Reference analysis for the p-dimensional linear calibration problem

Elias Chaibub Neto and Márcia D'Elia Branco[1]

Department of Statistics - University of São Paulo-
Caixa Postal 66281, CEP 05315-970, São Paulo, Brasil.

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 schedule 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 don't agree with Kubokawa and Robert's priors for general p.*

Key Words: *reference priors, linear calibration problem, ratio of two normal means and Gibbs Sampler.*

AMS classification: 62A15, 62J05, 62P30.

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 the first step, a set of values $x_1, x_2, \ldots, x_n$ are fixed for the accurate measurement X and others p related measurement 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

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

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

[1]Correspondence to: Márcia D. Branco, e-mail:mbranco@ime.usp.br

where (1) and (2) represent the calibration and prediction experiments, respectively. The error vectors $\boldsymbol{\epsilon}_i$ and $\boldsymbol{\epsilon}_{0j}$ are independent, σ^2 is unknown, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)'$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$ are the regression parameters. 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 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. 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 Berger and Bernardo(1989) and Berger and Bernardo(1992b) known as reference priors.

In this paper we obtain reference priors for this p-dimensional linear calibration problem, given by (1) and (2). 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. 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 ordination 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 ordinations and then chose 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 $\{x_0\}\{\boldsymbol{\alpha}\}\{\boldsymbol{\beta}\}\{\sigma^2\}$, $\{x_0\}\{\boldsymbol{\alpha}\}\{\sigma^2\}\{\boldsymbol{\beta}\}$, $\{x_0\}\{\boldsymbol{\beta}\}\{\boldsymbol{\alpha}\}\{\sigma^2\}$, $\{x_0\}\{\boldsymbol{\beta}\}\{\sigma^2\}\{\boldsymbol{\alpha}\}$, $\{x_0\}\{\sigma^2\}\{\boldsymbol{\alpha}\}\{\boldsymbol{\beta}\}$, $\{x_0\}\{\sigma^2\}\{\boldsymbol{\beta}\}\{\boldsymbol{\alpha}\}$ of the nuisance parameters and show that the 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 $\{x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2\}$. Note that, reference priors usually are not a density of probability function (improper prior), in that case it is necessary to verify if the posterior will be a density function (proper prior) in order to make Bayesian inference.

In section 3 we consider a data set from Johnson and Krishnamoorthy (1996)paper and implement the Bayesian inference comparing the results using the different reference priors. In section 4 we present a discussion of related work in the literature. In appendix A a Gibbs Sampler schedule for sampling from the posterior distribution is presented. Finally, in appendix B we present the proofs from the theorems from section 2.

2 References priors

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 $u(x_0) = (n+k)c_x + nk(x_0 - \overline{x})^2$.

Theorem 1 *Under the model given by (1) and (2), the Jeffreys prior is given by* $\pi_1(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) \ \propto \ u(x_0)^{(p-1)/2}\,(\sigma^2)^{-(2p+3)/2}\,(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}$.

Theorem 2 *Under the model given by (1) and (2), the two-group reference prior is given by* $\pi_2(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) \ \propto \ u(x_0)^{-1/2}\,(\sigma^2)^{-(p+1)}$.

Although the approach of more than two groups might produce different priors for different orderings of the nuisance parameters, Bernardo (1997) points 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 3 *Under the model given by (1) and (2), the four-group reference prior, associated with the six different orderings of nuisance parameters, doesn't depend on p and is given by* $\pi_4(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) \ \propto \ u(x_0)^{-1/2}\,(\sigma^2)^{-1}$.

The references priors given by Theorems 1, 2 and 3 are improper priors, as it is often in reference analysis, when the parameter space is unbounded. However the next result shows that the associated posterior distributions are proper.

Theorem 4 *The joint reference posterior distributions associated with reference priors* π_1, π_2 *and* π_4 *are proper.*

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

This 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 doesn't make distinction between interest and nuisance parameters. Berger and Bernardo (1992a) recommend the use of more than two-group approach based on their empirical experience. In particular 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 the next section we compare priors through an example.

3 An example

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 1, page 19, 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 $micromoles/cm^3$.

A cross-validation tecnique was considered to compare the different priors. Each time a value of x from Table 1 was considered unkown and it was estimated using the posterior distribution. On Table 1 columns 4, 5 and 6 present the posterior average and columns 7, 8 and 9 the equal-tail credibility interval with probability 0.95, respectively, for the four-group, two-group and Jeffreys priors.

As we can see from the table all the point estimate are very close and 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). These results agree with Ghosh et al (1995), for the case $p = 1$.

