BAYESIAN ANALYSIS OF THE CALIBRATION PROBLEM UNDER ELLIPTICAL DISTRIBUTIONS

Branco, M., Bolfarine, H.
Departamento de Estatística, Universidade de São Paulo
Caixa Postal 66281, CEP 05315-970, São Paulo, Brasil

Iglesias, P., Arellano-Valle, R. B.
Departamento de Estadística, Pontificia Universidad Católica de Chile
Casilla 306, Santiago 22, Chile

Abstract

In this paper we discuss calibration problems under the dependent and independent elliptical family of distributions. In the dependent case, it is shown that the posterior distribution of the quantity of interest is robust with respect to the distributions in the elliptical family. In particular, the results obtained by Hoadley (1970) showing that the inverse estimator is a Bayes estimator under normal models with a Student-t prior also holds under the dependent elliptical family of distributions. In the independent case, the use of the elliptical family allows the consideration of models which provide protection against possible outliers in the data. The multivariate calibration problem is also considered, where some results given in Brown (1993) are extended. Finally, the results of the paper are applied to a real data problem, showing that the Student-t model can be a valid alternative to normality.

Some key words: Calibration; Bayesian inference; Elliptical distributions ; Gibbs sampler.

AMS Subject Classifications: 62F15, 62F35, 62J05.

1 Introduction

Bayesian analysis of regression problems under elliptical distributions has received increasingly attention in the recent statistical literature. Essentially, in the case of the dependent model, it is shown that inference based on the posterior distribution using the usual improper reference prior distribution for the scale parameter coincide with those obtained under the normal model. A pioneer work in this area is the paper by Zellner (1976), in which a study based on the multivariate Student-t distribution is considered. Extensions of those results are considered in Osiewalski and Steel (1993). The independent Student-t regression model is treated in Geweke (1993) by using a Bayesian perspective. A problem of great practical interest associated with the regression model is the calibration problem. Most of the studies concerning the Bayesian approach to calibration problems are restricted to normal models (Hoadley, 1970, Hunter and Lamboy, 1981, among others). A review of the main results concerning calibration problems are presented in Brown (1993).

In this paper, we consider the linear calibration model specified by the equations

$$y_i = \beta_1 + \beta_2 x_i + \epsilon_i, \tag{1.1}$$

$i = 1, \ldots, n$, and

$$y_0 = \beta_1 + \beta_2 x_0 + \epsilon_0, \tag{1.2}$$

where $\epsilon_0, \epsilon_1, \ldots, \epsilon_n$ are uncorrelated random errors with $E[\epsilon_i] = 0$ and $Var[\epsilon_i] = \sigma^2$, $i = 1, \ldots, n$. It is assumed that $x_1, \ldots, x_n$ are known constants and that β_1, β_2, σ^2 and x_0 are unknown quantities (parameters). The main object is to make inference on x_0 (the calibration parameter) based on $y_0, y_1, \ldots, y_n$ and $x_1, \ldots, x_n$, which are observed. Bayesian solutions to the calibration problem under the dependent and independent elliptical models are developed. Here we use the notation $\boldsymbol{z} \sim El_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ to indicate that $\boldsymbol{z}$ is a p-dimensional random vector elliptically distributed with location vector $\boldsymbol{\mu} \in \Re^p$ and a $p \times p$ (positive definite) dispersion matrix $\boldsymbol{\Sigma}$, with density given by

$$p(\boldsymbol{z}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}|^{-1/2} g[(\mathbf{z} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{z} - \boldsymbol{\mu})],$$

for some density generator function $g(u)$, $u \geq 0$, that is,

$$\int_0^\infty u^{\frac{p}{2}-1} g(u) du = \frac{\Gamma(\frac{p}{2})}{\pi^{\frac{p}{2}}}, \tag{1.3}$$

which implies that $g(\boldsymbol{x}'\boldsymbol{x}), \boldsymbol{x} \in R^p$, is a p-dimensional spherical density. (A comprehensive review of the properties and characterizations of elliptical distributions can be found in

Fang et al., 1990). Thus, in (1.1)-(1.2) we have the dependent elliptical model by assuming that $\boldsymbol{\epsilon}_* = (\epsilon_0, \epsilon_1, \ldots, \epsilon_n)' \sim El_{n+1}(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n+1}; g)$, where $\boldsymbol{I}_{n+1}$ denotes the (n+1)-dimensional identity matrix, so that,

$$(1.4) \qquad p(\epsilon_0, \epsilon_1, \ldots, \epsilon_n | \sigma^2) = \Big(\frac{1}{\sigma^2}\Big)^{\frac{n+1}{2}} g\Big(\frac{1}{\sigma^2}\sum_{i=0}^{n} \epsilon_i^2\Big).$$

