Brazilian Journal of Probability and Statistics (2007), 21, pp. 121–139.

Reference analysis for the p-dimensional linear calibration problem

Elias Chaibub Neto and **Márcia D'Elia Branco**
Universidade de São Paulo

Abstract: In this paper we develop reference priors for the linear calibration problem in the context of multivariate response data. We proved that the joint reference posteriors are proper and the marginal posteriors for the interest parameter have finite moments of order strictly smaller than p, the dimension of the response vector $\mathbf{y}$. A Gibbs Sampler scheme is presented for sampling from the reference posterior. We call attention to the fact that our priors reduce to the priors obtained by Ghosh et al. (1995) in the special case of $p = 1$, but they do not agree with Kubokawa and Robert (1994) priors for a general p.

Key words: non informative prior, reference priors, linear calibration, ratio of two normal means and Gibbs Sampler.

1 Introduction

The linear calibration problem, also known as inverse regression problem, is motivated by the comparison of two or more measurement techniques related to the same characteristic of interest, where one of them is much more accurate and expensive than the others. The main objective is to use observations from the less accurate techniques to make inference about the more accurate and expensive one. Key references to this problem are Osborne (1991) and Brown (1993). More specifically, the linear controlled calibration problem can be described as follows: in a first step a set of values $x_1, x_2, \ldots, x_n$ are fixed for the accurate measurement X and other p related measurements are observed, $\mathbf{y}_i = (y_{i1}, \ldots, y_{ip})$, for each fixed value x_i, $i = 1, \ldots, n$. At a second stage, k replications of the response variable $\mathbf{y}_0$ (associated with an unknown covariate value x_0) are observed and the interest centers in estimating the accurate measurement x_0 given $(x_i, \mathbf{y}_i)$, $i = 1, \ldots, n$ and $\mathbf{y}_{0j}$, $j = 1, \ldots, k$. Under an additional linear supposition, the model is given by

$$\mathbf{y}_i = \boldsymbol{\alpha} + \boldsymbol{\beta} x_i + \boldsymbol{\epsilon}_i, \qquad \boldsymbol{\epsilon}_i \sim N_p(\mathbf{0}\,, \sigma^2 \mathbf{I}_p), \qquad i = 1, \ldots, n \tag{1.1}$$

$$\mathbf{y}_{0j} = \boldsymbol{\alpha} + \boldsymbol{\beta} x_0 + \boldsymbol{\epsilon}_{0j}, \qquad \boldsymbol{\epsilon}_{0j} \sim N_p(\mathbf{0}\,, \sigma^2 \mathbf{I}_p), \qquad j = 1, \ldots, k \tag{1.2}$$

where (1.1) and (1.2) represent the calibration and prediction experiments, respectively. The error vectors $\boldsymbol{\epsilon}_i$ and $\boldsymbol{\epsilon}_{0j}$ are independent, σ^2 is unknown, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)'$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$ are the regression parameters and $\mathbf{I}_p$ is the usual identity matrix, $p \times p$. Note that each covariate is associated with p response variables so that we are working in a multivariate context, although the interest parameter x_0 is a scalar.

Under the Bayesian point of view inference is based on posterior distribution of the parameters. In order to obtain these posterior distributions it is necessary to specify the prior distributions for the parameters, but sometimes no prior information is available and it is necessary to consider non subjective priors (or non informative). For a good review in that subject see Kass and Wasserman (1996), where different proposals are presented. We consider here the proposal given by Bernardo (1979), and developed by Berger and Bernardo (1989) and Berger and Bernardo (1992b) known as *reference priors*. Ghosh et al. (1995) obtained the reference prior for the univariante calibration problem ($p = 1$).

In this paper we obtain reference priors for the p-dimensional linear calibration problem, given by (1.1) and (1.2), and discuss the existence of the posterior distribution and posterior moments. Therefore, we extend Ghosh et al. (1995) results. Also, we discuss why Kubokawa and Robert (1994) prior is not the reference prior for this problem.

Reference priors are based on information-theoretical ideas and may be described as model-based positive functions that produce non subjective posteriors dominated by the data. Using these priors we try to minimize the prior influence under the posterior distribution. In Section 2, we develop reference priors through two different approaches: the first, proposed by Berger and Bernardo (1989), recommends the separation of the parameter vector in two groups, one composed by the interest parameter $\{x_0\}$ and the other by the nuisance parameters $\{\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2\}$. The second approach represents a refinement of the first one and recommends splitting the parameter vector into more than two groups ordered according to their inferential importance. A drawback of this approach is that sometimes it is not easy to decide which ordering to use, specially in relation to the nuisance parameters. In these cases, Berger and Bernardo (1992b) recommend to obtain the reference priors associated with all nuisance parameters orderings and then choose the best prior through a comparative study of their inferential performances. In relation to the linear calibration problem, we separated the parameter vector in four groups $\{x_0\}$ $\{\boldsymbol{\alpha}\}$ $\{\boldsymbol{\beta}\}$ $\{\sigma^2\}$, and considered the six possible different orderings of the nuisance parameters: $\{x_0\}\{\boldsymbol{\alpha}\}\{\boldsymbol{\beta}\}\{\sigma^2\}$, $\{x_0\}\{\boldsymbol{\alpha}\}\{\sigma^2\}\{\boldsymbol{\beta}\}$, $\{x_0\}\{\boldsymbol{\beta}\}\{\boldsymbol{\alpha}\}\{\sigma^2\}$, $\{x_0\}\{\boldsymbol{\beta}\}\{\sigma^2\}\{\boldsymbol{\alpha}\}$, $\{x_0\}\{\sigma^2\}\{\boldsymbol{\alpha}\}\{\boldsymbol{\beta}\}$ and $\{x_0\}\{\sigma^2\}\{\boldsymbol{\beta}\}\{\boldsymbol{\alpha}\}$. We show that for these six different orderings the same reference prior is obtained. Then, the four groups reference priors do not depend on the ordering. See Berger and Bernardo (1992b) for details on this approach. We also obtained Jeffreys prior, which can be seen as a reference prior obtained by grouping all parameters in a single group $\{x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2\}$. Often times the reference prior algorithm produces an improper prior and it is necessary to verify if the posterior is a proper density function in order to make Bayesian inference.

In Section 3 we present a discussion of related work in the literature, such as Kubokawa and Robert (1994) and Ghosh et al. (1995). In Section 4 a Gibbs Sampler scheme for sampling from the posterior distribution is presented. In Section 5 we consider a data set from Johnson and Krishnamoorthy (1996) paper and implement the Bayesian inference comparing the results using the different reference priors.

2 Reference priors for the calibration problem

In the next theorems we obtain the reference priors for each case discussed before and also give conditions for the existence of the posterior distributions and their moments. On the following we consider the notation: $c_x = \sum_{i=1}^n (x_i - \overline{x})^2$ and $u(x_0) = (n+k)c_x + nk(x_0 - \overline{x})^2$.