4 Discussion

In the statistical literature (see Yang and Berger 1997, Bernardo and Ramón 1998) the reference prior for the linear calibration problem is credited to Kubokawa and Robert (1994) that obtained the reference prior associated with the related problem of the estimation of the ratio of two normal means. In a first step, Kubokawa and Robert reduced the original model, given by 1 and 2), by sufficience 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 $\mathbf{w} \sim N_p(\boldsymbol{\beta}^*, \sigma^2\mathbf{I}_p)$, $\mathbf{z} \sim N_p(x_0^*\boldsymbol{\beta}^*, \sigma^2\mathbf{I}_p)$, $s \sim Gama\left(\frac{q}{2}, \frac{1}{2\sigma^2}\right)$. 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 $|\partial(x_0^*, \boldsymbol{\beta}^*, \sigma^2)/\partial(x_0, \boldsymbol{\beta}, \sigma^2)| = |\sqrt{nk/(n+k)}|$ we have that both Kubokawa and Robert's priors don't coincide with ours (see Table 2, page 20).

In spite of Kubokawa and Robert had worked with a model related to the linear calibration model, their reference priors don't 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 arouse 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 2, page 20). 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 $|\partial(x_0, \alpha_0, \beta, \sigma^2)/\partial(x_0, \alpha, \beta, \sigma^2)|$ and, as expected, the transformed reference prior coincides exactly with our priors.

5 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 revealed more conservative than the two-group and Jeffreys prior. It is in accordance with Ghosh *et al* results for $p = 1$.

Appendix A

Since it is not possible to find a closed form for the marginal posteriors, the inference used on section 3, was based on the Gibbs Sampler algorithm. It depends on the complete conditional posterior distributions. From the posterior distribution, given by

$$\pi_r \ \times \ (\sigma^{-2})^{(n+k)p/2} \exp\left\{ - \left[\sum_{i=1}^{n} (\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i) + \right.\right.$$

$$\left.\left. + \sum_{j=1}^{k} (\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta} x_0)'(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta} x_0) \right] \Big/ 2\sigma^2 \right\}, \qquad r = 1, 2 \text{ or } 4 \tag{3}$$

we can easily see that

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

$$\sigma^{-2} \mid x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{y}, \mathbf{X} \ \sim \ Gamma\left(\frac{(n+k+2)p+4}{2}, \frac{l_1}{2} \right) \qquad \text{for } r = 2, \tag{5}$$

$$\sigma^{-2} \mid x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{y}, \mathbf{X} \ \sim \ Gamma\left(\frac{(n+k+2)p+5}{2}, \frac{l_1}{2} \right) \qquad \text{for } r = 1, \tag{6}$$

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 a some algebra we have that

$$\boldsymbol{\alpha} \mid x_0, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X} \ \sim \ N_p\left(\boldsymbol{l}_2, \frac{\sigma^2}{n+k} \mathbf{I}_p \right) \qquad \text{for } r = 1, 2, 4, \tag{7}$$

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 some algebra we have that

$$\boldsymbol{\beta} \mid x_0, \boldsymbol{\alpha}, \sigma^2, \mathbf{y}, \mathbf{X} \ \sim \ N_p\left(\boldsymbol{l}_3, \frac{\sigma^2}{k x_0^2 + \sum_{i=1}^{n} x_i^2} \mathbf{I}_p \right) \qquad \text{for } r = 2 \text{ and } 4, \tag{8}$$

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

$$(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2} \ \times \ \exp\left\{ - \frac{(\boldsymbol{\beta} - \boldsymbol{l}_3)'(\boldsymbol{\beta} - \boldsymbol{l}_3)}{2\sigma^2/(k x_0^2 + \sum_{i=1}^{n} x_i^2)} \right\}.$$

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

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

$$u(x_0)^{-1/2} \times \exp\left\{ -\frac{(x_0 - l_4)^2}{2\sigma^2/k\boldsymbol{\beta}'\boldsymbol{\beta}} \right\}, \tag{9}$$

where $l_4 = \boldsymbol{\beta}' \sum_{j=1}^{k} (\mathbf{y}_{0j} - \boldsymbol{\alpha})/(k\boldsymbol{\beta}'\boldsymbol{\beta})$. Note that (9) is not a known distribution of probability. Following Kubokawa and Robert (1994) we introduce a latent variable t as follows

$$u(x_0)^{-1/2} \propto \Gamma(1/2)\, u(x_0)^{-1/2} = \int_0^\infty t^{\frac{1}{2}-1} \exp\{-u(x_0)\, t\}\, dt. \tag{10}$$

We now can rewrite (9) as

$$p(x_0 \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X}) = \int_0^\infty p(x_0, t \mid \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X})\, dt, \tag{11}$$

where

$$p(x_0, t \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X}) \propto t^{\frac{1}{2}-1} \exp\{-u(x_0)\, t\} \exp\left\{ -\frac{(x_0 - l_4)^2}{2\sigma^2/k\boldsymbol{\beta}'\boldsymbol{\beta}} \right\}. \tag{12}$$

From (12) we have that

$$t \mid x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X} \sim \mathit{Gamma}\left(\frac{1}{2},\, u(x_0) \right), \tag{13}$$

and

$$p(x_0 \mid t, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X}) \propto \exp\{-u(x_0)\, t\} \exp\left\{ -\frac{(x_0 - l_4)^2}{2\sigma^2/k\boldsymbol{\beta}'\boldsymbol{\beta}} \right\}. \tag{14}$$

