ACCEPTED

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Linear contrasts for the one way analysis of variance: A Bayesian approach

J.A. Cano [a,*], C. Carazo [b], D. Salmerón [c,d]

[a] Departamento de Estadística e Investigación Operativa, Universidad de Murcia, Spain
[b] Escuela de Arquitectura e Ingeniería de la Edificación, Universidad Católica San Antonio de Murcia, Spain
[c] CIBER Epidemiología y Salud Pública (CIBERESP), Spain
[d] Servicio de Epidemiología, Consejería de Sanidad, IMIB-Arrixaca, Murcia, Spain

HIGHLIGHTS

- Linear contrasts between means are dealt with for the first time as objective Bayesian model selection problems.
- A specific solution for the homoscedastic case is proposed.
- The p-value and the posterior probability of the null hypothesis are compared through calibration curves.

ARTICLE INFO

Article history:
Received 27 July 2015
Received in revised form 1 November 2015
Accepted 3 November 2015

Keywords:
Linear contrasts between means
Objective priors
Bayes factors
Intrinsic priors
Calibration curves

ABSTRACT

Linear contrasts between means for the one way analysis of variance are studied for the first time as objective model selection problems. For it, Bayes factors for intrinsic priors are used and classical and Bayesian measures of evidence are compared.

1. Introduction

Let us consider k normal populations $N(x_1|\mu_1, \sigma_1^2), \ldots, N(x_k|\mu_k, \sigma_k^2)$ and independent samples, $\mathbf{x}_i = (x_{i1}, \ldots, x_{in_i})$, from each population $i = 1, \ldots, k$. When the hypothesis of equality of means is rejected an analysis of certain linear contrasts between the means may be of interest. In the frequentist methodology there are several exact tests dealing with this topic, the methods of Scheffé and Tukey are the most commonly used in the homoscedastic case; however, just asymptotic solutions like the Welch's test or the Hotelling's test are obtained when heteroscedasticity is present. The first objective in this paper is to go one step further than in Cano et al. (2013), where the homoscedastic case for the one way ANOVA was dealt with using the intrinsic priors methodology. Here we solve as a model selection problem, linear contrasts like the following

$$H_0 : \varphi = 0 \quad \text{versus} \quad H_1 : \varphi \neq 0, \tag{1}$$

* Corresponding author.
E-mail address: jacano@um.es (J.A. Cano).

where $\varphi = \sum_{i=1}^{k} a_i \mu_i$, with $\sum_{i=1}^{k} a_i = 0$ and at least one $a_i \neq 0$. This interesting classical problem is treated here for the first time as an objective Bayesian model selection one, note that in Box and Tiao (1973) it was dealt with from a Bayesian estimation point of view. In Cano et al. (2013) we argue why this type of problems are better treated from a Bayesian model selection perspective. For it, because Bayes factors for improper noninformative priors are undefined we propose Bayes factors based on the intrinsic methodology. See Berger and Pericchi (1996), Moreno et al. (1998), and Bertolino et al. (2000), where contrasts are briefly discussed as estimation problems. The case of large k is not dealt with as it is behind the scope of this paper.

The second objective is to compare the Bayesian measure of evidence, the posterior probability of the null hypothesis, y, with the frequentist one, the p-value, p. Calibration is a simple means of establishing that comparison, see Girón et al. (2006). We state in this paper that in linear contrasts the posterior probability of the null hypothesis depends on the sample through sufficient statistics and the sample size, and the same is true for the p-value. That is,

$$y = P(H_0|\bar{\mathbf{x}}, \mathbf{s}^2, \mathbf{n}), \tag{2}$$

$$p = P_{H_0}(T \geq t(\bar{\mathbf{x}}, \mathbf{s}^2, \mathbf{n})), \tag{3}$$

where T is the contrast and t is its observed value, $\bar{\mathbf{x}} = (\bar{x}_1, \dots, \bar{x}_k)$, $\mathbf{s}^2 = (s_1^2, \dots, s_k^2)$ and $s_i^2 = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$; therefore we can define different calibration curves varying in (2) and (3) one of the sample means in an interval. Note that the posterior probability of the null hypothesis has been computed using the prior $p_1 = p_2 = 1/2$ for the hypotheses H_0 and H_1 and the p-values are the corresponding to the Scheffé test in the homoscedastic case and to the Welch test in the heteroscedastic one.

The paper is organized as follows. In Section 2 linear contrasts between means for homoscedastic populations are considered using the intrinsic priors methodology. In Cano et al. (2013) it is argued that it was necessary to study the case when homoscedasticity is present, since in this case a specific method can be used, similarly to what happens in the frequentist analysis.

