exp\left(-\nu s^2/2\sigma_0^2\right)}{\nu s^2},$$

if $\nu s^2 > 0$ and $\sigma > 0$. Then the transition $(\mu_1, \sigma_1) \to (\mu_1', \sigma_1')$ of the associated Markov chain is made of the following steps

1. (x_1, x_2, x_3) is simulated from the density proportional to

$$N(x_1|\mu_1,\sigma_1^2)N(x_2|\mu_1,\sigma_1^2)N(x_3|\mu_1,\sigma_1^2)\mathbb{I}_A(x_1,x_2,x_3)$$

2. $$\sigma_2^2 = -\frac{\nu s^2}{2\log\left(u(1-\alpha)+\alpha\right)},\ \alpha = \exp(-\nu s^2/2\sigma_0^2),\ u \sim U(0,1)$$

   and $\mu_2 \sim N(\overline{x}, \sigma_2^2/3)$, with $\overline{x} = (x_1+x_2+x_3)/3$

3. (x_1', x_2', x_3') is simulated from the density proportional to

$$N(x_1'|\mu_2,\sigma_2^2)N(x_2'|\mu_2,\sigma_2^2)N(x_3'|\mu_2,\sigma_2^2)\mathbb{I}_A(x_1',x_2',x_3')$$

4. $$\sigma_1'^2 = \frac{\nu s'^2\sigma_0^2}{\nu s'^2 - 2\sigma_0^2\log v},\ v \sim U(0,1)$$

   and $\mu_1' \sim N(\overline{x'}, \sigma_1'^2/3)$, with $\overline{x'} = (x_1'+x_2'+x_3')/3$.

For this example we consider $\sigma_0 = 1$, sample sizes $m = 10$ and $m = 20$, sample mean $\overline{\mathbf{x}} = 0$ and standard deviations s from 0.6 to 1.5 with step 0.05. For each value of s we have approximated the posterior probability of model M_2 using the integral methodology (blue/dashed lines in figure 3) and the empirical expected posterior priors (red/solid lines in figure 3). To compute each posterior probability associated with the integral priors we have approximated the Bayes factor B_{21}^A simulating the associated Markov chain (length 100000) with imaginary training samples space $A = [-5s, 5s]^3$. The posterior probabilities associated with the empirical expected posterior priors have been computed as follows. For each value of s we have drawn $z_1, ..., z_m \sim N(0, 1)$ and we have considered the data $x_i = (z_i - \overline{z})s/s_z$, $i = 1, ..., m$, with $(m-1)s_z^2 = \sum_{i=1}^m (z_i - \overline{z})^2$, see Berger and Pericchi (2004), example 7. The mean and the standard deviation of the data $x_1, ..., x_m$ are 0 and s, respectively, and using this sample we have built the empirical expected posterior priors that have been simulated 100000 times to approximate the posterior probability of M_2. Figure 3 shows that both methodologies practically agree