Therefore,

$$x_0 \mid t, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \mathbf{y}, \mathbf{X} \sim N\left(\frac{2\sigma^2 nt\overline{x} + l_4\boldsymbol{\beta}'\boldsymbol{\beta}}{2\sigma^2 nt + \boldsymbol{\beta}'\boldsymbol{\beta}},\, \frac{\sigma^2}{2\sigma^2 nkt + k\boldsymbol{\beta}'\boldsymbol{\beta}} \right). \tag{15}$$

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

Appendix B

In this appendix we present the proofs for theorems 1 to 5. The Fisher information matrix associated with the model given by (1) and (2) is

$$\mathbf{H}(\boldsymbol{\theta}) = \left(\begin{array}{c|c|c|c} k\boldsymbol{\beta}'\boldsymbol{\beta}/\sigma^2 & k\boldsymbol{\beta}'/\sigma^2 & kx_0\boldsymbol{\beta}'/\sigma^2 & 0 \\ \hline k\boldsymbol{\beta}/\sigma^2 & (n+k)/\sigma^2\mathbf{I}_p & (kx_0 + \sum\limits_{i=1}^{n} x_i)/\sigma^2\mathbf{I}_p & \mathbf{0} \\ \hline kx_0\boldsymbol{\beta}/\sigma^2 & (kx_0 + \sum\limits_{i=1}^{n} x_i)/\sigma^2\mathbf{I}_p & (kx_0^2 + \sum\limits_{i=1}^{n} x_i^2)/\sigma^2\mathbf{I}_p & \mathbf{0} \\ \hline 0 & \mathbf{0} & \mathbf{0} & (n+k)p/2\sigma^4 \end{array}\right). \tag{16}$$

Proof of Theorem 1: First we calculate the determinant of (16), partitioning the Fisher information matrix as

$$\mathbf{H}(\boldsymbol{\theta}) = \left(\begin{array}{c|c} \mathbf{h}_{11} & \mathbf{h}_{12} \\ \hline \mathbf{h}_{21} & \mathbf{h}_{22} \end{array}\right), \tag{17}$$

where

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

$$\mathbf{h}_{22} = \left(\begin{array}{c|cc} (n+k)/\sigma^2\mathbf{I}_p & (kx_0 + \sum\limits_{i=1}^{n} x_i)/\sigma^2\mathbf{I}_p & \mathbf{0} \\ \hline (kx_0 + \sum\limits_{i=1}^{n} x_i)/\sigma^2\mathbf{I}_p & (kx_0^2 + \sum\limits_{i=1}^{n} x_i^2)/\sigma^2\mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (n+k)p/2\sigma^4 \end{array}\right) = \left(\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right).$$

We have that $|\mathbf{H}(\boldsymbol{\theta})| = |\mathbf{h}_{22}||\mathbf{h}_{11.2}|$, where $|\mathbf{h}_{11.2}| = \mathbf{h}_{11.2} = \mathbf{h}_{11} - \mathbf{h}_{12}\mathbf{h}_{22}^{-1}\mathbf{h}_{21}$.

Observe 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 some algebra we have that

$$|\mathbf{h}_{22}| = \frac{u(x_0)^p(n+k)p}{2(\sigma^2)^{2p+2}}. \tag{18}$$

Now to calculate $\mathbf{h}_{11.2}$ we first need to invert matrix $\mathbf{h}_{22}$. Observe that matrix $\mathbf{h}_{22}^{-1}$ can be easily obtained as

$$\mathbf{h}_{22}^{-1} = \left(\begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} \end{array}\right), \tag{19}$$

where

$$\mathbf{B}_{11} = \mathbf{A}_{11.2}^{-1}, \qquad \mathbf{B}_{12} = -\mathbf{A}_{11.2}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1},$$
$$\mathbf{B}_{21} = -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11.2}^{-1}, \qquad \mathbf{B}_{22} = \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11.2}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1},$$

where the matrix $\mathbf{A}_{22}$ and $\mathbf{A}_{11.2}$ are diagonal. Then, after some algebra we have that

$$\mathbf{h}_{22}^{-1} = \left(\begin{array}{c|c|c} \sigma^2(kx_0^2 + \sum_{i=1}^{n} x_i^2)/a\mathbf{I}_p & -\sigma^2(kx_0 + \sum_{i=1}^{n} x_i)/a\mathbf{I}_p & \mathbf{0} \\ \hline -\sigma^2(kx_0 + \sum_{i=1}^{n} x_i)/a\mathbf{I}_p & \sigma^2(n+k)/a\mathbf{I}_p & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & 2\sigma^4/(n+k)p \end{array}\right). \tag{20}$$