In Section 3 Bayes factors for intrinsic priors are obtained for linear contrasts in the heteroscedastic case. The key idea to develop Sections 2 and 3 was to find a reparameterization allowing to formulate linear contrasts as nested Bayesian model selection problems for which the intrinsic methodology behaves satisfactorily, see Girón et al. (2006) and references therein. This provides us exact solutions even for the heteroscedastic case.

In Section 4 we illustrate the behavior of the calibration curves as the sample size of the involved populations increases. Finally, in Section 5 we briefly summarize the obtained results and we give some concluding remarks.

2. Linear contrasts between means for homoscedastic populations

In this section we consider k normal populations with unknown common variance σ^2 and we want to solve linear contrasts as (1) where, without loss of generality, we assume that $a_1 \neq 0$. The null hypotheses introduce a constraint on the parameters and considering the reparameterization

$$\varphi_1 = \sum_{i=1}^{k} a_i \mu_i, \varphi_2 = \mu_2, \dots, \varphi_k = \mu_k, \tag{4}$$

the linear contrast (1) can be expressed as a nested Bayesian model selection problem where the simple model M_1,

$$f_1(\mathbf{z}|\theta_1) = N_{n_1}\left(\mathbf{x}_1 \middle| \left(\sum_{i=2}^{k} d_i \beta_i\right) \mathbf{1}_{n_1}, \tau^2 I_{n_1}\right) \prod_{i=2}^{k} N_{n_i}(\mathbf{x}_i|\beta_i 1_{n_i}, \tau^2 I_{n_i}),$$

with the prior

$$\pi_1^N(\theta_1) = \frac{c_1}{\tau}, \tag{5}$$

is compared with the complex model M_2

$$f_2(\mathbf{z}|\theta_2) = \prod_{i=1}^{k} N_{n_i}\left(\mathbf{x}_i|\varphi_i 1_{n_i}, \sigma^2 I_{n_i}\right),$$

with the prior

$$\pi_2^N(\theta_2) = \frac{c_2}{\sigma}, \tag{6}$$

where $\mathbf{z} = (\mathbf{x}_1, \dots, \mathbf{x}_k)$, $\theta_1 = (\beta_2, \dots, \beta_k, \tau)$, $\theta_2 = (\varphi_1, \dots, \varphi_k, \sigma)$ and $d_i = -a_i/a_1$ for $i = 2, \dots, k$. Note that just $(k-1)$ means, $\beta_2, \dots, \beta_k$, have been left in model M_1 because of the constraint on the parameters.

To assign default priors we have always assumed that location and scale parameters are *a priori* independent, see Jeffreys (1961).

2.1. Objective Bayes factors for linear contrasts between means for homoscedastic populations

For the prior given in (6) the imaginary minimal training sample, $\mathbf{x}(l)$, is a $(k+1)$-dimensional random vector that consists in two observations from one of the populations and a single observation from each one of the remaining $(k-1)$ populations, that is $\mathbf{x}(l) = (x_1, x_2, \ldots, x_{j1}, x_{j2}, \ldots, x_k)$, with $1 \leq j \leq k$. Therefore it is not unique and this implies that intrinsic priors are not unique. To solve this drawback we use a pooling approach as follows.

2.1.1. Pooled intrinsic prior

The intrinsic priors for the models under comparison are $\{\pi_1^N(\theta_1), \pi_2^I(\theta_2)\}$, where $\pi_2^I(\theta_2) = \int \pi_2^I(\theta_2|\theta_1)\pi_1^N(\theta_1)d\theta_1$ and $\pi_2^I(\theta_2|\theta_1)$ is obtained as follows:

(i) if $\mathbf{x}(l) = (x_{11}, x_{12}, x_2, \ldots, x_j, \ldots, x_k)$, then

$$\pi_2^I(\theta_2|\theta_1) = N\left(\varphi_1 \middle| \sum_{i=2}^k d_i\beta_i, \frac{\tau^2+\sigma^2}{2}\right) HC^+(\sigma|0,\tau) \prod_{i=2}^k N\left(\varphi_i|\beta_i, \tau^2+\sigma^2\right), \tag{7}$$