First we remark that the Fisher information matrix associated with the model given by (1.1) and (1.2), considering the ordering $(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2)$, is given by

$$\mathbf{H}(\boldsymbol{\theta}) = \left(\begin{array}{c|c|c|c} \frac{k}{\sigma^2}\boldsymbol{\beta}'\boldsymbol{\beta} & \frac{k}{\sigma^2}\boldsymbol{\beta}' & \frac{k}{\sigma^2}x_0\boldsymbol{\beta}' & 0 \\ \hline \frac{k}{\sigma^2}\boldsymbol{\beta} & \frac{(n+k)}{\sigma^2}\mathbf{I}_p & \frac{(kx_0+\sum_{i=1}^n x_i)}{\sigma^2}\mathbf{I}_p & \mathbf{0} \\ \hline \frac{k}{\sigma^2}x_0\boldsymbol{\beta} & \frac{(kx_0+\sum_{i=1}^n x_i)}{\sigma^2}\mathbf{I}_p & \frac{(kx_0^2+\sum_{i=1}^n x_i^2)}{\sigma^2}\mathbf{I}_p & \mathbf{0} \\ \hline 0 & \mathbf{0} & \mathbf{0} & \frac{(n+k)p}{2\sigma^4} \end{array}\right). \tag{2.1}$$

Theorem 2.1 *Under the model given by (1.1) and (1.2), the Jeffreys prior is given by*

$$\pi_1(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) \propto u(x_0)^{(p-1)/2} (\sigma^2)^{-(2p+3)/2} (\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}.$$

Proof. We calculate the determinant of the Fisher information matrix, partitioning (2.1) as

$$\mathbf{H}(\boldsymbol{\theta}) = \left(\begin{array}{c|c} \mathbf{h}_{11} & \mathbf{h}_{12} \\ \hline \mathbf{h}_{21} & \mathbf{h}_{22} \end{array}\right), \tag{2.2}$$

where $\mathbf{h}_{11} = k\boldsymbol{\beta}'\boldsymbol{\beta}/\sigma^2$, $\mathbf{h}_{12} = \mathbf{h}_{21}' = \left(k\boldsymbol{\beta}'/\sigma^2\ ,\ kx_0\boldsymbol{\beta}/\sigma^2\ ,\ 0\right)$ and

$$\mathbf{h}_{22} = \left(\begin{array}{c|cc} (n+k)/\sigma^2\mathbf{I}_p & (kx_0 + \sum_{i=1}^{n} x_i)/\sigma^2\mathbf{I}_p & \mathbf{0} \\ \hline (kx_0 + \sum_{i=1}^{n} x_i)/\sigma^2\mathbf{I}_p & (kx_0^2 + \sum_{i=1}^{n} x_i^2)/\sigma^2\mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (n+k)p/2\sigma^4 \end{array}\right). \quad (2.3)$$

Then, using matrix proprieties, $|\mathbf{H}(\boldsymbol{\theta})| = |\mathbf{h}_{22}||\mathbf{h}_{11.2}|$, where $|\mathbf{h}_{11.2}| = \mathbf{h}_{11.2} = \mathbf{h}_{11} - \mathbf{h}_{12}\mathbf{h}_{22}^{-1}\mathbf{h}_{21}$. Now partitioning $\mathbf{h}_{22}$ as

$$\mathbf{h}_{22} = \left(\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right),$$

and associating each $\mathbf{A}_{ij}$ with the partition showed in (2.3), we note that $|\mathbf{h}_{22}|$ can be easily calculated since $\mathbf{A}_{22}$ is a diagonal matrix and $|\mathbf{h}_{22}| = |\mathbf{A}_{22}||\mathbf{A}_{11.2}|$ where $\mathbf{A}_{11.2} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$. After simplification we get

$$|\mathbf{h}_{22}| = \frac{u(x_0)^p(n+k)p}{2(\sigma^2)^{2p+2}}. \quad (2.4)$$

Note that the matrix $\mathbf{h}_{22}^{-1}$ can be obtained as

$$\mathbf{h}_{22}^{-1} = \left(\begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} \end{array}\right), \quad (2.5)$$

where

$$\mathbf{B}_{11} = \mathbf{A}_{11.2}^{-1}, \qquad \mathbf{B}_{12} = -\mathbf{A}_{11.2}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1},$$

$$\mathbf{B}_{21} = -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11.2}^{-1}, \qquad \mathbf{B}_{22} = \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11.2}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1},$$

and the matrices $\mathbf{A}_{22}$ and $\mathbf{A}_{11.2}$ are diagonal matrices. Then, after simplifications it follows that

$$\mathbf{h}_{22}^{-1} = \left(\begin{array}{c|c|c} \frac{\sigma^2}{a}(kx_0^2 + \sum_{i=1}^{n} x_i^2)\mathbf{I}_p & -\frac{\sigma^2}{a}(kx_0 + \sum_{i=1}^{n} x_i)\mathbf{I}_p & \mathbf{0} \\ \hline -\frac{\sigma^2}{a}(kx_0 + \sum_{i=1}^{n} x_i)\mathbf{I}_p & \frac{\sigma^2}{a}(n+k)\mathbf{I}_p & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \frac{2\sigma^4}{(n+k)p} \end{array}\right). \tag{2.6}$$

Then,

$$\mathbf{h}_{11.2} = \frac{nc_x k\boldsymbol{\beta}'\boldsymbol{\beta}}{u(x_0)\sigma^2}. \tag{2.7}$$

Therefore, the Jeffreys prior is given by

$$|\mathbf{H}(\boldsymbol{\theta})|^{1/2} \propto u(x_0)^{(p-1)/2}\,(\sigma^2)^{-(2p+3)/2}\,(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}.$$

Theorem 2.2 *Under the model given by (1.1) and (1.2), considering* $\boldsymbol{\Lambda} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2)$ *as a nuisance parameter, the two-group reference prior is given by* $\pi_2(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) \propto u(x_0)^{-1/2}\,(\sigma^2)^{-(p+1)}$.

Proof. Partitioning the Fisher information matrix as in (2.2) and following the notation from the corollary of theorem 5.29 from Bernardo and Smith (1994) we have that

$$\pi(\boldsymbol{\Lambda} \mid x_0) \propto |\mathbf{h}_{22}|^{\frac{1}{2}} = (\sigma^{-2})^{p+1}\left\{\frac{u(x_0)^p(n+k)p}{2}\right\}^{\frac{1}{2}} = g_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2)\, f_2(x_0). \tag{2.8}$$

Now observe that $\mathbf{h}_{x_0} = \mathbf{h}_{11} - \mathbf{h}_{12}\mathbf{h}_{22}^{-1}\mathbf{h}_{21} = \mathbf{h}_{11.2}$ was already calculated on (2.7). In this way

$$|\mathbf{h}_{x_0}|^{1/2} = \left\{\frac{nc_x k\boldsymbol{\beta}'\boldsymbol{\beta}}{\sigma^2}\right\}^{1/2} u(x_0)^{-1/2} = f_1(x_0)\, g_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2). \tag{2.9}$$

Choosing an increasing sequence of subsets $\{\boldsymbol{\Lambda}_i\}$ given by $\boldsymbol{\Lambda}_i = (\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \sigma_i^2)$ where $\boldsymbol{\alpha}_i = [-i, i]^p$, $\boldsymbol{\beta}_i = [-i, i]^p$ and $\sigma_i^2 = [e^{-i}, e^i]$ that do not depend on x_0, we have that the two-group reference prior is given by

$$\pi_2(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) \propto f_1(x_0)\, g_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) = u(x_0)^{-1/2}(\sigma^{-2})^{p+1}.$$

Although the approach of more than two groups might produce different priors for different orderings of the nuisance parameters, Bernardo (1997) points out that most often reference priors are invariant with respect to the nuisance parameters orderings. The next result shows that this invariance propriety is valid for the linear calibration model.

Theorem 2.3 *Under the model given by (1.1) and (1.2), the four-group reference prior, associated with the six different orderings of nuisance parameters does not depend on p. It is given by* $\pi_4(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) \propto u(x_0)^{-1/2} (\sigma^2)^{-1}$.