Therefore, we find that

$$\mathbf{h}_{11.2} = \frac{nc_x k\boldsymbol{\beta}'\boldsymbol{\beta}}{u(x_0)\sigma^2}. \tag{21}$$

The Jeffreys prior is then calculated as

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

Proof of Theorem 2: Partitioning the Fisher information matrix as in (17) and following the notation from the corollary of theorem 5.29 from Bernardo and Smith (1994) we have that

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

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 (21). In this way

$$|\mathbf{h}_{x_0}|^{1/2} = \left\{\frac{nc_x k\boldsymbol{\beta}'\boldsymbol{\beta}}{\sigma^2}\right\}^{1/2} u(x_0)^{-1/2} = f_1(x_0)\,g_1(\boldsymbol{\alpha},\boldsymbol{\beta},\sigma^2), \tag{23}$$

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 two-group reference prior is given by

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

Proof of Theorem 3: Here we are following the algorithm proposed by Berger and Bernardo (1992b). Firstly we have that $\mathbf{H}_4(\boldsymbol{\theta})$ is equal to the Fisher information matrix given by (16). Recall that $\mathbf{h}_j(\boldsymbol{\theta})$ is the right down submatrix of $\mathbf{H}_j(\boldsymbol{\theta})$ and then,

$$|\mathbf{h}_4(\boldsymbol{\theta})|^{1/2} = \sigma^{-2}\{(n+k)p/2\}^{1/2} = a_4(\sigma^2)b_4(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}). \tag{24}$$

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 (17) we have that

$$\mathbf{S}_4(\boldsymbol{\theta}) = \left(\begin{array}{c|c} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \hline \mathbf{C}_{21} & \mathbf{C}_{22} \end{array}\right), \tag{25}$$

where

$$\begin{aligned} &\mathbf{C}_{11} = \mathbf{h}_{11.2}^{-1}, && \mathbf{C}_{12} = -\mathbf{h}_{11.2}^{-1}\mathbf{h}_{12}\mathbf{h}_{22}^{-1}, \\ &\mathbf{C}_{21} = -\mathbf{h}_{22}^{-1}\mathbf{h}_{21}\mathbf{h}_{11.2}^{-1}, && \mathbf{C}_{22} = \mathbf{h}_{22}^{-1}\mathbf{h}_{21}\mathbf{h}_{11.2}^{-1}\mathbf{h}_{12}\mathbf{h}_{22}^{-1} + \mathbf{h}_{22}^{-1}, \\ &\mathbf{h}_{11.2} = \mathbf{h}_{x_0} = \mathbf{h}_{11} - \mathbf{h}_{12}\mathbf{h}_{22}^{-1}\mathbf{h}_{21}. \end{aligned}$$

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

$$\mathbf{S}_4(\boldsymbol{\theta}) = \left(\begin{array}{c|c|c|c} u(x_0)\sigma^2/nc_x k\boldsymbol{\beta}'\boldsymbol{\beta} & -\sigma^2\boldsymbol{\beta}' b/nc_x\boldsymbol{\beta}'\boldsymbol{\beta} & -\sigma^2\boldsymbol{\beta}' c/nc_x\boldsymbol{\beta}'\boldsymbol{\beta} & 0 \\ \hline -\sigma^2\boldsymbol{\beta}' b/nc_x\boldsymbol{\beta}'\boldsymbol{\beta} & f(x_0,\boldsymbol{\beta},\sigma^2) & g(x_0,\boldsymbol{\beta},\sigma^2) & \mathbf{0} \\ \hline -\sigma^2\boldsymbol{\beta}' c/nc_x\boldsymbol{\beta}'\boldsymbol{\beta} & g(x_0,\boldsymbol{\beta},\sigma^2) & h(x_0,\boldsymbol{\beta},\sigma^2) & \mathbf{0} \\ \hline 0 & \mathbf{0} & \mathbf{0} & 2\sigma^4/(n+k)p \end{array}\right), \tag{26}$$

where

$$\begin{aligned} b &= \sum_{i=1}^n x_i^2 - x_0\sum_{i=1}^n x_i, & c &= nx_0 - \sum_{i=1}^n x_i, \\ d &= \sum_{i=1}^n x_i^2, & e &= \sum_{i=1}^n x_i, \end{aligned}$$

$$f(x_0,\boldsymbol{\beta},\sigma^2) = \frac{\sigma^2 k\boldsymbol{\beta}\boldsymbol{\beta}' b^2}{u(x_0)nc_x\boldsymbol{\beta}'\boldsymbol{\beta}} + \frac{\sigma^2(kx_0^2+d)}{u(x_0)}\mathbf{I}_p,$$

$$g(x_0,\boldsymbol{\beta},\sigma^2) = \frac{\sigma^2 k\boldsymbol{\beta}\boldsymbol{\beta}' bc}{u(x_0)nc_x\boldsymbol{\beta}'\boldsymbol{\beta}} - \frac{\sigma^2(kx_0+e)}{u(x_0)}\mathbf{I}_p,$$