(ii) if $\mathbf{x}(l) = (x_1, x_2, \ldots, x_{j1}, x_{j2}, \ldots, x_k)$, with $2 \leq j \leq k$, then

$$\pi_2^I(\theta_2|\theta_1) = N\left(\varphi_1 \middle| \sum_{i=2}^k d_i\beta_i, \tau^2+\sigma^2\right) N\left(\varphi_j|\beta_j, \frac{\tau^2+\sigma^2}{2}\right) HC^+(\sigma|0,\tau) \prod_{\substack{i=2\\(i\neq j)}}^k N\left(\varphi_i|\beta_i, \tau^2+\sigma^2\right), \tag{8}$$

where $HC^+(\sigma|0, \tau)$ denotes the half Cauchy density. The conditional intrinsic priors given in (7) and (8) are obtained in a similar way to that in Theorem 2.1 of Cano et al. (2013), where these intrinsic conditional priors were obtained by computing the analytical integrals involved with the global hypothesis of equality of means. Note that although the imaginary minimal training sample is integrated out to obtain the conditional intrinsic priors given in (7) and (8), the φ_i corresponding to the population from which two observations are taken has a variance half of the variance of the other $(k-1)$ remaining populations. This implies a labeling problem and to avoid it we propose a pooled conditional distribution of θ_2 given θ_1, that is

$$\pi_2^I(\theta_2|\theta_1) = N\left(\varphi_1 \middle| \sum_{i=2}^k d_i\beta_i, (1-1/(2k))\left(\tau^2+\sigma^2\right)\right) HC^+(\sigma|0,\tau) \prod_{i=2}^k N\left(\varphi_i|\beta_i, (1-1/(2k))\left(\tau^2+\sigma^2\right)\right), \tag{9}$$

where the common variance $(1-1/(2k))\left(\tau^2+\sigma^2\right)$ is the average of the φ_i's variances. Note that in linear models it is a common practice to average over all possible training samples of minimal size and this is the idea motivating our pooled conditional prior.

2.1.2. Bayes factors for intrinsic priors

For the data $\mathbf{z}$, let $m_i^I(\mathbf{z})$, $i = 1, 2$, be the marginal densities for the intrinsic priors $\{\pi_1^N(\theta_1), \pi_2^I(\theta_2)\}$. The corresponding Bayes factor turns out to be:

(i) when $m_2^I(\mathbf{z})$ comes from the conditional distribution (7),

$$B_{21}^I(\mathbf{z}) = \frac{2^{\frac{2}{3}} I_2}{\pi \int \left(s^2 + n_1\left(\bar{x}_1 - \sum_{i=2}^k d_i\beta_i\right)^2 + \sum_{i=2}^k n_i(\bar{x}_i-\beta_i)^2\right)^{-N/2} d\beta_2 \ldots d\beta_k}, \tag{10}$$

where $I_2 = \int_{\mathbb{R}^{k-1}} I_1 d\beta_2 \ldots d\beta_k$, with

$$I_1 = \int_0^{\pi/2} \frac{(h(n_1,\theta))^{-1/2} \prod_{i=2}^k (g(n_i,\theta))^{-1/2}}{(\sin\theta)^{N-k}\left(\frac{s^2}{\sin^2\theta} + \frac{2n_1\left(\bar{x}_1 - \sum_{i=2}^k d_i\beta_i\right)^2}{h(n_1,\theta)} + \sum_{i=2}^k \frac{n_i(\bar{x}_i-\beta_i)^2}{g(n_i,\theta)}\right)^{\frac{N}{2}}} d\theta,$$

$g(n,\theta) = n + \sin^2\theta$ and $h(n,\theta) = n + 2\sin^2\theta$,

(ii) when $m_2^I(\mathbf{z})$ comes from the conditional distribution (8),

$$B_{21}^I(\mathbf{z}) = \frac{2^{\frac{2}{3}} I_2}{\pi \int \left(s^2 + n_1\left(\bar{x}_1 - \sum_{i=2}^{k} d_i\beta_i\right)^2 + \sum_{i=2}^{k} n_i(\bar{x}_i - \beta_i)^2\right)^{-N/2} d\beta_2 \dots d\beta_k}, \tag{11}$$

where $I_2 = \int_{\mathbb{R}^{k-1}} I_1 d\beta_2 \dots d\beta_k$, with

$$I_1 = \int_0^{\pi/2} \frac{1}{(\sin\theta)^{N-k} (h(n_j,\theta))^{1/2} \prod_{\substack{i=1 \\ (i\neq j)}}^{k} (g(n_i,\theta))^{1/2}}$$

$$\times \left(\frac{s^2}{\sin^2\theta} + \frac{n_1\left(\bar{x}_1 - \sum_{i=2}^{k} d_i\beta_i\right)^2}{g(n_1,\theta)} + \frac{2n_j(\bar{x}_j - \beta_j)^2}{h(n_j,\theta)} + \sum_{\substack{i=2 \\ (i\neq j)}}^{k} \frac{n_i(\bar{x}_i-\beta_i)^2}{g(n_i,\theta)}\right)^{-\frac{N}{2}} d\theta,$$