Proof. Here we are following the algorithm proposed by Berger and Bernardo (1992b). First, we consider the ordering $(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2)$ and denote the Fisher information matrix given in (2.1) by $\mathbf{H}_4(\boldsymbol{\theta})$. Using Berger and Bernardo (1992b) notation, we recall that $\mathbf{h}_j(\boldsymbol{\theta})$ is the right down submatrix of $\mathbf{H}_j(\boldsymbol{\theta})$, $j = 1, 2, 3, 4$. Then,

$$|\mathbf{h}_4(\boldsymbol{\theta})|^{1/2} = \sigma^{-2}\{(n+k)p/2\}^{1/2} = a_4(\sigma^2) b_4(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}). \tag{2.10}$$

We need now to calculate $\mathbf{S}_4(\boldsymbol{\theta}) = \mathbf{H}_4^{-1}(\boldsymbol{\theta})$. Remembering that $\mathbf{H}_4(\boldsymbol{\theta})$ can be partitioned as in (2.2) we have that

$$\mathbf{S}_4(\boldsymbol{\theta}) = \left(\begin{array}{c|c} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \hline \mathbf{C}_{21} & \mathbf{C}_{22} \end{array}\right), \tag{2.11}$$

where

$$\mathbf{C}_{11} = \mathbf{h}_{11.2}^{-1}, \qquad \mathbf{C}_{12} = -\mathbf{h}_{11.2}^{-1}\mathbf{h}_{12}\mathbf{h}_{22}^{-1},$$

$$\mathbf{C}_{21} = -\mathbf{h}_{22}^{-1}\mathbf{h}_{21}\mathbf{h}_{11.2}^{-1}, \qquad \mathbf{C}_{22} = \mathbf{h}_{22}^{-1}\mathbf{h}_{21}\mathbf{h}_{11.2}^{-1}\mathbf{h}_{12}\mathbf{h}_{22}^{-1} + \mathbf{h}_{22}^{-1},$$

$$\mathbf{h}_{11.2} = \mathbf{h}_{x_0} = \mathbf{h}_{11} - \mathbf{h}_{12}\mathbf{h}_{22}^{-1}\mathbf{h}_{21}.$$

After some algebra $\mathbf{S}_4(\boldsymbol{\theta})$ is given by

$$\mathbf{S}_4(\boldsymbol{\theta}) = \left(\begin{array}{c|c|c|c} \frac{u(x_0)\sigma^2}{nc_x k}\boldsymbol{\beta}'\boldsymbol{\beta} & -\frac{\sigma^2\boldsymbol{\beta}' b}{nc_x}\boldsymbol{\beta}'\boldsymbol{\beta} & -\frac{\sigma^2\boldsymbol{\beta}' c}{nc_x}\boldsymbol{\beta}'\boldsymbol{\beta} & 0 \\ \hline -\frac{\sigma^2\boldsymbol{\beta}' b}{nc_x}\boldsymbol{\beta}'\boldsymbol{\beta} & f(x_0, \boldsymbol{\beta}, \sigma^2) & g(x_0, \boldsymbol{\beta}, \sigma^2) & \mathbf{0} \\ \hline -\frac{\sigma^2\boldsymbol{\beta}' c}{nc_x}\boldsymbol{\beta}'\boldsymbol{\beta} & g(x_0, \boldsymbol{\beta}, \sigma^2) & h(x_0, \boldsymbol{\beta}, \sigma^2) & \mathbf{0} \\ \hline 0 & \mathbf{0} & \mathbf{0} & \frac{2\sigma^4}{(n+k)p} \end{array}\right), \tag{2.12}$$

where

$$b = \sum_{i=1}^{n} x_i^2 - x_0 \sum_{i=1}^{n} x_i, \qquad c = n x_0 - \sum_{i=1}^{n} x_i,$$

$$d = \sum_{i=1}^{n} x_i^2, \qquad e = \sum_{i=1}^{n} x_i,$$

$$f(x_0, \boldsymbol{\beta}, \sigma^2) = \frac{\sigma^2 k \boldsymbol{\beta}\boldsymbol{\beta}' b^2}{u(x_0) n c_x \boldsymbol{\beta}'\boldsymbol{\beta}} + \frac{\sigma^2 (k x_0^2 + d)}{u(x_0)} \mathbf{I}_p,$$

$$g(x_0, \boldsymbol{\beta}, \sigma^2) = \frac{\sigma^2 k \boldsymbol{\beta}\boldsymbol{\beta}' bc}{u(x_0) n c_x \boldsymbol{\beta}'\boldsymbol{\beta}} - \frac{\sigma^2 (k x_0 + e)}{u(x_0)} \mathbf{I}_p,$$

$$h(x_0, \boldsymbol{\beta}, \sigma^2) = \frac{\sigma^2 k \boldsymbol{\beta}\boldsymbol{\beta}' c^2}{u(x_0) n c_x \boldsymbol{\beta}'\boldsymbol{\beta}} + \frac{\sigma^2 (n + k)}{u(x_0)} \mathbf{I}_p.$$

Obtaining $\mathbf{H}_3(\boldsymbol{\theta}) = \mathbf{S}_3(\boldsymbol{\theta})^{-1}$ (where $\mathbf{S}_3(\boldsymbol{\theta})$ is obtained by the elimination of the last row and column from matrix $\mathbf{S}_4(\boldsymbol{\theta})$) we find that

$$|\mathbf{h}_3(\boldsymbol{\theta})|^{1/2} = 1 \times \left\{ \frac{\left\{k x_0^2 + \sum_{i=1}^{n} x_i^2\right\}^p}{\sigma^{2p}} \right\}^{1/2} = a_3(\boldsymbol{\beta}) b_3(x_0, \boldsymbol{\alpha}, \sigma^2). \tag{2.13}$$

Similarly, obtaining $\mathbf{H}_2(\boldsymbol{\theta}) = \mathbf{S}_2(\boldsymbol{\theta})^{-1}$ (where $\mathbf{S}_2(\boldsymbol{\theta})$ is obtained by the elimination of the last $p+1$ rows and columns from matrix $\mathbf{S}_3(\boldsymbol{\theta})$) we find that

$$|\mathbf{h}_2(\boldsymbol{\theta})|^{1/2} = 1 \times |\mathbf{h}_2(\boldsymbol{\theta})|^{1/2} = a_2(\boldsymbol{\alpha}) b_2(x_0, \boldsymbol{\beta}, \sigma^2). \tag{2.14}$$

Note that we do not really need to invert $\mathbf{S}_2(\boldsymbol{\theta})$ and calculate $\mathbf{h}_2(\boldsymbol{\theta})$ explicitly at this step because $\mathbf{S}_2(\boldsymbol{\theta})$ is not a function of $\boldsymbol{\alpha}$. Finally, since

$$\mathbf{S}_1(\boldsymbol{\theta}) = \frac{a\sigma^2}{n c_x k \boldsymbol{\beta}'\boldsymbol{\beta}},$$

we get

$$\mathbf{h}_1(\boldsymbol{\theta}) = \mathbf{H}_1(\boldsymbol{\theta}) = \mathbf{S}_1^{-1}(\boldsymbol{\theta}) = \frac{n c_x k \boldsymbol{\beta}'\boldsymbol{\beta}}{a\sigma^2}. \tag{2.15}$$

Therefore

$$|\mathbf{h}_1(\boldsymbol{\theta})|^{1/2} = u(x_0)^{-1/2} \left\{ \frac{n c_x k \boldsymbol{\beta}'\boldsymbol{\beta}}{\sigma^2} \right\}^{1/2} = a_1(x_0) b_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2). \tag{2.16}$$

Choosing an increasing sequence of subsets $\{\boldsymbol{\Lambda}_i\}$ given by $\boldsymbol{\Lambda}_i = (\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \sigma_i^2)$ where $\boldsymbol{\alpha}_i = [-i, i]^p$, $\boldsymbol{\beta}_i = [-i, i]^p$ e $\sigma_i^2 = [e^{-i}, e^i]$ that do not depend on x_0, we have that the four-group reference prior is given by

$$\pi_4(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) \;\propto\; a_1(x_0)\; a_2(\boldsymbol{\alpha})\; a_3(\boldsymbol{\beta})\; a_4(\sigma^2) = u(x_0)^{-1/2} \sigma^{-2}.$$

Using the same ideas developed above, for the other five orderings of the parameter vector given in the introduction, we obtained the same expression for the prior reference. The details are omited here to save space.