Moreover, (1.4) implies that (provided they exist) $E[\boldsymbol{\epsilon}_*] = \mathbf{0}$ and $Var[\boldsymbol{\epsilon}_*] = a_g \sigma^2 \boldsymbol{I}$, where a_g denotes an expectation associated with a positive random variable having density function $\{\pi^{p/2}/\Gamma(p/2)\}u^{p/2-1}g(u), u \geq 0$, with $p = n+1$. In the independent case, we assume in (1.1)-(1.2) that $\epsilon_i \overset{iid}{\sim} El_1(0, \sigma^2; g)$, $i = 0, 1, \ldots, n$, with g being now an univariate elliptical density generator, that is, g is such that (1.3) holds for $p = 1$. Thus, for this model $\epsilon_0, \epsilon_1, \ldots, \epsilon_n$ have a joint density given by

$$(1.5) \qquad p(\epsilon_0, \epsilon_1, \ldots, \epsilon_n | \sigma^2) = \Big(\frac{1}{\sigma^2}\Big)^{\frac{n+1}{2}} \prod_{i=0}^{n} g\Big(\frac{\epsilon_i^2}{\sigma^2}\Big),$$

so that (provided they exist) $E[\epsilon_i] = 0$ and $Var[\epsilon_i] = a_g \sigma^2$, for $i = 0, 1, \ldots, n$.

Here we consider that the density generator g is known in both cases, namely, the dependent and independent elliptical models. In Section 2, we discuss Bayesian solutions to the linear calibration problem under dependent elliptical models. We show that inference does not depend on the density generator function when the prior distribution of the scale parameter is the usual improper reference prior distribution. Solutions for the independent elliptical model are considered in Section 3. Finally, we present an analysis of a real data problem in Section 4.

2 A Bayesian solution to the dependent case

In this section we present Bayesian solutions to the linear calibration problem under the dependent elliptical model. For simplicity the following notation is used in the sequel. Let $\boldsymbol{y} = (y_1, \ldots, y_n)'$ denote the response vector corresponding to the first step of the calibration experiment and $\boldsymbol{y}_* = (y_0, \boldsymbol{y}')'$. Similarly, we can define the vectors $\boldsymbol{x} = (x_1, \ldots, x_n)'$, $\boldsymbol{x}_* = (x_0, \boldsymbol{x}')'$ and the matrices

$$\boldsymbol{X} = [\mathbf{1}_n \quad \boldsymbol{x}] \quad \text{and} \quad \boldsymbol{X}_* = [\mathbf{1}_{n+1} \quad \boldsymbol{x}_*],$$

where $\mathbf{1}_k$ denotes the k-dimensional vector of ones. Assuming that $\boldsymbol{\epsilon}_* \sim El_{n+1}(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n+1}; g)$, it follows from (1.1) and (1.2) that $\boldsymbol{y}_* | \boldsymbol{X}_*, \boldsymbol{\beta}, \sigma^2 \sim El_{n+1}(\boldsymbol{X}_*\boldsymbol{\beta}, \sigma^2 \boldsymbol{I}_{n+1}; g)$, where $\boldsymbol{\beta} = (\beta_1, \beta_2)'$. Moreover, from (1.4) and from the fact that

$$(\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{\beta})'(\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{\beta}) = (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_*)'\boldsymbol{X}_*'\boldsymbol{X}_*(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_*) + (\boldsymbol{y}_* - \boldsymbol{X}_*\hat{\boldsymbol{\beta}}_*)'(\boldsymbol{y}_* - \boldsymbol{X}_*\hat{\boldsymbol{\beta}}_*),$$

where $\hat{\boldsymbol{\beta}}_* = (\boldsymbol{X}'_*\boldsymbol{X}_*)^{-1}\boldsymbol{X}'_*\boldsymbol{y}_*$, it follows that the likelihood function of the dependent case can be written as

$$p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{\beta}, \sigma^2) = \Big(\frac{1}{\sigma^2}\Big)^{\frac{n+1}{2}} g\Big(\Big\{1 + \frac{Q_*(\boldsymbol{\beta})}{n-1}\Big\}\Big\{\frac{(n-1)S_*^2}{\sigma^2}\Big\}\Big), \tag{2.1}$$

where

$$Q_*(\beta) = (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_*)'\boldsymbol{X}'_*\mathbf{X}_*(\beta - \hat{\boldsymbol{\beta}}_*)/S_*^2 \quad \text{and} \quad S_*^2 = (\boldsymbol{y}_* - \mathbf{X}_*\hat{\boldsymbol{\beta}}_*)'(\boldsymbol{y}_* - \boldsymbol{X}_*\hat{\boldsymbol{\beta}}_*)/(n-1). \tag{2.2}$$