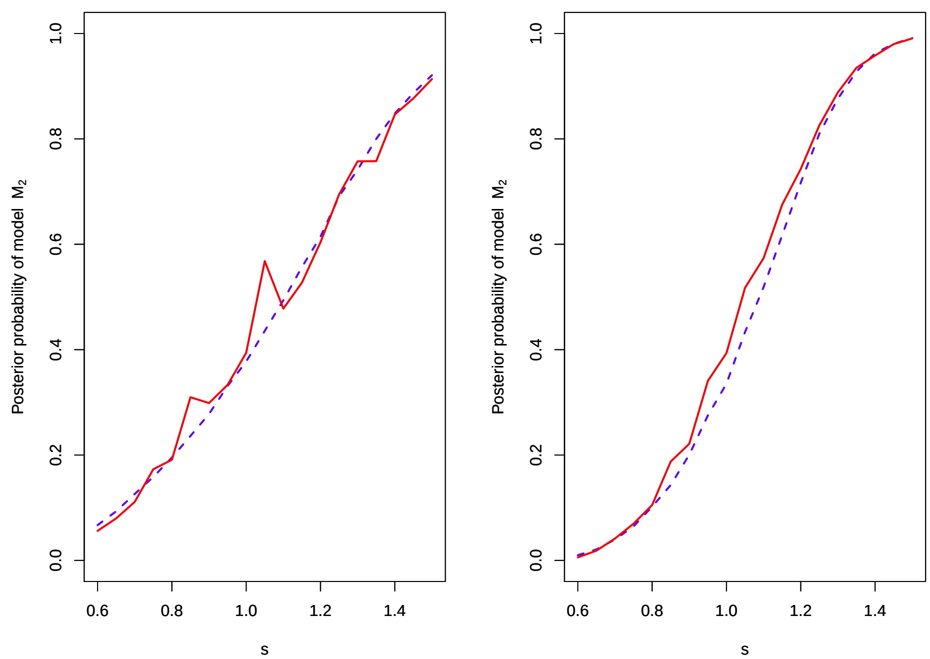

Figure 3: Posterior probability of model M_2 for several values of the sample standard deviation s, $\overline{\mathbf{x}} = 0$ and sample size $m = 10$ (left) and $m = 20$ (right), using the integral methodology (blue/dashed lines) and the methodology of the empirical expected posterior priors (red/solid lines).

but computations derived from integral priors are more stable than those obtained with the empirical expected posterior priors, perhaps due to the fact that samples $(x_1, ..., x_m)$ are varying as s does. Finally, we have checked robustness of the integral priors methodology in this case. Table 3 shows that the posterior probability of model M_2 is not sensitive with respect to the choice of A for several values of m and s.

4.3 Testing the location parameter of a Cauchy distribution

We consider a Cauchy density with location θ and scale σ

$$\mathcal{C}(x|\theta, \sigma) = \frac{1}{\pi\sigma\left(1 + \left(\frac{x-\theta}{\sigma}\right)^2\right)},$$

and the simpler model is $\theta = 0$.

m	s	$A=[-5s,5s]^3$	$A=[-10s,10s]^3$
10	1.1	0.50	0.49
	1.2	0.62	0.63
	1.3	0.74	0.75
20	1.1	0.51	0.52
	1.2	0.72	0.73
	1.3	0.88	0.88
40	1.1	0.59	0.59
	1.2	0.86	0.86
	1.3	0.98	0.97

Table 3: Posterior probability of model M_2 for several values of the sample standard deviation s, $\overline{\mathbf{x}}=0$, two compact sets A, $\sigma_0=1$ and sample size $m=10$, 20 and 40. Each probability is based on a Markov chain of length 500000.

We express the Cauchy density as a mixture of the normal and the gamma:

$$\mathcal{C}(x|\theta,\sigma)=\int_0^{+\infty} N(x|\theta,\sigma^2/\lambda)\mathcal{G}(\lambda|1/2,2)d\lambda.$$

The posterior distribution for the Cauchy parameters $\pi(\theta,\sigma|x)$ given the imaginary minimal training sample $x=(x_1,x_2)$ and the prior $\pi(\theta,\sigma)\propto 1/\sigma$, is the marginal of

$$\pi(\theta,\sigma,\lambda_1,\lambda_2|x)\propto\frac{1}{\sigma}N(x_1|\theta,\sigma^2/\lambda_1)N(x_2|\theta,\sigma^2/\lambda_2)\mathcal{G}(\lambda_1|1/2,2)\mathcal{G}(\lambda_2|1/2,2).$$

Therefore to simulate $(\theta,\sigma)\sim\pi(\theta,\sigma|x)$ we can proceed with the following three steps:

1. $\lambda=(\lambda_1,\lambda_2)\sim\pi(\lambda|x)$
2. $\sigma\sim\pi(\sigma|\lambda,x)$
3. $\theta\sim\pi(\theta|\sigma,\lambda,x)$

First, to simulate $\pi(\lambda|x)$, note that

$$\pi(\lambda|x)\propto\mathcal{G}(\lambda_1|1/2,2)\mathcal{G}(\lambda_2|1/2,2)\int\frac{1}{\sigma}N(x_1|\theta,\sigma^2/\lambda_1)N(x_2|\theta,\sigma^2/\lambda_2)d\theta d\sigma$$

$$=\mathcal{G}(\lambda_1|1/2,2)\mathcal{G}(\lambda_2|1/2,2)\frac{1}{2|x_1-x_2|}$$

and therefore the simulation is straightforward. Then, to simulate $\pi(\sigma|\lambda,x)$ and $\pi(\theta|\sigma,\lambda,x)$ note that