The reference priors given by Theorems 2.1, 2.2 and 2.3 are improper priors, as it is often the case in reference analysis when the parameter space is unbounded. However the next result shows that the associated posterior distributions are proper.

Theorem 2.4 *The joint reference posterior distributions associated with reference priors π_1, π_2 and π_4 are proper posterior densities.*

Proof. The proof of this theorem is in Appendix.

Ghosh et al. (1995) have shown that for $p = 1$ the marginal reference posterior distribution of x_0, although proper, have infinite mean (and hence all higher order moments). Next result represents a generalization of that result to arbitrary p.

Theorem 2.5 *The x_0's marginal reference posterior distributions associated with reference priors π_2 and π_4 only have finite moments of order strictly smaller than p.*

Proof. We will first show the result for the marginal posterior distribution associated with the two-group reference prior.

Observing that

$$E[x_0^p \mid \mathbf{y}, \mathbf{X}] \propto E_{x_0}\left[x_0^p \left[(n+k)c_x + nk(x_0 - \overline{x})^2\right]^{\frac{(n+k-1)p+3}{2}}\right], \tag{2.17}$$

and

$$x_0^p \left[(n+k)c_x + nk(x_0 - \overline{x})^2\right]^{\frac{(n+k-1)p+3}{2}} = a_0 + a_1\, x_0 + a_2\, x_0^2 + \ldots + a_{(n+k)p+3}\, x_0^{(n+k)p+3}$$

represents a polinomial of order $(n+k)p+3$, where

$$x_0 \sim t_v\left(l_7\,,\, \frac{l_8}{v(nkl_6 - \boldsymbol{l}_5'\boldsymbol{l}_5)}\right), \tag{2.18}$$

with $\boldsymbol{l}_5, l_6, l_7$ and l_8 defined in the proof of the Theorem 2.4 in Appendix. Where $t_v(a, b)$ denoted a t-Student distribution with location a, scale b and degree of freedom v. The result (2.18) is showed in the proof of Theorem 2.4. It follows that $E[x_0^p \mid \mathbf{y}, \mathbf{X}]$ can be rewritten as a linear combination of moments of order less or equal to $(n+k)p+3$ from the distribution (2.18). Since x_0 has degree of freedom equal to $v = (n+k)p+3$, and from t-Student proprieties the moment of order $(n+k)p+3$ diverges, then $E[x_0^p \mid \mathbf{y}, \mathbf{X}] = \infty$. Recall that, when the moment of order i does not exists all other moments of order greater than i also do not exist. Therefore, $\pi_2(x_0 \mid \mathbf{y}, \mathbf{X})$ does not have any moment of order greater or equal to p.

Now, observing that

$$E[x_0^{p-1} \mid \mathbf{y}, \mathbf{X}] \propto E_{x_0}\left[x_0^{p-1} \left[(n+k)c_x + nk(x_0 - \overline{x})^2\right]^{\frac{(n+k-1)p+3}{2}}\right], \tag{2.19}$$

and that

$$x_0^{p-1}\left[(n+k)c_x+nk(x_0-\overline{x})^2\right]^{\frac{(n+k-1)p+3}{2}}=a_0+a_1\,x_0+\ldots+a_{(n+k)p+2}\,x_0^{(n+k)p+2},$$

represents a polynomial of order $(n+k)p+2$, we have that $E[x_0^{p-1}\mid\mathbf{y},\mathbf{X}]$ can be rewritten as a linear combination of moments from a t-Student distribution of order less or equal to $(n+k)p+2$. Since $v=(n+k)p+3$, then $E[x_0^{p-1}\mid\mathbf{y},\mathbf{X}]<\infty$.

To show the marginal posterior distribution associated with the four-group prior π_4 is proper, we only have to change n by $n-2$ in all equations of the above proof.

The last theorem shows that as we increase p, we automatically increase the number of finite moments of the reference posteriors, that is, we produce more informative posterior distributions. This result is expected because when we increase p, we increase the quantity of information embodied in the model (remember that p represents the number of less accurate measurement techniques used to infer the value of x_0), as a consequence, we produce a more informative reference prior.

A question that arises naturally at this point is what prior should be used. Jeffreys prior may be criticized on the grounds that it does not make distinction between interest and nuisance parameters. Berger and Bernardo (1992a) recommend to use the more than two-group approach based on their empirical experience, especially when that approach is invariant to the nuisance parameters orderings (producing a unique prior). However, Ghosh et al. (1995) argues that the two-group approach appears to produce the most intuitively appealing reference prior, because in the lack of additional information about the nuisance parameters importance it seems more natural to group them all together. In Section 4 we compare these priors through an example.

3 Comparison with other reference priors

In the statistical literature (see Yang and Berger, 1997 and Bernardo and Ramón, 1998) the reference prior for the p-dimensional linear calibration problem is credited to Kubokawa and Robert (1994). In fact, they obtained the reference prior associated with the related problem of the estimation of the ratio of two normal means. However, we remark here that it is not the reference prior for the real multivariate calibration problem. Kubokawa and Robert reduced by sufficience the original model, given by (1.1) and (1.2), to the following model $\overline{\mathbf{y}}\sim N_p\left(\boldsymbol{\alpha}+\boldsymbol{\beta}\overline{x}\,,\,\frac{\sigma^2}{n}\mathbf{I}_p\right)$, $\overline{\mathbf{y}}_0\sim N_p\left(\boldsymbol{\alpha}+\boldsymbol{\beta}x_0\,,\,\frac{\sigma^2}{k}\mathbf{I}_p\right)$, $\hat{\boldsymbol{\beta}}\sim N_p\left(\boldsymbol{\beta}\,,\,\frac{\sigma^2}{c_x}\mathbf{I}_p\right)$ and $s\sim Gama\left(\frac{q}{2}\,,\,\frac{1}{2\sigma^2}\right)$ where $\hat{\boldsymbol{\beta}}$, $\overline{\mathbf{y}}$, $\overline{\mathbf{y}}_0$ and s are mutually independent and $q=(n+k-3)p$, $\hat{\boldsymbol{\beta}}=\sum_{i=1}^n(x_i-\overline{x})(y_i-\overline{\mathbf{y}})/c_x$, $\overline{x}=\sum_{i=1}^n x_i/n$, $\overline{\mathbf{y}}=\sum_{i=1}^n\mathbf{y}_i/n$, $\overline{\mathbf{y}}_0=\sum_{j=1}^k\mathbf{y}_{0j}/k$, $c_x=\sum_{i=1}^n(x_i-\overline{x})^2$. From this reduced model, Kubokawa and Robert considered the reparametrization $\boldsymbol{\beta}^*=\sqrt{c_x}\boldsymbol{\beta}$, $x_0^*=(x_0-\overline{x})/\sqrt{c_x(\frac{1}{n}+\frac{1}{k})}$, $\mathbf{w}=\sqrt{c_x}\hat{\boldsymbol{\beta}}$, $\mathbf{z}=(\overline{\mathbf{y}}_0-\overline{\mathbf{y}})/\sqrt{\frac{1}{n}+\frac{1}{k}}$ where

$\mathbf{w}$, $\mathbf{z}$ and s are independent and distributed as $\mathbf{w} \sim N_p\left(\boldsymbol{\beta}^*\ ,\ \sigma^2\mathbf{I}_p\right)$, $\mathbf{z} \sim N_p\left(x_0^*\boldsymbol{\beta}^*\ ,\ \sigma^2\mathbf{I}_p\right)$, $s \sim Gama\left(\frac{q}{2}\ ,\ \frac{1}{2\sigma^2}\right)$. From this reduced and reparametrized model Kubokawa and Robert obtained the Jeffreys prior $\pi_j^k(x_0^*, \boldsymbol{\beta}^*, \sigma^2) \propto (1 + x_0^{*2})^{(p-1)/2}\,(\sigma^2)^{-(p+3)/2}\,\boldsymbol{\beta}^{*\prime}\boldsymbol{\beta}^*$ and the two-group reference prior $\pi_2^k(x_0^*, \boldsymbol{\beta}^*, \sigma^2) \propto (1 + x_0^{*2})^{-1/2}(\sigma^{-2})^{(p+2)/2}$. Transforming back to the original parametrization through the jacobian element $\left|\partial(x_0^*, \boldsymbol{\beta}^*, \sigma^2)/\partial(x_0, \boldsymbol{\beta}, \sigma^2)\right| = |\sqrt{nk/(n+k)}|$, Kubokawa and Robert's priors do not coincide with ours. Table 1 presents a summary of the reference priors for the calibration problem.