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

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

$$|\mathbf{h}_3(\boldsymbol{\theta})|^{1/2} = 1 \times \left\{ \frac{\{kx_0^2 + \sum_{i=1}^n x_i^2\}^p}{\sigma^{2p}} \right\}^{1/2} = a_3(\boldsymbol{\beta}) b_3(x_0, \boldsymbol{\alpha}, \sigma^2). \tag{27}$$

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 $\mathbf{S}_4(\boldsymbol{\theta})$) we find that

$$|\mathbf{h}_2(\boldsymbol{\theta})|^{1/2} = 1 \times |\mathbf{h}_2(\boldsymbol{\theta})|^{1/2} = a_2(\boldsymbol{\alpha}) b_2(x_0, \boldsymbol{\beta}, \sigma^2). \tag{28}$$

Observe that we don't 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, from

$$\mathbf{S}_1(\boldsymbol{\theta}) = \frac{a\sigma^2}{nc_x k \boldsymbol{\beta}'\boldsymbol{\beta}} \quad \Rightarrow \quad \mathbf{h}_1(\boldsymbol{\theta}) = \mathbf{H}_1(\boldsymbol{\theta}) = \mathbf{S}_1^{-1}(\boldsymbol{\theta}) = \frac{nc_x k \boldsymbol{\beta}'\boldsymbol{\beta}}{a\sigma^2}, \tag{29}$$

follows that

$$|\mathbf{h}_1(\boldsymbol{\theta})|^{1/2} = u(x_0)^{-1/2} \left\{ \frac{nc_x k \boldsymbol{\beta}'\boldsymbol{\beta}}{\sigma^2} \right\}^{1/2} = a_1(x_0) b_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2). \tag{30}$$

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

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

Proof of Theorem 4: We will first prove the result for the posterior distribution associated with π_2. The posterior distribution $\pi_2(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}, \mathbf{X})$ is proportional to

$$\begin{aligned} u(x_0)^{-1/2} (\sigma^{-2})^{\frac{(n+k+2)p+2}{2}} \exp \Bigg\{ - \Bigg[& \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i) \ + \\ & + \sum_{j=1}^k (\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta} x_0)'(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta} x_0) \Bigg] \Big/ 2\sigma^2 \Bigg\}. \end{aligned} \tag{31}$$

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

$$\pi_2(x_0, \boldsymbol{\alpha}, \boldsymbol{\beta} \mid \mathbf{y}, \mathbf{X}) \ \propto \ u(x_0)^{-1/2} \Bigg\{ \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta} x_i) \ +$$

$$+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\alpha}-\boldsymbol{\beta}x_0)'(\mathbf{y}_{0j}-\boldsymbol{\alpha}-\boldsymbol{\beta}x_0)\Bigg\}^{-\frac{(n+k+2)p+4}{2}}. \tag{32}$$

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

$$u(x_0)^{-1/2}\Bigg\{(n+k)\left[\sum_{i=1}^{n}(\mathbf{y}_i-\boldsymbol{\beta}x_i)'(\mathbf{y}_i-\boldsymbol{\beta}x_i)+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)'(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)\right]$$
$$-\left(\sum_{i=1}^{n}(\mathbf{y}_i-\boldsymbol{\beta}x_i)+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)\right)'\left(\sum_{i=1}^{n}(\mathbf{y}_i-\boldsymbol{\beta}x_i)+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)\right)\Bigg\}^{-\frac{(n+k+1)p+4}{2}}. \tag{33}$$

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

$$\pi_2(x_0\mid\mathbf{y},\mathbf{X})\;\propto\;u(x_0)^{\frac{(n+k-1)p+3}{2}}\left(1+\frac{1}{v}\frac{(x_0-l_7)^2}{l_8/(nkl_6-\boldsymbol{l}_5'\boldsymbol{l}_5)v}\right)^{-\frac{v+1}{2}}, \tag{34}$$

where

$$\begin{aligned}
v &= (n+k)p+3,\\
\boldsymbol{l}_4 &= (n+k)\sum_{i=1}^{n}\mathbf{y}_ix_i-\sum_{i=1}^{n}x_i\left(\sum_{i=1}^{n}\mathbf{y}_i+\sum_{j=1}^{k}\mathbf{y}_{0j}\right).\\
\boldsymbol{l}_5 &= n\sum_{j=1}^{k}\mathbf{y}_{0j}-k\sum_{i=1}^{n}\mathbf{y}_i,\\
l_6 &= (n+k)(\sum_{i=1}^{n}\mathbf{y}_i'\mathbf{y}_i+\sum_{j=1}^{k}\mathbf{y}_{0j}'\mathbf{y}_{0j})-\left(\sum_{i=1}^{n}y_i+\sum_{j=1}^{k}\mathbf{y}_{0j}\right)'\left(\sum_{i=1}^{n}y_i+\sum_{j=1}^{k}\mathbf{y}_{0j}\right),\\
l_7 &= \frac{nkl_6\overline{x}+\boldsymbol{l}_4'\boldsymbol{l}_5}{nkl_6-\boldsymbol{l}_5'\boldsymbol{l}_5},\\
l_8 &= nk\overline{x}^2l_6+(n+k)c_xl_6-\boldsymbol{l}_4'\boldsymbol{l}_4-(nkl_6-\boldsymbol{l}_5'\boldsymbol{l}_5)l_7^2.
\end{aligned}$$

