OBJECTIVE BAYESIAN COMPARISON OF CONSTRAINED ANALYSIS OF VARIANCE MODELS

Guido Consonni¹ and Roberta Paroli²

¹ Dipartimento di Scienze Statistiche
Universit Cattolica del Sacro Cuore
Largo Gemelli 1, 20133, Milano, Italy
(e-mail: guido.consonni@unicatt.it)

² Dipartimento di Scienze Statistiche
Universit Cattolica del Sacro Cuore
Largo Gemelli 1, 20133, Milano, Italy
(e-mail: roberta.paroli@unicatt.it)

ABSTRACT. We consider the comparison of models specified by parametric equality or inequality constraints, involving a combination of non-nested models, as well as nested models having the same dimension. We adopt an objective Bayesian approach, and describe a general procedure to obtain the posterior probability of each model under consideration. Our method combines in a unified way both the intrinsic prior methodology -which is especially appropriate for comparing nested models having different dimensions- and the encompassing prior approach -which was devised to compare models specified by inequality constraints. We apply our method to constrained ANOVA models, and evaluate its performance, relative to currently available techniques, through simulation studies.

1 INTRODUCTION

We consider the comparison of models specified by inequality or equality constraints on its parameters, or possibly by a combination of them. These models are common in the social sciences; see Klugkist et al. (2005) and Wesel et al. (2011). For instance, consider a four-way (normal) ANOVA with group means $\mu_j$. One possible model is $M_1: \mu_1 < \mu_2 < \mu_3 < \mu_4$, while another one is $M_2: \{\mu_1 = \mu_3\} < \{\mu_2 = \mu_4\}$. Two special models stand out: the full model $M_f: \mu_1; \mu_2; \mu_3; \mu_4$, wherein no constraint is imposed on the parameters, and the null model $M_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$. From a Bayesian perspective, models are compared in terms of their posterior probability (conditional on the available data $y$). The posterior probability of an arbitrary model $M_c$, say, can be deduced from the marginal density of $y$ conditional on $M_c$ (also called the marginal likelihood of model $M_c$), and the prior probability of $M_c$, $\Pr(M_c)$. Another possibility, which we follow here, is to single out a benchmark model (e.g. $M_0$), and compute the Bayes factor (BF) of $M_c$ against $M_0$, $BF_{c0}(y)$. The collection of $BF_{c0}(y)$, and their prior probabilities $\Pr(M_c)$, as $c$ varies, can be combined to return the posterior probability over the entire model space. Recently Klugkist et al. (2005), Klugkist et al. (2007), Laudy et al. (2007) have introduced a methodology, named encompassing prior approach, for dealing specifically with inequality constrained models. With some ad hoc modification, their methodology can be applied also to models specified through equality constraints (basically approximating an equality constraint with an “about equality constraint”). Recently Wesel et al. (2011) have presented an objective Bayesian method to compare constrained ANOVA
models using an encompassing prior approach. Their basic idea is to start with an improper default prior on the parameter of the encompassing model; next to turn this prior into a proper one using training data, so as to apply the encompassing prior machinery. Specifically they adopt an expected posterior prior (EPP) approach, as described in Pérez and Berger (2002), using as base measure the empirical distribution of a minimal training sample. They name their method empirical EPP (EEPP). In this paper we sketch a procedure for the objective Bayesian comparison of constrained models, which deals simultaneously with inequality and equality constraints, without resorting to approximations for the latter. Our method uses intrinsic priors, see for instance Girón et al. (2006), as well as the encompassing prior approach, and for this reason we name it Intrinsic Encompassing Prior (IEP). A full description of the methodology, as well as a thorough comparison with existing alternative methods, will be shortly reported in a paper by the same Authors.