$$\pi(\theta,\sigma|\lambda,x) \propto \frac{1}{\sigma}N(x_1|\theta,\sigma^2/\lambda_1)N(x_2|\theta,\sigma^2/\lambda_2)$$
$$\propto \frac{1}{\sigma^3}\exp\left(-\frac{1}{2\sigma^2}\left(\lambda_1(x_1-\theta)^2+\lambda_2(x_2-\theta)^2\right)\right)$$
$$= \frac{1}{\sigma^3}\exp\left(-\frac{1}{2\sigma^2}\left(H_1(\lambda,x)+(\lambda_1+\lambda_2)(\theta-H_2(\lambda,x))^2\right)\right)$$
$$= \frac{1}{\sigma^3}\exp\left(-\frac{H_1(\lambda,x)}{2\sigma^2}\right)\exp\left(-\frac{(\theta-H_2(\lambda,x))^2}{2\sigma^2/(\lambda_1+\lambda_2)}\right),$$

where
$$H_1(\lambda,x)=\lambda_1\lambda_2(x_1-x_2)^2/(\lambda_1+\lambda_2)$$
and
$$H_2(\lambda,x)=(\lambda_1x_1+\lambda_2x_2)/(\lambda_1+\lambda_2).$$
Therefore $\pi(\theta|\sigma,\lambda,x)$ is the normal density with mean $H_2(\lambda,x)$ and variance $\sigma^2/(\lambda_1+\lambda_2)$. Moreover
$$\pi(\sigma|\lambda,x)=\int\pi(\theta,\sigma|\lambda,x)d\theta\propto\frac{1}{\sigma^2}\exp\left(-\frac{H_1(\lambda,x)}{2\sigma^2}\right)$$
and to simulate $\pi(\sigma|\lambda,x)$, we made $v\sim\mathcal{G}(1/2,2/H_1(\lambda,x))$ and we take $\sigma=1/\sqrt{v}$.

For the simpler model we just have the parameter σ and the posterior distribution is
$$\pi(\sigma|x_1,x_2)\propto\frac{\sigma}{(\sigma^2+x_1^2)(\sigma^2+x_2^2)}$$
which simulation can be carried out using the probability integral transform as follows
$$\int_0^t\pi(\sigma|x)d\sigma=\frac{H(t,x)-H(0,x)}{\lim_{t\to+\infty}H(t,x)-H(0,x)}=1-\frac{H(t,x)}{H(0,x)},$$
where
$$H(\sigma,x)=\frac{\log\left(\frac{\sigma^2+x_2^2}{\sigma^2+x_1^2}\right)}{2(x_1^2-x_2^2)}.$$
Then if $u\sim U(0,1)$, the solution t of the equation
$$H(t,x)=(1-u)H(0,x)$$
is a simulation from $\pi(\sigma|x)$. This equation is equivalent to
$$\log\left(\frac{t^2+x_2^2}{t^2+x_1^2}\right)=(1-u)\log(x_2^2/x_1^2)$$
$$\frac{t^2+x_2^2}{t^2+x_1^2}=\left(\frac{x_2^2}{x_1^2}\right)^{1-u}$$

dataset	min	max	b=30	b=80	b=150
$(\theta = 3)$	-3.3	42.9	0.0192	0.0197	0.0200
$(\theta = 1)$	-3.7	19.0	0.5346	0.5649	0.5625
$(\theta = 0)$	-8.9	15.8	1.8986	1.9103	1.9663

Table 4: Bayes factor B_{12}^{A} in favor of the simpler model $\theta = 0$, using integral priors for different constraints and for 3 simulated datasets.

from which we can obtain t easily.

We have simulated three samples of size 20 from the Cauchy $\mathcal{C}(\theta, 2)$ for $\theta = 3$, $\theta = 1$ and $\theta = 0$, respectively. We have simulated the associated Markov chains for 3 different constraints $|x_i| \leq b$, $i = 1, 2$, with $b = 30$, $b = 80$ and $b = 150$, respectively. The length of these chains was 200000. The values of the Bayes factors that are obtained and the range of the data are shown in table (4).

The p-value associated with a t-test for the simulated data with $\theta = 0$ that is 0.418 and the Bayes factors in the last row of table (4) provide evidence in favor of $\theta = 0$. On the other hand, the p-values that are obtained with the data simulated from $\theta = 3$ and $\theta = 1$ that are 0.05715 and 0.007039, respectively, do not provide support for $\theta = 0$. The same happens with their corresponding Bayes factors, although we observe that the behavior of these Bayes factors is more robust and consistent. Moreover, another dataset was simulated from $\theta = 0$ for which the data range was from -28 to 2.8 and a p-value of 0.029 and Bayes factors of 4.0444, 4.0031 and 4.0401, respectively, were obtained, showing more intensively the lack of robustness and consistency of the t-test to departures from normality.