Table 1 *Reference priors in the original parametrization* $(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2)$

	Kubokawa and Robert	Chaibub and Branco	Ghosh et al.
Jeffreys	$u(x_0)^{\frac{(p-1)}{2}}\,(\sigma^2)^{-\frac{(p+3)}{2}}\,\boldsymbol{\beta}'\boldsymbol{\beta}$	$u(x_0)^{\frac{(p-1)}{2}}\,(\sigma^2)^{-\frac{(2p+3)}{2}}\,(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}$	$(\sigma^2)^{-5/2}\,\|\beta\|$
Two-group	$u(x_0)^{-1/2}\,(\sigma^2)^{-\frac{(p+2)}{2}}$	$u(x_0)^{-1/2}\,(\sigma^2)^{-(p+1)}$	$u(x_0)^{-1/2}\,(\sigma^2)^{-2}$
Four-group	-	$u(x_0)^{-1/2}\,(\sigma^2)^{-1}$	$u(x_0)^{-1/2}\,(\sigma^2)^{-1}$

Although Kubokawa and Robert worked with a model related to the linear calibration model, their reference priors do not coincide with ours after the adequate jacobian transformation is done. This result is expected because, although reference priors are invariant under sufficient statistics and one-to-one parameter transformations, they are not invariant under transformations on the parametric model. In fact, Kubokawa and Robert first reduced the original model to its sufficient statistics $\hat{\boldsymbol{\beta}}$, $\overline{\mathbf{y}}$, $\overline{\mathbf{y}}_0$ and s, what is fine since reference priors are invariant under this type of transformation. The problem arose when the authors changed the model considering the reparametrization $\mathbf{z} = (\overline{\mathbf{y}}_0 - \overline{\mathbf{y}})/\sqrt{\frac{1}{n} + \frac{1}{k}}$. We can also see that the likelihood function associated with the last model is not proportional to the original likelihood.

On the other hand, in the particular case $p = 1$, our reference priors exactly match with the results obtained by Ghosh et al. (1995) (see Table 1). These authors presented the complete list of reference priors associated with all possible groupings and ordinations of nuisance parameters for the linear calibration model when $p = 1$. In their paper Ghosh et al. considered the reparametrization $(x_0, \alpha, \beta, \sigma^2) \to (x_0, \alpha_0, \beta, \sigma^2)$, where $\alpha_0 = \alpha + \beta\overline{x}$. Nonetheless, since α_0 is an one-to-one reparametrization of α, Ghosh's reference prior can be transformed back to the original parametrization through the jacobian element $\left|\partial(x_0, \alpha_0, \beta, \sigma^2)/\partial(x_0, \alpha, \beta, \sigma^2)\right|$ and, as expected, the transformed reference prior coincides exactly with our priors.

4 The Gibbs sampler algorithm

Since there is not a closed form for the marginal posterior distributions, to implement the Bayesian inference under the reference priors we will consider a Monte Carlo Markov Chain (MCMC) method. The existence of the joint posterior density, proved in Theorem 4, is a essential condition to use a MCMC method. Here we will develop a Gibbs Sampler (GS) algortihm with some Metropolis Hastings steps.

Note that, for all reference priors π_1, π_2 and π_4, the joint posterior distribution is proportional to

$$\pi_r \times (\sigma^{-2})^{(n+k)p/2} \exp\left\{-\left[\sum_{i=1}^{n}(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i) + \right.\right.$$
$$\left.\left. + \sum_{j=1}^{k}(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)'(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)\right] \Big/ 2\sigma^2\right\}, \qquad r = 1, 2 \text{ or } 4 \tag{4.1}$$

Then, we can easily see that

$$\sigma^{-2} \mid x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{y}, \mathbf{X} \sim \mathit{Gamma}\left(\frac{(n+k)p+4}{2}, \frac{l_1}{2}\right) \qquad \text{for } r = 4, \tag{4.2}$$

$$\sigma^{-2} \mid x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{y}, \mathbf{X} \sim \mathit{Gamma}\left(\frac{(n+k+2)p+4}{2}, \frac{l_1}{2}\right) \qquad \text{for } r = 2, \tag{4.3}$$

$$\sigma^{-2} \mid x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{y}, \mathbf{X} \sim \mathit{Gamma}\left(\frac{(n+k+2)p+5}{2}, \frac{l_1}{2}\right) \qquad \text{for } r = 1, \tag{4.4}$$

where $l_1 = \sum_{i=1}^{n}(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i) + \sum_{j=1}^{k}(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)'(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)$. The conditional distribution of $\boldsymbol{\alpha}$ is proportional to $\exp\left\{-\frac{l_1}{2\sigma^2}\right\}$ and after simplification it follows that

$$\boldsymbol{\alpha} \mid x_0, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X} \sim N_p\left(\boldsymbol{l}_2, \frac{\sigma^2}{n+k}\mathbf{I}_p\right) \qquad \text{for } r = 1, 2, 4, \tag{4.5}$$

where $\boldsymbol{l}_2 = [\sum_{i=1}^{n}(\mathbf{y}_i - \boldsymbol{\beta}x_i) + \sum_{j=1}^{k}(\mathbf{y}_{0j} - \boldsymbol{\beta}x_0)]/(n+k)$.
The conditional distribution of $\boldsymbol{\beta}$ is also proportional to $\exp\left\{-\frac{l_1}{2\sigma^2}\right\}$ for $r = 2, 4$ and after simplification it follows that

$$\boldsymbol{\beta} \mid x_0, \boldsymbol{\alpha}, \sigma^2, \mathbf{y}, \mathbf{X} \sim N_p\left(\boldsymbol{l}_3, \frac{\sigma^2}{kx_0^2 + \sum_{i=1}^{n} x_i^2}\mathbf{I}_p\right) \qquad \text{for } r = 2 \text{ and } 4, \tag{4.6}$$

where $\boldsymbol{l}_3 = [\sum_{i=1}^{n} x_i(\mathbf{y}_i - \boldsymbol{\alpha}) + x_0\sum_{j=1}^{k}(\mathbf{y}_{0j} - \boldsymbol{\alpha})]/(kx_0^2 + \sum_{i=1}^{n} x_i^2)$. When $r = 1$ the conditional distribution of $\boldsymbol{\beta}$ is not a known and is proportional to

$$(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2} \times \exp\left\{-\frac{(\boldsymbol{\beta} - \boldsymbol{l}_3)'(\boldsymbol{\beta} - \boldsymbol{l}_3)}{2\sigma^2/(kx_0^2 + \sum_{i=1}^{n} x_i^2)}\right\}.$$

However we can easily generate from this conditional using a Metropolis-Hastings step inside the GS algorithm.