$g(n,\theta) = n + \sin^2\theta$ and $h(n,\theta) = n + 2\sin^2\theta$,

(iii) when $m_2^I(\mathbf{z})$ comes from the pooled conditional distribution (9),

$$B_{21}^I(\mathbf{z}) = \frac{2 I_2}{\pi \int \left(s^2 + n_1\left(\bar{x}_1 - \sum_{i=2}^{k} d_i\beta_i\right)^2 + \sum_{i=2}^{k} n_i(\bar{x}_i - \beta_i)^2\right)^{-N/2} d\beta_2 \dots d\beta_k}, \tag{12}$$

where $I_2 = \int_{\mathbb{R}^{k-1}} I_1 d\beta_2 \dots d\beta_k$, with

$$I_1 = \int_0^{\pi/2} \frac{\prod_{i=1}^{k} (g(\theta, n_i, k))^{-1/2}}{(\sin\theta)^{N-k}\left(\frac{s^2}{\sin^2\theta} + \frac{n_1\left(\bar{x}_1 - \sum_{i=2}^{k} d_i\beta_i\right)^2}{g(n_1,\theta,k)} + \sum_{i=2}^{k}\frac{n_i(\bar{x}_i-\beta_i)^2}{g(n_i,\theta,k)}\right)^{\frac{N}{2}}} d\theta$$

and $g(n,\theta,k) = (1 - 1/(2k))\, n + \sin^2\theta$.

Note that expressions (10)–(12) are obtained by direct integration on the φ_i's and changing (τ,σ) to polar coordinates. However, the integrals I_2 need to be solved by numerical integration using standard software. This approximation is unfeasible when the sample sizes of the populations considered are large because numerical computations with standard software fail, then an estimate of $m_2^I(\mathbf{z})$ based on simulation techniques is required. In Cano et al. (2013) is stated how to proceed in this case. Finally, we note that Bayes factors (10)–(12) depend on the sample through sufficient statistics.

2.1.3. Simulated examples in homoscedastic populations

From expressions (10)–(12) it is clear that Bayes factors for linear contrasts under homoscedasticity depend on the minimal training sample and therefore we recommend the pooled prior (9) as a sensible approach. Despite this, in the next example we show that the impact of the training sample seems small.

Example 1. We simulated the three normal populations $N(2.5, 1)$, $N(1, 1)$ and $N(1, 1)$, with sample sizes $(n_1, n_2, n_3) = (30, 20, 60)$, and the following values were obtained: $(\bar{x}_1, \bar{x}_2, \bar{x}_3) = (2.204, 1.091, 1.086)$, $(s_1^2, s_2^2, s_3^2) = (35.665, 20.660, 74.211)$. Pairwise comparisons of the means were done and whether the mean of the first population was equal to the average of the means of the other two populations was studied too.

Table 1 shows that the values of the posterior probabilities of M_1 for different selections of the population chosen to take two observations are similar. Furthermore, these values are similar to the value of the posterior probability of M_1 for the pooled conditional distribution given in (9). Finally, we note that the p-values and the posterior probabilities of M_1 convey the same reasonable message in each case; both measures favor the simpler model in the third linear contrast and they favor the complex model in the remaining linear contrasts.

Example 2. Three normal populations of variance 1 were considered, with sample sizes $(n_1, n_2, n_3) = (30, 20, 60)$, sample means $\bar{x}_1 \in (0, 3.5)$, $(\bar{x}_2, \bar{x}_3) = (1.091, 1.086)$ and $(s_1^2, s_2^2, s_3^2) = (35.665, 20.660, 74.211)$. For each one of the infinite values of $\bar{x}_1$ it is desired to test if:

Table 1
P-values, values of the posterior probability of M_1 when the population chosen to take two observations is the first, the second or the third one ($j = 1, 2, 3$), and values of the posterior probability of M_1 corresponding to the pooled conditional distribution given in (9).

Linear contrasts	*P*-values	$P_{j=1}(M_1)$	$P_{j=2}(M_1)$	$P_{j=3}(M_1)$	$P(M_1)$
$\mu_1 - \mu_2$	1.70×10^{-5}	0.024	0.023	0.026	0.024
$\mu_1 - \mu_3$	2.92×10^{-8}	0.0008	0.0009	0.0008	0.0008
$\mu_2 - \mu_3$	0.99	0.882	0.897	0.882	0.888
$-2\mu_1 + \mu_2 + \mu_3$	3.12×10^{-8}	0.0008	0.0008	0.0008	0.0008