4.4 Integral priors for homoscedastic one way ANOVA

The one way homoscedastic ANOVA problem was studied using the intrinsic priors in Cano et al. (2012a) and now we apply the integral priors methodology to this problem. Roughly speaking we are comparing the means of k homoscedastic normal populations and then the models are

$$M_1 : \mu_1 = \mu_2 = ... = \mu_k = \mu, \quad vs. \quad M_2 : \text{all the } \mu_i \text{ are not equal.}$$

The default priors are $\pi_1^N(\mu, \tau) \propto 1/\tau$ and $\pi_2^N(\mu_1, ..., \mu_2, \sigma) \propto 1/\sigma$, and the data are $z = (x_1, ..., x_k)$ with $x_i \sim N_{n_i}(\mu_i \mathbf{1}, \sigma^2 \mathbf{I})$, $i = 1, ..., k$, under model M_2 and $z \sim N_n(\mu, \tau^2 \mathbf{I})$ under model M_1. Note that

$$\pi_2^N(\mu_1, ..., \mu_k, \sigma | x_1, ..., x_k) \propto \frac{1}{\sigma} \prod_{i=1}^{k} N_{n_i}(x_i | \mu_i \mathbf{1}, \sigma^2 \mathbf{I})$$

and therefore $\mu_1, ..., \mu_k$ are independent given $(\sigma, x_1, ..., x_k)$ and

$$\pi_2^N(\mu_i | \sigma, x_1, ..., x_k) = N(\mu_i | \overline{x}_i, \sigma^2 / n_i).$$

On the other hand

$$\pi_2^N(\sigma|x_1,...,x_k) = \int \pi_2^N(\mu_1,...,\mu_k,\sigma|x_1,...,x_k)d\mu_1...d\mu_k$$

$$\propto \frac{1}{\sigma}\int \prod_{i=1}^{k} N_{n_i}(x_i|\mu_i\mathbf{1},\sigma^2\mathbf{I})d\mu_1...d\mu_k = \frac{1}{\sigma}\prod_{i=1}^{k}\int N_{n_i}(x_i|\mu_i\mathbf{1},\sigma^2\mathbf{I})d\mu_i$$

$$\propto \frac{1}{\sigma}\prod_{i=1}^{k}\sigma^{1-n_i}\exp\left(-\frac{(n_i-1)s_i^2}{2\sigma^2}\right) = \sigma^{k-n-1}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{k}(n_i-1)s_i^2\right)$$

where $n = n_1 + ... + n_k$ and s_i^2 is the sample variance of the sample x_i. The minimal training sample consists of two observations from one of the populations and a single observation from the remaining $k-1$ populations, see Cano et al. (2012a). Therefore, if two observations are considered from the population j with $j \in \{1,...,k\}$ then $n_i = 1\,\forall\, i \neq j$, $n_j = 2$ and

$$\pi_2^N(\sigma|x_1,...,x_k) \propto \sigma^{-2}\exp\left(-\frac{s_j^2}{2\sigma^2}\right).$$

For the simpler model the posterior $\pi_1^N(\mu,\tau|x_1,...,x_k)$ can be easily simulated using that

$$\pi_1^N(\mu|x_1,...,x_k,\tau) = N(\mu|\overline{x},\tau^2/n)$$

and

$$\pi_1^N(\tau|x_1,...,x_k) \propto \tau^{-n}\exp\left(-\frac{(n-1)s^2}{2\tau^2}\right),$$

where s^2 is the sample variance of z.

For $k = 3$ populations we simulated the Markov chain with the constraint [-10,10] on the training samples space. For the simulation of the Markov chain at steps 1 and 3 we simulated the imaginary training sample with $j \sim U(\{1,2,3\})$. Using this Markov chain we computed the Bayes factor B_{12}^A for $\overline{\mathbf{x}}_1 = \overline{\mathbf{x}}_2 = 0$ and

$$\overline{\mathbf{x}}_3 \in \{-2,-1.5,-1,-0.5,-0.25,0,0.25,0.5,1,1.5,2\}.$$

The sample size was 10 for each population and the sample variances were 0.9, 1.3 and 1.2, respectively. Figure (4) shows that the Bayes factor is very well behaved, diminishing as the sample mean of the third population goes away from zero.