Now we are going to show that the marginal posterior distribution of x_0 is proper. First observe that (34) is proportional to a function of x_0 multiplied by the kernel of t-Student distribution. In this way

$$\int_{-\infty}^{+\infty}\pi_2(x_0\mid\mathbf{y},\mathbf{X})\,dx_0\;\propto\;E_{x_0}\left[u(x_0)^{\frac{(n+k-1)p+3}{2}}\right], \tag{35}$$

where

$$x_0 \ \sim\ t_v\left(l_7\,,\ \frac{l_8}{(nkl_6 - \boldsymbol{l}_5'\boldsymbol{l}_5)v}\right). \tag{36}$$

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 easily show that (35) is finite. Observe 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 polinomial of degree $(n+k-1)p+3$. Since

$$E\left[\sum_{i=0}^{(n+k-1)p+3} a_i x_0^i\right] = \sum_{i=0}^{(n+k-1)p+3} a_i E\left[x_0^i\right],$$

and from (36), the number of degrees of freedom $v = (n+k)p+3$ is strictly higher than $(n+k-1)p+3$, we have that (35) can be rewritten as a linear combination of finite moments, so that

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

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. $\triangle$

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

$$(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2} u(x_0)^{\frac{p-1}{2}} (\sigma^{-2})^{\frac{(n+k+2)p+3}{2}} \exp\left\{-\left[\sum_{i=1}^{n}(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i)'(\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i) + \right.\right.$$
$$\left.\left. + \sum_{j=1}^{k}(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)'(\mathbf{y}_{0j} - \boldsymbol{\alpha} - \boldsymbol{\beta}x_0)\right] \Big/ 2\sigma^2\right\}. \tag{37}$$

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

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

$$+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\alpha}-\boldsymbol{\beta}x_0)'(\mathbf{y}_{0j}-\boldsymbol{\alpha}-\boldsymbol{\beta}x_0)\Bigg\}^{-\frac{(n+k+2)p+5}{2}}. \qquad (38)$$

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

$$\pi_1(x_0,\boldsymbol{\beta}\mid\mathbf{y},\mathbf{X}) \;\propto\; (\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}u(x_0)^{\frac{p-1}{2}}\Bigg\{(n+k)\Bigg[\sum_{i=1}^{n}(\mathbf{y}_i-\boldsymbol{\beta}x_i)'(\mathbf{y}_i-\boldsymbol{\beta}x_i)+$$
$$+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)'(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)\Bigg]-\left(\sum_{i=1}^{n}(\mathbf{y}_i-\boldsymbol{\beta}x_i)+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)\right)'$$
$$\times\left(\sum_{i=1}^{n}(\mathbf{y}_i-\boldsymbol{\beta}x_i)+\sum_{j=1}^{k}(\mathbf{y}_{0j}-\boldsymbol{\beta}x_0)\right)\Bigg\}^{-\frac{(n+k+1)p+5}{2}}. \qquad (39)$$

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

$$\pi_1(x_0\mid\mathbf{y},\mathbf{X}) \;\propto\; E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}\right]u(x_0)^{\frac{(n+k)p+4}{2}}\left(1+\frac{1}{v}\frac{(x_0-l_7)^2}{l_8/(nkl_6-\boldsymbol{l}_5'\boldsymbol{l}_5)v_1}\right)^{-\frac{v+1}{2}}, \qquad (40)$$

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

$$E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}\right]=\int_{\boldsymbol{\beta}} c'\;(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}\left(1+\frac{1}{v_2}\frac{(\boldsymbol{\beta}-\mathbf{m})'(\boldsymbol{\beta}-\mathbf{m})}{s/v_2}\right)^{-\frac{v_2+p}{2}} d\boldsymbol{\beta}, \qquad (41)$$

$$c'=\frac{\Gamma(\frac{v_2+p}{2})}{\Gamma(\frac{v_2}{2})\,v_2^{p/2}\;\pi^{p/2}\,|s\mathbf{I}_p|^{1/2}} \qquad\qquad \mathbf{m}=\frac{\boldsymbol{l}_4+\boldsymbol{l}_5x_0}{a},$$

$$s=\frac{a\,l_6-(\boldsymbol{l}_4+\boldsymbol{l}_5x_0)'(\boldsymbol{l}_4+\boldsymbol{l}_5x_0)}{a^2v_2}, \qquad\qquad v_2=(n+k)p+5.$$