In the following theorem we show that if $(\boldsymbol{\beta}, \sigma^2, x_0)$ has a prior distribution of the form

$$p(x_0, \boldsymbol{\beta}, \sigma^2|\boldsymbol{X}) \propto \frac{1}{\sigma^2} p(x_0, \boldsymbol{\beta}|\boldsymbol{X}), \tag{2.3}$$

then the posterior distribution of x_0 given $\boldsymbol{y}_*$ and $\boldsymbol{X}$ does not depend on g, whatever the prior distribution $p(x_0, \boldsymbol{\beta}|\boldsymbol{X}) = p(\boldsymbol{\beta}|\boldsymbol{X}_*)p(x_0|\boldsymbol{X})$ assigned to $(x_0, \boldsymbol{\beta})$. The proof follows from the fact that

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(\boldsymbol{y}_*|\boldsymbol{X}_*)p(x_0|\boldsymbol{X}), \tag{2.4}$$

where

$$p(\boldsymbol{y}_*|\boldsymbol{X}_*) = \int_{\boldsymbol{\beta}} \int_0^\infty p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{\beta}, \sigma^2) p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}_*) d\sigma^2 \mathbf{d}\boldsymbol{\beta},$$

which is independent of g (see Kelker, 1970 and Osiewalski and Steel, 1993).

Theorem 2.1 *Under the assumptions (2.1) and (2.3), the posterior density of x_0 is given by*

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(x_0|\boldsymbol{X}) \int_{\boldsymbol{\beta}} \Big\{\frac{1}{S_*^2}\Big\}^{\frac{n+1}{2}} \Big\{1 + \frac{Q_*(\boldsymbol{\beta})}{n-1}\Big\}^{-\frac{n+1}{2}} p(\beta|\boldsymbol{X}_*)\mathbf{d}\boldsymbol{\beta},$$

for any elliptical density generator g, where $Q_*(\boldsymbol{\beta})$ and S_*^2 are as in (2.2).

The above theorem shows that, under the prior distribution (2.3), Bayesian inference on x_0 is independent of the particular elliptical distribution under consideration. This includes the fact that the results derived for the normal model can be used for any elliptical model. Following next, the posterior distribution of x_0 for a special and important case of the elliptical family is considered as a special case of Theorem 2.1. The result is directly related to the results obtained by Hoadley (1970). Let $\hat{x}_{0,I}$ and $\hat{x}_0$ be the inverse and classical estimators of x_0, respectively. As is well known, $\hat{x}_{0,I}$ can be written as (see (3.7) in Hoadley, 1970)

$$\hat{x}_{0,I} = \Big(\frac{F}{n-2+F}\Big)\hat{x}_0 + \Big(1 - \frac{F}{n-2+F}\Big)\bar{x},$$

where

$$F = nS_x^2\hat{\beta}_2^2/S^2 \quad \text{and} \quad \hat{x}_0 = (y_0 - \hat{\beta}_1)/\hat{\beta}_2,$$

with

$$S_x^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2/n, \quad \bar{x} = \sum_{i=1}^{n} x_i/n,$$

$$S^2 = (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})/(n-2) \quad \text{and} \quad \hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)' = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}.$$

We call attention to the fact that, as in the normal model, under the dependent elliptical model (2.1), $\hat{\boldsymbol{\beta}}$ and $\hat{x}_0$ are the likelihood estimators of $\boldsymbol{\beta}$ and x_0, respectively.