2 BAYES FACTORS AND INTRINSIC ENCOMPASSING PRIORS FOR CONSTRAINED ANOVA MODELS

Consider a normal ANOVA model with J groups and group means $\mu_j$, $j = 1, \ldots, J$ and let $y_{ij}$ be the observation of subject $i$ in group $j$, so that

$$y_{ij} = \mu_j + \epsilon_i,$$

with $\epsilon_i \sim N(0, \sigma^2)$. Under the full model, the mean structure is unconstrained so that $(\mu_1, \ldots, \mu_J) \in \mathbb{R}^J$. A constrained model is specified by equality and inequality constraints among the $\{\mu_j\}$ through the symbols “<”, “>”, “=”. Consider a constrained ANOVA model $M_c$. Let $M_{e(c)}$ be the corresponding encompassing-$M_c$ model (this model has the same equality constraints as $M_c$; on the other hand it imposes no structure on the parameters involved in the inequality constraints under $M_c$, which are therefore free to move). If $\Theta_c$ denotes the parameter space of $M_c$, and similarly for $\Theta_{e(c)}$ with regard to $M_{e(c)}$, we clearly get $\Theta_c \subset \Theta_{e(c)}$. Denote with $\alpha_0$ the group-mean under $M_0$ (the same for all $J$ groups), and let $\sigma_0^2$ be the error variance. The goal is to compute the BF of $M_c$ against the null model $M_0$ based on the conditional intrinsic prior procedure. This will be done in three steps (superscript “IP” stands for intrinsic prior).

i) Intrinsic step: Compute

$$BF^{IP}_{M_{e(c)}, M_0}(y | \alpha_0, \sigma_0).$$

ii) Encompassing step: Compute

$$BF^{IP}_{M_c, M_{e(c)}}(y | \alpha_0, \sigma_0) = \frac{Pr^{IP}[\Theta | y, \alpha_0, \sigma_0, M_{e(c)}]}{Pr^{IP}[\Theta | \alpha_0, \sigma_0, M_{e(c)}]}.$$

iii) Finally obtain

$$BF^{IP}_{M_c, M_0}(y | \alpha_0, \sigma_0) = BF^{IP}_{M_c, M_{e(c)}}(y | \alpha_0, \sigma_0) \times BF^{IP}_{M_{e(c)}, M_0}(y | \alpha_0, \sigma_0).$$
As far as step i) is concerned, we need to compute the marginal likelihood for model $M_{c(e)}$ under the conditional intrinsic prior. This can be expressed as the expectation of a particular function of a Beta-distributed random variable. Although no closed form expression for this expectation is available, we provide an estimate of the marginal likelihood using an MCMC algorithm based on the technique proposed by Chib and Jeliazkov (2001). The evaluation of $BF_{M_{c(e)}}^{IP}(y|\alpha_0, \sigma_0)$ in step ii) naturally lends itself to a Monte Carlo approximation. Specifically, if parameter values are sampled from the conditional intrinsic prior, and posterior, under $M_{c(e)}$, the empirical frequencies of $\Theta_e$ obtained under either distributions can be used to estimate both the numerator and denominator.

Eventually, in order to implement our procedure, one has to assign a value to $\alpha_0$ and $\sigma_0$, the overall expectation and standard deviation under $M_0$. The most obvious way is to estimate both parameters. Typically a maximum likelihood estimate $\hat{\alpha}_0, \hat{\sigma}_0$ is sufficient, and this will be close to the Bayesian estimate based on a noninformative prior.

3 Simulation study

To evaluate the performance of the IEP approach, we provide in this section results based on a preliminary simulation study. In particular we compare the IEP results with those of the EEPP of Wesel et al. (2011) using one of their simulation studies. Specifically, we consider ANOVA models with equal group sizes ($n_j = 25$ or $n_j = 50$, for $j = 1, \ldots, 5$) and homogeneous group variances. The datasets are simulated under three different models, with parameter specifications described in Table 1.