Observe that (41) is a function of x_0 and that

$$\int_{-\infty}^{+\infty}\pi_1(x_0\mid\mathbf{y},\mathbf{X})\,dx_0 \;\propto\; E_{x_0}\left[E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}\right]u(x_0)^{\frac{(n+k)p+4}{2}}\right], \qquad (42)$$

where

$$x_0 \;\sim\; t_v\left(l_7\,,\;\frac{l_8}{v_1(nkl_6-\boldsymbol{l}_5'\boldsymbol{l}_5)}\right). \qquad (43)$$

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

$$E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})^{1/2}\right] \;\leq\; E\left[(\boldsymbol{\beta}'\boldsymbol{\beta})\right]^{1/2}. \tag{44}$$

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

$$E\left[\mathbf{x}'\mathbf{A}\mathbf{x}\right] = tr(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}, \tag{45}$$

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

$$\begin{aligned} E\left[\boldsymbol{\beta}'\boldsymbol{\beta}\right] &= tr(s\mathbf{I}_p) + \mathbf{m}'\mathbf{m} = \\ &= \frac{\{(n+k)c_x + nk(x_0-\overline{x})^2\}l_6 - m_3(x_0+m_1)^2 - m_2}{\{(n+k)c_x + nk(x_0-\overline{x})^2\}^2 v}p+ \\ &\quad + \frac{m_3(x_0+m_1)^2 + m_2}{\{(n+k)c_x + nk(x_0-\overline{x})^2\}^2} \end{aligned} \tag{46}$$

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. So, we have that

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

and so

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

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.

And as a consequence from the Jensen inequality

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

so,

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

Proof of Theorem 5: We will first show the result for the marginal posterior distribution associated with the two-group reference prior.

To prove this result we just need to show that

$$E[x_0^p \mid \mathbf{y}, \mathbf{X}] = \int_{-\infty}^{+\infty} x_0^p \pi_2(x_0 \mid \mathbf{y}, \mathbf{X})\, dx_0 = \infty, \tag{47}$$

and that

$$E[x_0^{p-1} \mid \mathbf{y}, \mathbf{X}] = \int_{-\infty}^{+\infty} x_0^{p-1} \pi_2(x_0 \mid \mathbf{y}, \mathbf{X})\, dx_0 < \infty. \tag{48}$$

Observing that

$$E[x_0^p \mid \mathbf{y}, \mathbf{X}] \ \propto \ E_{x_0}\left[x_0^p\left[(n+k)c_x + nk(x_0-\overline{x})^2\right]^{\frac{(n+k-1)p+3}{2}}\right], \tag{49}$$

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

$$x_0 \ \sim \ t_v\left(l_7\,,\ \frac{l_8}{v(nkl_6 - \boldsymbol{l}_5'\boldsymbol{l}_5)}\right), \tag{50}$$

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 (50). Nonetheless, as (50) has $v = (n+k)p+3$, we have that the moment of order $(n+k)p+3$ diverges, and so $E[x_0^p \mid \mathbf{y}, \mathbf{X}] = \infty$. Recalling that, when the moment of order i doesn't exists all other moments of order superior to i also don't exist, we have that $\pi_2(x_0 \mid \mathbf{y}, \mathbf{X})$ doesn't have any moment of order superior or equal to p.

Now, observing that

$$E[x_0^{p-1} \mid \mathbf{y}, \mathbf{X}] \;\propto\; E_{x_0}\left[x_0^{p-1} \left[(n+k)c_x + nk(x_0 - \overline{x})^2 \right]^{\frac{(n+k-1)p+3}{2}} \right], \tag{51}$$

and that

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

represents a polinomial 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 the distribution 50) of order less or equal to $(n+k)p+2$. Nonetheless, since (50) has $v = (n+k)p+3$, we have that all moments of order less or equal to $(n+k)p+2$ are finite, and so $E[x_0^{p-1} \mid \mathbf{y}, \mathbf{X}] < \infty$. In this way, we proved $\pi(x_0 \mid \mathbf{y}, \mathbf{X})$ is proper, but has finite moments of order less or equal to $p-1$. $\triangle$

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

References

[1] Berger, J. O. and Bernardo, J. M. (1989). Estimating a product of means: Bayesian analysis with reference priors. *J. Amer. Statist. Assoc.* **84**, 200-207.

[2] 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).

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

[4] Bernardo, J. M. (1979). Reference posterior distributions for bayesian inference. *J. R. Statist. Soc. B* **41**, 113-147 (with discussion).

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

[6] Bernardo, J. M. (1997). Non-informative priors do not exist. *J. Statist. Planning Inference* **65**, 113-147 (with discussion).

[7] 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.

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

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

[10] Ghosh, M. Carlin, B. P. and Srivastava, M. S. (1995). Probability matching priors for linear calibration. *Test* **4**, 2, 333-357.

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

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

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

[14] Kubokawa, T. and Robert, C. (1994). New perspectives on linear calibration. *J. Multivariate Analysis* **50**, 178-200.