When $r = 2, 4$ the conditional distribution of x_0 is proportional to

$$u(x_0)^{-1/2} \times \exp\left\{-\frac{(x_0 - l_4)^2}{2\sigma^2/k\boldsymbol{\beta}'\boldsymbol{\beta}}\right\}, \tag{4.7}$$

where $l_4 = \boldsymbol{\beta}' \sum_{j=1}^{k} (\mathbf{y}_{0j} - \boldsymbol{\alpha})/(k\boldsymbol{\beta}'\boldsymbol{\beta})$. Note that (4.7) is not a known distribution of probability. Following Kubokawa and Robert (1994) we introduce a latent variable t as follows

$$u(x_0)^{-1/2} \propto \Gamma(1/2)\, u(x_0)^{-1/2} = \int_0^\infty t^{\frac{1}{2}-1} \exp\{-u(x_0)\, t\}\, dt. \tag{4.8}$$

We now can rewrite (4.7) as

$$p(x_0 \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X}) = \int_0^\infty p(x_0, t \mid \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X})\, dt, \tag{4.9}$$

where

$$p(x_0, t \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X}) \propto t^{\frac{1}{2}-1} \exp\{-u(x_0)\, t\} \exp\left\{-\frac{(x_0 - l_4)^2}{2\sigma^2/k\boldsymbol{\beta}'\boldsymbol{\beta}}\right\}. \tag{4.10}$$

From (4.10) we have that

$$t \mid x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X} \sim \mathit{Gamma}\left(\frac{1}{2}\,,\, u(x_0)\right), \tag{4.11}$$

and

$$p(x_0 \mid t, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X}) \propto \exp\{-u(x_0)\, t\} \exp\left\{-\frac{(x_0 - l_4)^2}{2\sigma^2/k\boldsymbol{\beta}'\boldsymbol{\beta}}\right\}. \tag{4.12}$$

Therefore,

$$x_0 \mid t, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X} \sim N\left(\frac{2\sigma^2 nt\overline{x} + l_4\boldsymbol{\beta}'\boldsymbol{\beta}}{2\sigma^2 nt + \boldsymbol{\beta}'\boldsymbol{\beta}}\,,\, \frac{\sigma^2}{2\sigma^2 nkt + k\boldsymbol{\beta}'\boldsymbol{\beta}}\right). \tag{4.13}$$

When $r = 1$ it is not possible to consider the latent variable t. However again we can use a Metropolis-Hastings step inside the GS algorithm to generate from the conditional distribution $u(x_0)^{(p-1)/2} \times \exp\left\{-\frac{(x_0 - l_4)^2}{2\sigma^2/k\boldsymbol{\beta}'\boldsymbol{\beta}}\right\}$.

5 Application

We consider here a data set given by Johnson and Krishnamoorthy (1996) where a controlled experiment was conducted at National Biological Service, Louisiana, to predict the amount of sodium chloride solution in dionized water based on electric conductivity. Two machines were used to measure the electric conductivity: the Conductivity Controller(CC) and the Fisher Conductivity Meter(FCM). The data are presented in Table 2, where x is the amount of sodium chloride solution (in ml), y_1 is the CC measure and y_2 is the FCM measure, the last two given in $micromoles/cm^3$.

Table 2 *R0.95 posterior credibility intervals for* x_0

x_0	y_1	y_2	Four-group	Two-group	Jeffreys
0.0	1.6	1.5	(−1.45, 0.11)	(−1.41, 0.09)	(−1.39, 0.07)
0.5	1.8	1.9	(−0.76, 0.77)	(−0.75, 0.74)	(−0.73, 0.75)
1.0	2.0	2.2	(−0.24, 1.30)	(−0.19, 1.28)	(−0.21, 1.28)
1.5	2.2	2.6	(0.44, 1.99)	(0.46, 1.94)	(0.47, 1.93)
2.0	2.4	2.9	(0.97, 2.49)	(1.01, 2.46)	(1.01, 2.46)
2.5	2.6	3.2	(1.50, 3.00)	(1.52, 3.00)	(1.54, 3.02)
3.0	2.8	3.6	(2.17, 3.70)	(2.18, 3.67)	(2.18, 3.66)
3.5	3.0	3.9	(2.67, 4.20)	(2.71, 4.17)	(2.70, 4.18)
4.0	3.2	4.2	(3.21, 4.73)	(3.24, 4.69)	(3.24, 4.70)
4.5	3.4	4.5	(3.73, 5.24)	(3.76, 5.24)	(3.77, 5.23)
5.0	3.6	4.8	(4.26, 5.78)	(4.26, 5.75)	(4.31, 5.72)
5.5	3.8	5.2	(4.91, 6.42)	(4.94, 6.40)	(4.96, 6.38)
6.0	3.9	5.5	(5.39, 6.88)	(5.41, 6.86)	(5.40, 6.85)
6.5	4.1	5.8	(5.91, 7.40)	(5.92, 7.39)	(5.94, 7.38)
7.0	4.3	6.1	(6.42, 7.92)	(6.45, 7.91)	(6.45, 7.90)
7.5	4.5	6.4	(6.96, 8.44)	(6.97, 8.42)	(6.99, 8.43)
8.0	4.6	6.7	(7.40, 8.91)	(7.44, 8.88)	(7.42, 8.87)
8.5	4.8	7.0	(7.92, 9.41)	(7.96, 9.40)	(7.97, 9.41)
9.0	5.0	7.3	(8.45, 9.95)	(8.47, 9.94)	(8.49, 9.91)
9.5	5.1	7.6	(8.91, 10.41)	(8.91, 10.37)	(8.93, 10.37)
10.0	5.3	7.9	(9.44, 10.93)	(9.43, 10.89)	(9.44, 10.90)
11.0	5.6	8.5	(10.41, 11.90)	(10.41, 11.88)	(10.43, 11.87)
12.0	6.0	9.1	(11.45, 12.94)	(11.47, 12.94)	(11.49, 12.92)
13.0	6.3	9.7	(12.42, 13.93)	(12.44, 13.90)	(12.44, 13.91)
14.0	6.6	11.0	(14.39, 15.82)	(14.42, 15.81)	(14.44, 15.79)
15.0	6.9	11.4	(15.07, 16.55)	(15.11, 16.52)	(15.10, 16.52)
16.0	7.2	11.6	(15.47, 17.00)	(15.50, 16.99)	(15.53, 16.99)
17.0	7.5	12.0	(16.17, 17.71)	(16.17, 17.71)	(16.18, 17.67)
18.0	7.7	13.0	(17.62, 19.21)	(17.67, 19.18)	(17.67, 19.15)
20.0	8.2	14.0	(19.19, 20.80)	(19.22, 20.75)	(19.22, 20.75)
24.0	9.1	15.0	(20.90, 21.76)	(20.92, 21.73)	(20.92, 21.74)

A cross-validation tecnique was considered to compare the different priors. Each time a value of x from Table 2 was considered unkown and it was estimated using the posterior distribution. For all the priors considered the posterior means (and also medians) were very close to each other. On Table 2 columns 4, 5 and 6 present the equal-tail credibility interval with probability 0.95 for the four-group, two-group and Jeffreys priors, respectively. As we can see that the four-group posterior analysis gave intervals slighty wider than the others two, although smaller than the classical intervals given by Johnson and Krishnamoorthy (1996). This indicate that the four-group prior is the least informative prior, in the sense to be closer of the likelihood inference methods. These results agree with Ghosh et al. (1995), for the case $p = 1$.

6 Final comments

Although the reference priors developed by Kubokawa and Robert (1994) are known in the literature as the reference priors for the linear calibration problem (see Yang and Berger, 1997), we discuss in this paper that in reality they are reference priors associated with the related problem of the estimation of the ratio of two normal means. These two problems are not equivalent because the likelihood functions are not proportional. Also, we extend the results given by Ghosh et al. (1995) proving that the reference posteriors are proper but the existence of the posterior moments depends on the response vector dimension. In our example, the four-group reference prior is less informative than the two-group and Jeffreys prior. It is in accordance with Ghosh et al. results for $p = 1$.