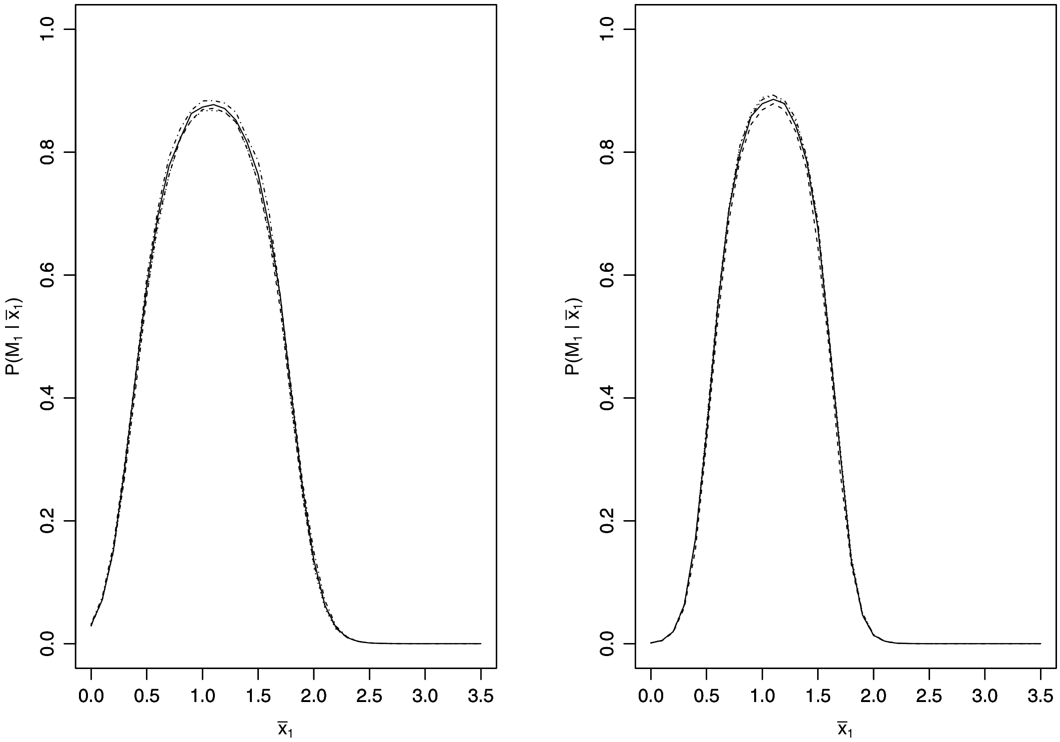

Fig. 1. To the left and to the right four curves are plotted (one for each prior under consideration) that represent the value of the posterior probability of M_1 for linear contrasts (a) and (b) in Example 2, respectively, as $\bar{x}_1$ varies.

(a) $\mu_1 - \mu_2 = 0$,
(b) $-2\mu_1 + \mu_2 + \mu_3 = 0$.

The results are illustrated in Fig. 1. To the left and to the right the graphics correspond to the linear contrasts (a) and (b), respectively. In the graphics of Fig. 1 the solid curve, the dashed curve and the dotted curve represent the value of the posterior probability of M_1 as $\bar{x}_1$ varies and the population chosen to take two observations is the first, the second and the third one, respectively, while the long dashed curve represents the value of this probability when the pooled prior distribution given in (9) is considered.

It is clear that the four curves to the left in Fig. 1 practically agree and the same happens with the curves to the right. The posterior probability of M_1 convey the same reasonable message in each case, when $\bar{x}_1$ is close to 1 the value of this probability is close to 1 and decreases progressively when $\bar{x}_1$ is going away from 1.

3. Linear contrasts between means for heteroscedastic populations

Now we consider k normal populations $N(x_1|\mu_1, \sigma_1^2), \ldots, N(x_k|\mu_k, \sigma_k^2)$ and the reparameterization (4). This yields to the nested selection problem of the simple model M_1

$$f_1(\mathbf{z}|\theta_1) = N_{n_1}\left(\mathbf{x}_1 \left| \left(\sum_{i=2}^{k} d_i\beta_i\right) \mathbf{1}_{n_1}, \tau_1^2 I_{n_1}\right.\right) \prod_{i=2}^{k} N_{n_i}(\mathbf{x}_i|\beta_i \mathbf{1}_{n_i}, \tau_i^2 I_{n_i}),$$

with

$$\pi_1^N(\theta_1) = c_1 / \prod_{i=1}^{k} \tau_i, \tag{13}$$

versus the complex model M_2

$$f_2(\mathbf{z}|\theta_2) = \prod_{i=1}^{k} N_{n_i}(\mathbf{x}_i|\varphi_i \mathbf{1}_{n_i}, \sigma_i^2 I_{n_i}),$$

with

$$\pi_2^N(\theta_2) = c_2 / \prod_{i=1}^{k} \sigma_i, \tag{14}$$

where $\mathbf{z} = (\mathbf{x}_1, \ldots, \mathbf{x}_k)$, $\theta_1 = (\beta_2, \ldots, \beta_k, \tau_1, \ldots, \tau_k)$, $\theta_2 = (\varphi_1, \ldots, \varphi_k, \sigma_1, \ldots, \sigma_k)$ and $d_i = -a_i/a_1$ for $i = 2, \ldots, k$. Again, $(k-1)$ means, $\beta_2, \ldots, \beta_k$, have been left in model M_1 as a consequence of the constraint on the parameters. Note that, to obtain (13) and (14) assumptions are similar to those needed to obtain (5) and (6).