- $M_0$: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$
- $M_1$: $\mu_1 < \mu_2 < \mu_3 < \mu_4 < \mu_5$
- $M_2$: $\mu_2 < \mu_1 < \mu_4 < \{\mu_3 = \mu_5\}$

Notice that in model $M_1$ the group means are in increasing order (inequality constraint). On the other hand, under model $M_2$ the group means are in increasing order but for the equality constraint $\{\mu_3 = \mu_5\}$. The group standard deviations are assumed equal to 1, 3 and 1.55, respectively. For each of the three models described in Table 1, and for each group size $n_j = 25$ and $n_j = 50$ for $j = 1, \ldots, 5$, 500 datasets are randomly generated.

For each simulated dataset, we computed the Bayes factor $BF_{M_{c(e)}}^{IP}(y|\hat{\alpha}_0, \hat{\sigma}_0)$. The results are summarized in Table 2 which reports the percentage of times (relative to the 500 datasets) in which each of the four models ($M_0, M_1, M_2, M_f$) was assigned the largest Bayes factor. Values corresponding to the correct model are printed in bold, while the corresponding percentages computed using the EEPP approach are reported in brackets (see Tables 5, 6, 8 and 9 of Wesel et al., 2011). Values are reported for their optimal minimal training sample size equal to 2. Finally, for each of the 500 datasets, the posterior model probabilities for the four competing models are calculated under the assumptions that all models have equal prior probabilities. In the last column of Table 2 the medians of the posterior model probabilities ($PMP_{med}$) for the correct model are provided.
Table 1. Simulation study: models with corresponding group means $\mu_j$ and standard deviations $\sigma_j$

<table>
<thead>
<tr>
<th>Model</th>
<th>$\mu_j$</th>
<th>$\sigma_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_0$</td>
<td>0, 0, 0, 0</td>
<td>1, 1, 1, 1</td>
</tr>
<tr>
<td>$M_1$</td>
<td>0, 1.4, 2.8, 4.2, 5.6</td>
<td>3, 3, 3, 3, 3</td>
</tr>
<tr>
<td>$M_2$</td>
<td>2.23, 1.33, 3.23, 2.33, 3.23</td>
<td>1.55, 1.55, 1.55, 1.55, 1.55</td>
</tr>
</tbody>
</table>

Table 2. Percentage of times (relative to the 500 datasets) in which each of the four models is assigned the largest Bayes factor, and posterior model probability medians for the correct model. Results of Wesel et al. (2011) in brackets.

<table>
<thead>
<tr>
<th>Model</th>
<th>$n_j$</th>
<th>$M_0$</th>
<th>$M_1$</th>
<th>$M_2$</th>
<th>$M_f$</th>
<th>PMPmed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_0$</td>
<td>25</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>(89)</td>
<td>(2)</td>
<td>(4)</td>
<td>(5)</td>
<td>(0.80)</td>
</tr>
<tr>
<td>$M_1$</td>
<td>25</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0</td>
<td>(0)</td>
<td>(100)</td>
<td>(0)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>$M_2$</td>
<td>25</td>
<td>4</td>
<td>0</td>
<td>96</td>
<td>0</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0</td>
<td>(0)</td>
<td>(0)</td>
<td>(96)</td>
<td>(4)</td>
</tr>
</tbody>
</table>

From the results of Table 2 it can be seen that both our IEP and the EEPP methods perform rather well in this simulation study both in terms of the percentage of times the true model scores the highest Bayes factor, and in terms of the median posterior probability for the true model.

Our method appears to perform better than the approach based on EEPP. Clearly more extensive investigations are needed to assess the relative merits.

4 DISCUSSION

From a methodological perspective, some broad conclusions can be gathered: i) IEP deals directly with inequality and equality constraints, and does not require approximations for the latter; ii) from a computational perspective IEP only requires a one-dimensional integration to compute each Bayes factor followed by straightforward Monte Carlo draws to estimate the Bayes factor; it is therefore far less demanding than EEPP which instead requires more elaborate simulations to approximate the mixture distribution defining the expected posterior prior; iii) IEP does not require tuning parameters, as opposed to EEPP.

REFERENCES