Appendix

In this Appendix we present with details the proof of the Theorem 2.4. First we prove the result for the posterior distribution associated with π_2. Note that, the posterior distribution $\pi_2(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}, \mathbf{X})$ is proportional to

$$u(x_0)^{-1/2}(\sigma^{-2})^{\frac{(n+k+2)p+2}{2}} \exp\left\{-\left[\sum_{i=1}^{n}(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i) + \right.\right.$$
$$\left.\left. + \sum_{j=1}^{k}(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)'(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)\right] \Big/ 2\sigma^2\right\}. \qquad (A.1)$$

Integrating the above expression in relation to σ^{-2} and using integration results associated with the Gamma family we have that

$$\pi_2(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta} \mid \mathbf{y}, \mathbf{X}) \propto u(x_0)^{-1/2}\left\{\sum_{i=1}^{n}(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i) + \right.$$
$$\left. + \sum_{j=1}^{k}(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)'(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)\right\}^{-\frac{(n+k+2)p+4}{2}}. \qquad (A.2)$$

After some algebra and using integration results associated with the multivariate t-Student family of distributions to integrate in $\boldsymbol{\alpha}$ we have that $\pi_2(x_0, \boldsymbol{\beta} \mid \mathbf{y}, \mathbf{X})$ is proportional to

$$u(x_0)^{-1/2}\left\{(n+k)\left[\sum_{i=1}^{n}(\mathbf{y}_i-\boldsymbol{\beta}x_i)'(\mathbf{y}_i-\boldsymbol{\beta}x_i)+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)'(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)\right]\right.$$
$$\left.-\left(\sum_{i=1}^{n}(\mathbf{y}_i-\boldsymbol{\beta}x_i)+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)\right)'\left(\sum_{i=1}^{n}(\mathbf{y}_i-\boldsymbol{\beta}x_i)+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)\right)\right\}^{-\frac{(n+k+1)p+4}{2}}. \tag{A.3}$$

Finally, after some algebra and using integration results associated with the multivariate t-Student family of distributions to integrate in $\boldsymbol{\beta}$ we have that

$$\pi_2(x_0 \mid \mathbf{y}, \mathbf{X}) \;\propto\; u(x_0)^{\frac{(n+k-1)p+3}{2}}\left(1+\frac{1}{v}\frac{(x_0-l_7)^2}{l_8/(nkl_6-\boldsymbol{l}_5'\boldsymbol{l}_5)v}\right)^{-\frac{v+1}{2}}, \tag{A.4}$$

where
$v=(n+k)p+3$,

$$\boldsymbol{l}_4=(n+k)\sum_{i=1}^{n}\mathbf{y}_i x_i-\sum_{i=1}^{n}x_i\left(\sum_{i=1}^{n}\mathbf{y}_i+\sum_{j=1}^{k}\mathbf{y}_{0j}\right).$$

$$\boldsymbol{l}_5=n\sum_{j=1}^{k}\mathbf{y}_{0j}-k\sum_{i=1}^{n}\mathbf{y}_i,$$

$$l_6=(n+k)(\sum_{i=1}^{n}\mathbf{y}_i'\mathbf{y}_i+\sum_{j=1}^{k}\mathbf{y}_{0j}'\mathbf{y}_{0j})-\left(\sum_{i=1}^{n}y_i+\sum_{j=1}^{k}\mathbf{y}_{0j}\right)'\left(\sum_{i=1}^{n}y_i+\sum_{j=1}^{k}\mathbf{y}_{0j}\right),$$

$$l_7=\frac{nkl_6\overline{x}+\boldsymbol{l}_4'\boldsymbol{l}_5}{nkl_6-\boldsymbol{l}_5'\boldsymbol{l}_5},$$

$$l_8=nk\overline{x}^2l_6+(n+k)c_xl_6-\boldsymbol{l}_4'\boldsymbol{l}_4-(nkl_6-\boldsymbol{l}_5'\boldsymbol{l}_5)l_7^2.$$

Note that (A.4) is proportional to a function of x_0 multiplied by the kernel of t-Student distribution. Therefore,

$$\int_{-\infty}^{+\infty}\pi_2(x_0 \mid \mathbf{y}, \mathbf{X})\,dx_0 \;\propto\; E_{x_0}\left[u(x_0)^{\frac{(n+k-1)p+3}{2}}\right], \tag{A.5}$$

where

$$x_0 \;\sim\; t_v\left(l_7\,,\;\frac{l_8}{(nkl_6-\boldsymbol{l}_5'\boldsymbol{l}_5)v}\right). \tag{A.6}$$

Now, using the fact that a t-Student distribution only have moments of order strictly lower than its degrees of freedom (Johnson, Kotz and Balakrishnan, 1995) we can show that (A.5) is finite, observing that

$$\left[(n+k)c_x+nk(x_0-\overline{x})^2\right]^{\frac{(n+k-1)p+3}{2}}=a_0+a_1\,x_0+a_2\,x_0^2+\ldots+a_{(n+k-1)p+3}\,x_0^{(n+k-1)p+3},$$

can be rewritten as a polynomial of degree $(n+k-1)p+3$, and $v = (n+k)p+3$. Therefore,

$$\int_{-\infty}^{+\infty} \pi_2(x_0 \mid \mathbf{y}, \mathbf{X})\, dx_0 \quad \propto \quad \sum_{i=0}^{(n+k-1)p+3} a_i E\left[x_0^i\right] < \infty.$$

To show that the posterior distribution associated with the four-group prior π_4 is proper, we just have to change n by $n-2$ in all equations from the above proof.

Now we are going to show that the posterior distribution associated with the Jeffreys prior π_1 is proper.

The posterior distribution $\pi_1(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}, \mathbf{X})$ is proportional to

$$(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2} u(x_0)^{\frac{p-1}{2}} (\sigma^{-2})^{\frac{(n+k+2)p+3}{2}} \exp\left\{ - \left[\sum_{i=1}^{n} (\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i) + \right.\right.$$
$$\left.\left. + \sum_{j=1}^{k} (\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta} x_0)'(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta} x_0) \right] \Big/ 2\sigma^2 \right\}. \qquad (A.7)$$

Using integration results associated with the family of gamma distributions we have that

$$\pi_1(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta} \mid \mathbf{y}, \mathbf{X}) \propto (\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2} u(x_0)^{\frac{p-1}{2}} \left\{ \sum_{i=1}^{n} (\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i) + \right.$$
$$\left. + \sum_{j=1}^{k} (\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta} x_0)'(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta} x_0) \right\}^{-\frac{(n+k+2)p+5}{2}}. \qquad (A.8)$$

Using integration results associated with the family of multivariate t-Student distributions we have that

$$\pi_1(x_0, \boldsymbol{\beta} \mid \mathbf{y}, \mathbf{X}) \propto (\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2} u(x_0)^{\frac{p-1}{2}} \left\{ (n+k) \left[\sum_{i=1}^{n} (\mathbf{y}_i - \boldsymbol{\beta} x_i)'(\mathbf{y}_i - \boldsymbol{\beta} x_i) + \right.\right.$$
$$\left. + \sum_{j=1}^{k} (\mathbf{y}_{0j} - \boldsymbol{\beta} x_0)'(\mathbf{y}_{0j} - \boldsymbol{\beta} x_0) \right] - \left(\sum_{i=1}^{n} (\mathbf{y}_i - \boldsymbol{\beta} x_i) + \sum_{j=1}^{k} (\mathbf{y}_{0j} - \boldsymbol{\beta} x_0) \right)'$$
$$\left. \times \left(\sum_{i=1}^{n} (\mathbf{y}_i - \boldsymbol{\beta} x_i) + \sum_{j=1}^{k} (\mathbf{y}_{0j} - \boldsymbol{\beta} x_0) \right) \right\}^{-\frac{(n+k+1)p+5}{2}}. \qquad (A.9)$$