3.1. Objective Bayes factors for linear contrasts between means for heteroscedastic populations

Now the minimal training sample is a $2k$ dimensional random vector that consists in two observations from each population, that is $\mathbf{x}(l) = (x_{ij})$ with $i = 1, \ldots, k$ and $j = 1, 2$. The intrinsic priors are now $\{\pi_1^N(\theta_1), \pi_2^I(\theta_2)\}$, with $\pi_2^I(\theta_2) = \int \pi_2^I(\theta_2|\theta_1)\, \pi_1^N(\theta_1) d\theta_1$ where $\pi_2^I(\theta_2|\theta_1)$ is given by

$$N\left(\varphi_1 \middle| \sum_{i=2}^{k} d_i \beta_i, \frac{\tau_1^2 + \sigma_1^2}{2}\right) \prod_{i=2}^{k} N\left(\varphi_i | \beta_i, \frac{\tau_i^2 + \sigma_i^2}{2}\right) \prod_{i=1}^{k} HC^+(\sigma_i|0, \tau_i),$$

that is again obtained in a similar way to the conditional intrinsic priors (7) and (8). The corresponding Bayes factor $B_{21}^I(\mathbf{z})$ is

$$\frac{2^{\frac{3k}{2}} \pi^{-k} I_2}{\int \left(s_1^2 + n_1\left(\bar{x}_1 - \sum_{i=2}^{k} d_i \beta_i\right)^2\right)^{-n_1/2} \prod_{i=2}^{k} (s_i^2 + n_i(\bar{x}_i - \beta_i)^2)^{-n_i/2} d\beta_2 \ldots d\beta_k}, \tag{15}$$

where $I_2 = \int_{\mathbb{R}^{k-1}} I_1\, d\beta_2 \ldots d\beta_k$, with

$$I_1 = \int_0^{\pi/2} \frac{d\theta}{(\sin\theta)^{n_1-1}(g(\theta, n_1))^{1/2}\left(\frac{s_1^2}{\sin^2\theta} + \frac{2n_1}{g(\theta,n_1)}\left(\bar{x}_1 - \sum_{i=2}^{k} d_i\beta_i\right)^2\right)^{\frac{n_1}{2}}}$$
$$\times \prod_{i=2}^{k} \int_0^{\pi/2} \frac{d\theta}{(\sin\theta)^{n_i-1}(g(\theta, n_i))^{1/2}\left(\frac{s_i^2}{\sin^2\theta} + \frac{2n_i}{g(\theta,n_i)}(\bar{x}_i - \beta_i)^2\right)^{\frac{n_i}{2}}}$$

and $g(\theta, n) = n + 2\sin^2\theta$.

Expression (15) is obtained by direct integration on the φ_i's and changing (τ_i, σ_i) to polar coordinates and I_2 can be carried out by numerical integration except when the sample sizes of the populations are large; then an estimate of $m_2^I(\mathbf{z})$ using simulation techniques is needed. Concretely, these integrals have been carried out using importance sampling with a normal distribution on the location parameters and an inverse gamma on the variances. Note that Bayes factor (15) depends on the sample through sufficient statistics.

3.2. Simulated examples in heteroscedastic populations

In this subsection it is illustrated with two examples that the Bayes factor given in (15) is well behaved.

Example 3. Three cases, each one involving sampling from the three normal populations, $N(1, 1)$, $N(1, 2.25)$ and $N(2.5, 2)$ were considered. The corresponding sufficient statistics are shown in Table 2. It is desired to test in each case if:

(a) $\mu_1 - \mu_2 = 0$,
(b) $\mu_1 - \mu_3 = 0$.

Table 2
Sufficient statistics for three different cases involving the three normal populations, $N(1, 1)$, $N(1, 2.25)$ and $N(2.5, 2)$.

	n	$\bar{\mathbf{x}}$	$\mathbf{s}^2$
Case 1	{35, 30, 35}	{0.921, 0.989, 2.108}	{27.082, 73.200, 44.366}
Case 2	{60, 55, 60}	{0.920, 1.049, 2.804}	{72.556, 144.9, 137.961}
Case 3	{85, 80, 85}	{0.995, 0.965, 2.589}	{71.164, 173.070, 167.738}

Table 3
Posterior probabilities of M_1 and p-values corresponding to the linear contrasts considered in Example 3. The second and third columns correspond to the linear contrast $\mu_1 - \mu_2 = 0$ and the fourth and fifth columns correspond to $\mu_1 - \mu_3 = 0$.