[15] Liseo, B. (1993). Elimination of nuisance parameters with reference priors. *Biometrika* **80**, 295-304.

[16] Osborne, C. (1991). Statistical calibration: a review. *Internat. Statist. Review* **59**, 309-336.

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

Table 1: Posterior average and credibility intervals (0.95)

x_0	y_1	y_2	$\hat{x}_0^4$	$\hat{x}_0^2$	$\hat{x}_0^J$	I^4	I^2	I^J
0.0	1.6	1.5	-0.66	-0.66	-0.65	$(-1.45, 0.11)$	$(-1.41, 0.09)$	$(-1.39, 0.07)$
0.5	1.8	1.9	0.01	0.01	0.01	$(-0.76, 0.77)$	$(-0.75, 0.74)$	$(-0.73, 0.75)$
1.0	2.0	2.2	0.54	0.54	0.55	$(-0.24, 1.30)$	$(-0.19, 1.28)$	$(-0.21, 1.28)$
1.5	2.2	2.6	1.22	1.21	1.21	(0.44, 1.99)	(0.46, 1.94)	(0.47, 1.93)
2.0	2.4	2.9	1.74	1.73	1.73	(0.97, 2.49)	(1.01, 2.46)	(1.01, 2.46)
2.5	2.6	3.2	2.26	2.26	2.27	(1.50, 3.00)	(1.52, 3.00)	(1.54, 3.02)
3.0	2.8	3.6	2.93	2.92	2.92	(2.17, 3.70)	(2.18, 3.67)	(2.18, 3.66)
3.5	3.0	3.9	3.45	3.45	3.45	(2.67, 4.20)	(2.71, 4.17)	(2.70, 4.18)
4.0	3.2	4.2	3.98	3.97	3.97	(3.21, 4.73)	(3.24, 4.69)	(3.24, 4.70)
4.5	3.4	4.5	4.50	4.50	4.50	(3.73, 5.24)	(3.76, 5.24)	(3.77, 5.23)
5.0	3.6	4.8	5.02	5.02	5.02	(4.26, 5.78)	(4.26, 5.75)	(4.31, 5.72)
5.5	3.8	5.2	5.68	5.67	5.68	(4.91, 6.42)	(4.94, 6.40)	(4.96, 6.38)
6.0	3.9	5.5	6.13	6.13	6.14	(5.39, 6.88)	(5.41, 6.86)	(5.40, 6.85)
6.5	4.1	5.8	6.65	6.65	6.65	(5.91, 7.40)	(5.92, 7.39)	(5.94, 7.38)
7.0	4.3	6.1	7.18	7.18	7.18	(6.42, 7.92)	(6.45, 7.91)	(6.45, 7.90)
7.5	4.5	6.4	7.70	7.70	7.70	(6.96, 8.44)	(6.97, 8.42)	(6.99, 8.43)
8.0	4.6	6.7	8.15	8.15	8.15	(7.40, 8.91)	(7.44, 8.88)	(7.42, 8.87)
8.5	4.8	7.0	8.67	8.67	8.67	(7.92, 9.41)	(7.96, 9.40)	(7.97, 9.41)
9.0	5.0	7.3	9.19	9.20	9.20	(8.45, 9.95)	(8.47, 9.94)	(8.49, 9.91)
9.5	5.1	7.6	9.65	9.64	9.65	(8.91, 10.41)	(8.91, 10.37)	(8.93, 10.37)
10.0	5.3	7.9	10.17	10.17	10.17	(9.44, 10.93)	(9.43, 10.89)	(9.44, 10.90)
11.0	5.6	8.5	11.15	11.14	11.15	(10.41, 11.90)	(10.41, 11.88)	(10.43, 11.87)
12.0	6.0	9.1	12.20	12.19	12.20	(11.45, 12.94)	(11.47, 12.94)	(11.49, 12.92)
13.0	6.3	9.7	13.17	13.17	13.17	(12.42, 13.93)	(12.44, 13.90)	(12.44, 13.91)
14.0	6.6	11.0	15.11	15.11	15.11	(14.39, 15.82)	(14.42, 15.81)	(14.44, 15.79)
15.0	6.9	11.4	15.82	15.81	15.82	(15.07, 16.55)	(15.11, 16.52)	(15.10, 16.52)
16.0	7.2	11.6	16.24	16.24	16.24	(15.47, 17.00)	(15.50, 16.99)	(15.53, 16.99)
17.0	7.5	12.0	16.93	16.93	16.93	(16.17, 17.71)	(16.17, 17.71)	(16.18, 17.67)
18.0	7.7	13.0	18.41	18.41	18.40	(17.62, 19.21)	(17.67, 19.18)	(17.67, 19.15)
20.0	8.2	14.0	19.98	19.98	19.98	(19.19, 20.80)	(19.22, 20.75)	(19.22, 20.75)
24.0	9.1	15.0	21.33	21.32	21.33	(20.90, 21.76)	(20.92, 21.73)	(20.92, 21.74)

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

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