5 Conclusions

We have explained how integral priors operate in Bayesian model selection and we have illustrated their use with some complex problems including one-sided testing situations that are non-nested problems for which automatic solutions had not been previously

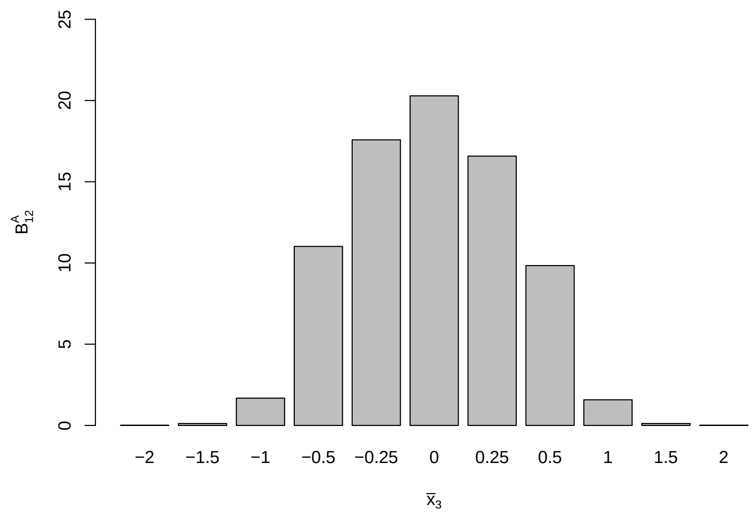

Figure 4: Bayes factors B_{12}^{A} according to the sample mean $\bar{\mathbf{x}}_3$, when $\bar{\mathbf{x}}_i = 0$, $i = 1, 2$.

found. Integral priors are an automatic tool to compute Bayes factors since we only have to simulate from the involved models and their posterior distributions once a default prior has been assigned to each model. Our methodology has been proved to be competitive to solve complex problems like the one-sided testing for the exponential distribution and the one way homoscedastic ANOVA problem.

Several situations may arise when applying this methodology. If we are able to obtain the unique invariant distribution we can straightforwardly compute the unique integral Bayes factor, this is the case of the problem of testing a normal mean with known variance that was considered in Cano et al. (2012b). If we can just establish the positive recurrence of the associated Markov chain we can approximate the unique integral Bayes factor simulating this Markov chain, this is the case of the one-sided testing for the exponential distribution. On the other hand, we have satisfactorily dealt with other problems needing constrained imaginary training samples to assess the recurrence of the associated Markov chain.

References

Athreya, K. B., Doss, H., and Sethuraman, J. (1996). "On the convergence of the Markov chain simulation method." *The Annals of Statistics*, 24: 69–100. 365

Berger, J. and Pericchi, L. R. (1996). "The intrinsic Bayes factor for model selection and prediction." *Journal of the American Statistical Association*, 91: 109–122. 362, 363

— (2004). "Training samples in objective Bayesian model selection." *The Annals of Statistics*, 32: 841–869. 372

Cano, J. A., Carazo, C., and Salmerón, D. (2012a). "Bayesian model selection approach to the one way analysis of variance under homoscedasticity." *Computational Statistics*, doi:10.1007/s00180-012-0339-8. 376, 377

Cano, J. A., Iniesta, M., and Salmerón, D. (2012b). "Integral priors for Bayesian model selection: illustrative implementations and perspectives." Technical report, Department of Statistics and Operational Research, University of Murcia. 361, 362, 366, 370, 378

Cano, J. A., Kessler, M., and Moreno, E. (2004). "On intrinsic priors for nonnested models." *Test*, 13: 445–463. 362, 364

Cano, J. A., Kessler, M., and Salmerón, D. (2007a). "Integral priors for the one way random effects model." *Bayesian Analysis*, 2: 59–68. 362, 363

— (2007b). "A synopsis of integral priors for the one way random effects model." In *Bayesian Statistics 8*, 577–582. Oxford University Press. 362, 363

Cano, J. A., Salmerón, D., and Robert, C. P. (2008). "Integral equation solutions as prior distributions for Bayesian model selection." *Test*, 17: 493–504. 361, 362, 363

Moreno, E. (2005). "Objective Bayesian methods for one-sided testing." *Test*, 14: 181–198. 363, 364, 365, 366

Moreno, E., Bertolino, F., and Racugno, W. (1998). "An intrinsic limiting procedure for model selection and hypotheses testing." *Journal of the American Statistical Association*, 93: 1451–1460. 362

Pérez, J. M. and Berger, J. (2002). "Expected posterior priors for model selection." *Biometrika*, 89: 491–511. 362

Acknowledgments

This research was supported by the Séneca Foundation Programme for the Generation of Excellence Scientific Knowledge under Project 15220/PI/10.