	$P(M_1)$	p-values	$P(M_1)$	p-values
Case 1	0.87	0.82	0.0005	8.78×10^{-6}
Case 2	0.89	0.62	1.23×10^{-9}	6.74×10^{-12}
Case 3	0.92	0.86	1.72×10^{12}	6.66×10^{-15}

Table 4
Sufficient statistics for three cases involving three normal populations, $N(3, 1)$, $N(1, 2.25)$ and $N(2, 2)$.

	n	$\bar{\mathbf{x}}$	$\mathbf{s}^2$
Case 1	{35, 30, 35}	{2.733, 0.994, 1.999}	{41.630, 50.785, 42.617}
Case 2	{60, 55, 60}	{2.788, 1.063, 1.846}	{85.360, 177.735, 113.571}
Case 3	{85, 80, 85}	{2.991, 1.013, 1.993}	{102.887, 194.969, 150.857}

Table 5
Posterior probabilities of M_1 for the linear contrast $\mu_1 + \mu_2 - 2\mu_3 = 0$ for the situations (a) and (b) in Example 4.

	$P(M_1)$, situation (a)	$P(M_1)$, situation (b)
Case 1	0.001893	0.88
Case 2	1.87×10^{-9}	0.91
Case 3	7.30×10^{-13}	0.93

Table 6
Data from three homoscedastic normal populations with variance 1.

	n_i	$\bar{x}_i$	s_i^2
Population 1	20	(−4, 1)	13.058
Population 2	30	−0.146	25.163
Population 3	15	−0.483	15.366

The results are shown in Table 3. Numbers in Table 3 show that the posterior probability of M_1 and the p-value convey the same reasonable message for each case. In the linear contrast $\mu_1 - \mu_2 = 0$ both measures favor the simpler model (the means of the first and the second population are equal), whereas in the linear contrast $\mu_1 - \mu_3 = 0$ both measures favor the complex model (the means of the first and the third population are not equal). Furthermore, the value of the posterior probability of M_1 varies with the sample size providing more evidence in favor of the true model as the sample size increases.

Example 4. Three cases, each one involving sampling from three normal populations were considered and it was tested in each case if $\mu_1 + \mu_2 - 2\mu_3 = 0$ in the following situations:

(a) the populations and data are the same as in Example 3,
(b) the populations are the $N(3, 1)$, $N(1, 2.25)$ and $N(2, 2)$, and the corresponding sufficient statistics are shown in Table 4.

The results are shown in Table 5. Numbers in Table 5 show a reasonable message, the simpler model is rejected in situation (a) and it is accepted in situation (b). Furthermore, again, the value of the posterior probability of M_1 is very well behaved.

4. Calibration of *p*-values

The classic p-values and the Bayes factors for linear contrasts depend on the sample through sufficient statistics. This allows us to calibrate the p-values and the posterior probability of the simple model through the so called calibration curves, see Cano et al. (2013). In this section calibration curves for the data in Tables 6 and 7 are presented.

Table 7
Data from three heteroscedastic normal populations with variances 1, 1.5 and 2, respectively.

	n_i	$\bar{x}_i$	s_i^2
Population 1	20	(−2, 2)	28.162
Population 2	30	1.037	50.602
Population 3	15	2.229	17.729