Using again integration results associated with the family of multivariate t-Student distributions we have that

$$\pi_1(x_0 \mid \mathbf{y}, \mathbf{X}) \propto E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2} \right] u(x_0)^{\frac{(n+k)p+4}{2}} \left(1 + \frac{1}{v_1} \frac{(x_0 - l_7)^2}{l_8/(nkl_6 - \boldsymbol{l}_5'\boldsymbol{l}_5)v_1} \right)^{-\frac{v_1+1}{2}}, \qquad (A.10)$$

where $v_1 = (n+k)p+4$ and

$$E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}\right] = \int_{\boldsymbol{\beta}} c' \, (\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2} \left(1 + \frac{1}{v_2}\frac{(\boldsymbol{\beta}-\mathbf{m})'(\boldsymbol{\beta}-\mathbf{m})}{s/v_2}\right)^{-\frac{v_2+p}{2}} d\boldsymbol{\beta}, \qquad (A.11)$$

$$c' = \frac{\Gamma(\frac{v_2+p}{2})}{\Gamma(\frac{v_2}{2})\, v_2^{p/2}\, \pi^{p/2}\, |s\mathbf{I}_p|^{1/2}} \qquad \mathbf{m} = \frac{\boldsymbol{l}_4 + \boldsymbol{l}_5 x_0}{a},$$

$$s = \frac{a\, l_6 - (\boldsymbol{l}_4 + \boldsymbol{l}_5 x_0)'(\boldsymbol{l}_4 + \boldsymbol{l}_5 x_0)}{a^2 v_2}, \qquad v_2 = (n+k)p + 5.$$

Observe that (A.11) is a function of x_0 and that

$$\int_{-\infty}^{+\infty} \pi_1(x_0 \mid \mathbf{y}, \mathbf{X})\, dx_0 \;\propto\; E_{x_0}\left[E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}\right] u(x_0)^{\frac{(n+k)p+4}{2}}\right], \qquad (A.12)$$

where

$$x_0 \;\sim\; t_{v_1}\left(l_7\,,\, \frac{l_8}{v_1(nkl_6 - \boldsymbol{l}_5'\boldsymbol{l}_5)}\right). \qquad (A.13)$$

Since the square root is a concave function, it follows from Jensen inequality that

$$E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}\right] \;\le\; E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})\right]^{1/2}. \qquad (A.14)$$

Using the result (Fang and Zhang, 1990, pg 42)

$$E\left[\mathbf{x}'\mathbf{A}\mathbf{x}\right] = tr(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}, \qquad (A.15)$$

where the distribution of vector $\mathbf{x}$ has a vector of means $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, we have for $\mathbf{A} = \mathbf{I}_p$ that

$$E\left[\boldsymbol{\beta}'\boldsymbol{\beta}\right] = tr(s\mathbf{I}_p) + \mathbf{m}'\mathbf{m} =$$

$$= \frac{\{(n+k)c_x + nk(x_0 - \overline{x})^2\}l_6 - m_3(x_0+m_1)^2 - m_2}{\{(n+k)c_x + nk(x_0-\overline{x})^2\}^2 v} p + \\ + \frac{m_3(x_0+m_1)^2 + m_2}{\{(n+k)c_x + nk(x_0-\overline{x})^2\}^2} \qquad (A.16)$$

where

$$m_1 = \frac{\boldsymbol{l}_4'\boldsymbol{l}_5}{\boldsymbol{l}_5'\boldsymbol{l}_5}, \qquad m_2 = \boldsymbol{l}_4'\boldsymbol{l}_4 - \frac{(\boldsymbol{l}_4'\boldsymbol{l}_5)^2}{\boldsymbol{l}_5'\boldsymbol{l}_5}, \qquad m_3 = \boldsymbol{l}_5'\boldsymbol{l}_5,$$

are not functions of x_0. In this way, we have that $E\left[\boldsymbol{\beta}'\boldsymbol{\beta}\right]$ corresponds to a polinomial of degree -2, when it is seen as a function of x_0. Then, we have that

$$E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}\right] \;\le\; \underbrace{E\left[\boldsymbol{\beta}'\boldsymbol{\beta}\right]^{1/2}}_{\text{polinomial of degree -1}}.$$

Therefore,

$$E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{\frac{1}{2}}\right]\left[(n+k)c_x + nk(x_0-\overline{x})^2\right]^{\frac{(n+k)p+4}{2}}$$
$$\leq \underbrace{E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})\right]^{\frac{1}{2}}\left[(n+k)c_x + nk(x_0-\overline{x})^2\right]^{\frac{(n+k)p+4}{2}}}_{\text{polinomial of degree } (n+k)p+3}.$$

Recalling that the t-Student distribution have finite moments of order strictly lower than its degrees of freedom, we have that

$$E_{x_0}\left[E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})\right]^{1/2}\left[(n+k)c_x + nk(x_0-\overline{x})^2\right]^{\frac{(n+k)p+4}{2}}\right] < \infty$$

since the above equation can be rewritten as a linear combination of moments of order lower or equal to $(n+k)p+3$ from a t-Student distribution with $v_1 = (n+k)p+4$ degrees of freedom. As a consequence from the Jensen inequality

$$\int_{-\infty}^{+\infty} \pi_1(x_0 \mid \mathbf{y}, \mathbf{X})\, dx_0 \ < \infty.$$

Acknowledgements

The first author was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) during his master program under supervision of the second author.

(Received September, 2005. Accepted November, 2006.)

References

Berger, J. O. and Bernardo, J. M. (1989). Estimating a product of means: Bayesian analysis with reference priors. *Journal of the American Statistical Association*, **84**, 200-207.

Berger, J. O. and Bernardo, J. M. (1992a). On the development of reference priors. *Bayesian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawid, e A. F. M. Smith, eds). Oxford University Press, 35-60 (with discussion).

Berger, J. O. and Bernardo, J. M. (1992b). Ordered group reference priors with application to the multinomial problem. *Biometrika*, **79**, 25-37.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. *Journal of Royal Statistical Society B*, **41**, 113-147 (with discussion).

Bernardo, J. M. (1997). Non-informative priors do not exist. *Journal of Statistics Planning and Inference*, **65**, 113-147 (with discussion).

Bernardo, J. M. and Ramón, J. M. (1998). An introduction to Bayesian reference analysis: inference on the ratio of multinomial parameters. *The Statistician*, **47**, 101-135.

Bernardo, J. M. and Smith, A. (1994). *Bayesian Theory.* Wiley.

Brown, P. J. (1993). *Measurement, Regression and Calibration.* Oxford University Press.

Fang, K. T. and Zhang, Y.,T.(1990). *Generalized Multivariate Analysis.* Beijing: Springer-Verlag.

Ghosh, M. Carlin, B. P. and Srivastava, M. S. (1995). Probability matching priors for linear calibration. *Test*, **4**, 333-357.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions.* vol. 2, 2nd edition. New York: Wiley.

Johnson, D.J. and Krishnamoorthy, K. (1996). Combining independent studies in calibration problem. *Journal of the American Statistical Association*, **91**, 1707-1715.

Kass, R.E. and Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, **91**, 1343-1370.

Kubokawa, T. and Robert, C. (1994). New perspectives on linear calibration. *Journal of Multivariate Analysis*, **50**, 178-200.

Osborne, C. (1991). Statistical calibration: a review. *International Statistical Review*, **59**, 309-336.

Yang, R. and Berger, J. O. (1997). A catalog of noninformative priors. *ISDS Discussion Paper* 97-42.

Elias Chaibub Neto and **Márcia D'Elia Branco**
Departamento de Estatística
Instituto de Matemática e Estatística
Universidade de São Paulo
Rua do Matão, 1010 Cidade Universitária
05508-090 São Paulo/SP, Brazil
E-mails: echaibub@hotmail.com and mbranco@ime.usp.br