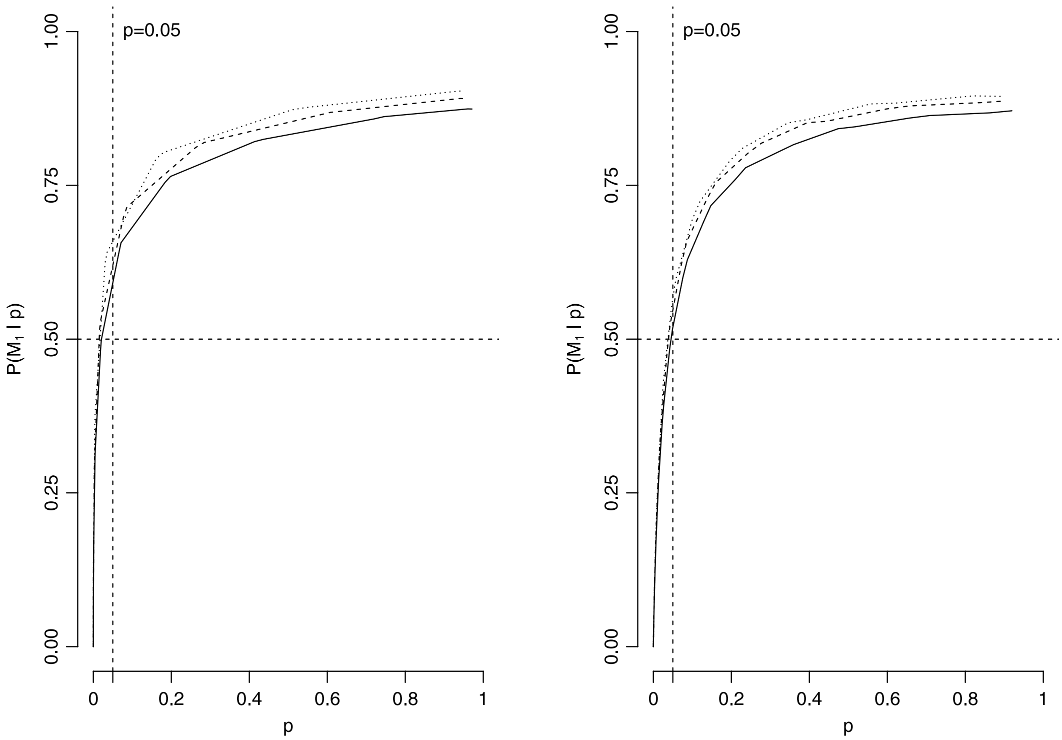

Fig. 2. Calibration curves corresponding to the linear contrast $\mu_1 - \mu_2 = 0$ for the data in Table 6 (to the left) and Table 7 (to the right) and the changes stated above.

Fig. 2 illustrates the calibration curves corresponding to solve the linear contrast $\mu_1 - \mu_2 = 0$ for the data in Tables 6 and 7 and when the sample size of the first population is multiplied by 2 and 4, respectively. In both graphics the solid curve is the calibration curve corresponding to the data shown in Table 6 (to the left) and Table 7 (to the right), respectively. The dashed and dotted curves correspond, respectively, to the modified sample sizes of the first population stated above. Note that s_1^{*2}'s are proportionally modified as $s_1^{*2} = s_1^2 n_1^*/n_1$ too.

Calibration curves presented in Fig. 2 are monotonic increasing functions of the p-value with a *disagreement region* that is non empty. Assuming that the usual critical values to reject the simple model are $p = 0.05$ and $P(M_1) = 0.5$ the disagreement region is the arc of the calibration curve intersecting the quadrant $Q_U = \{(x, y) : x < 0.05, y > 0.5\}$. As a consequence of this both criteria only will convey the same message if the critical p-value is adequately reduced. The order of this diminution should be given by the x-value of the intersecting point of the calibration curve with the line $y = 1/2$. On the other hand, the behavior of the calibration curves depend on the sample size. Generally, when n increases the disagreement region becomes larger. Therefore, when n increases the critical p-value should be diminished accordingly too; otherwise, we could reject hypotheses with high posterior probabilities.

5. Summary and conclusions

We have developed an original objective Bayesian model selection procedure for linear contrasts between means for the one way analysis of variance. Our development is in the context of both homoscedasticity and heteroscedasticity, where a frequentist exact test does not exist. Bayes factors for intrinsic priors have been used because they behave very well in the nested case. In the homoscedastic case, these Bayes factors depend on the minimal training sample chosen and we propose a pooling approach to solve this drawback. In the heteroscedastic case we have illustrated that the posterior probability of M_1 conveys a reasonable message and provides more evidence in favor of the true model as the sample size increases.

The classical p-value and the posterior probability of M_1 for some linear contrasts have been compared through the so-called calibration curves. The behavior of the calibration curves as a function of the sample size has been studied too. The case of a null hypothesis with two or more contrasts is ongoing research that deserves special attention in its own right. All the computations have been programmed in Mathematica and are available under request to the authors

Acknowledgment

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

References

Berger, J.O., Pericchi, L.R., 1996. The intrinsic bayes factor for model selection and prediction. J. Amer. Statist. Assoc. 91, 109–122.
Bertolino, F., Racugno, W., Moreno, E., 2000. Bayesian model selection approach to analysis of variance under heteroscedasticity. Statistician 49, 503–517.
Box, G.E.P., Tiao, G.C., 1973. Bayesian Inference in Statistical Analysis. Addison Wesley.
Cano, J.A., Carazo, C., Salmerón, D., 2013. Bayesian model selection approach to the one way analysis of variance under homoscedasticity. Comput. Statist. 28, 919–931.
Girón, F.J., Martínez, M.L., Moreno, E., Torres, F., 2006. Objetive testing procedures in linear models: calibration of the p-values. Scand. J. Statist. 33, 765–784.
Jeffreys, H., 1961. Theory of Probability. Oxford University Press.
Moreno, E., Bertolino, F., Racugno, W., 1998. An intrinsic limiting procedure for model selection and hypotheses testing. J. Amer. Statist. Assoc. 93, 1451–1460.