Corollary 2.1 *If in (2.3), $p(x_0, \boldsymbol{\beta}|\mathbf{X}) = p(\boldsymbol{\beta}|\boldsymbol{X}_*)p(x_0|\boldsymbol{X})$ is such that $p(\boldsymbol{\beta}|\boldsymbol{X}_*) \propto$ constant and $x_0|\boldsymbol{X} \sim t(\bar{x}, (n+1)S_x^2/(n-3); n-3)$, $n > 3$, then*

$$x_0|\boldsymbol{X}, \boldsymbol{y}_* \sim t\Big(\hat{x}_{0,I}, \frac{n+1}{n-2+F} + \frac{\hat{x}_{0,I}^2}{F}; n-2\Big). \tag{2.5}$$

where $t(\mu, \sigma; \nu)$ denotes the Student-t distribution with location μ, dispersion σ and ν degrees of freedom.

As a consequence of Corollary 2.1, it follows that the inverse estimator is a robust Bayes estimator in the family of the elliptical distributions. Moreover, Theorem 2.1 also illustrates the fact that this result depends on the prior distribution and not on the likelihood function (2.1), which is a member of the elliptical family. The prior specifications for x_0 indicates symmetry about $\bar{x}$, and the posterior density becomes more concentrated as n increases. This is a consequence of the fact that $Var[x_0|\boldsymbol{X}] = (n+1)S_x^2/(n-3)$, $n > 3$. It is also true that the results obtained by Hunter and Lamboy (1981) and Kubokawa and Robert (1994) remain valid under the assumptions of Theorem 2.1.

Remark 2.1 *Brown (1982, 1993) obtain similar results for the multivariate calibration problem under the matrix normal distribution. These results can be also extended to the class of (left) matrix elliptical distribution, namely, $\boldsymbol{Z} \sim El_{m\times q}(\mathbf{L}, \boldsymbol{D}; g)$, with density given by (see Fang and Zhang, 1990)*

$$p(\boldsymbol{Z}|\boldsymbol{M}, \boldsymbol{D}) = |\boldsymbol{D}|^{-\frac{m}{2}} g\left(\boldsymbol{D}^{-\frac{1}{2}}(\boldsymbol{Z} - \mathbf{L})'(\boldsymbol{Z} - \mathbf{L})\boldsymbol{D}^{-\frac{1}{2}}\right),$$

where $\mathbf{L}$ and $\boldsymbol{D}$ are location and dispersion matrices with dimension $m \times q$ and $q \times q$, respectively, and g is a generator density function satisfying

$$\int |\boldsymbol{W}|^{\frac{m-q-1}{2}} g(\boldsymbol{W}) d\boldsymbol{W} = \frac{\Gamma_q(\frac{m}{2})}{\pi^{\frac{mq}{2}}},$$

where $\boldsymbol{W}$ *is a* $q \times q$ *matrix and* $\Gamma_q(.)$ *is the q-variate gamma function, that is,* $\Gamma_q(m/2) = \pi^{q(1-q)/4} \prod_{i=1}^{q} \Gamma[(m-i+1)/2]$. *In fact, let consider the linear calibration model:* $\boldsymbol{y}_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \boldsymbol{e}_i$, $i = 0, 1, .., n$, *where* $\boldsymbol{y}_i$ *and* $\boldsymbol{e}_i$ *are* $q \times 1$ *random vectors,* $\boldsymbol{x}_i$, $i = 1, \ldots, n$, *are* $p \times 1$ *known vectors,* $\boldsymbol{x}_0$ *is a* $p \times 1$ *unknown vector and* $\boldsymbol{\beta}$ *is a* $p \times q$ *unknown matrix. In matrix notation, this model can be written as* $\boldsymbol{Y}_* = \boldsymbol{X}_*\boldsymbol{\beta} + \boldsymbol{E}_*$, *where*

$$\boldsymbol{Y}_* = \begin{pmatrix} \boldsymbol{y}_0 \\ \boldsymbol{Y} \end{pmatrix}, \quad \boldsymbol{E}_* = \begin{pmatrix} \boldsymbol{e}_0 \\ \boldsymbol{E} \end{pmatrix} \quad \text{and} \quad \boldsymbol{X}_* = \begin{pmatrix} \boldsymbol{x}_0 \\ \boldsymbol{X} \end{pmatrix},$$

with $\boldsymbol{Y} = [\boldsymbol{y}_1', ..., \boldsymbol{y}_n']'$ *and* $\boldsymbol{E} = [\boldsymbol{e}_1', ..., \boldsymbol{e}_n']'$ *being* $n \times q$ *matrices and* $\boldsymbol{X} = [\boldsymbol{x}_1', ..., \boldsymbol{x}_n']'$ *being a* $n \times p$ *matrix. Thus, assuming that* $\boldsymbol{E}_* \sim El_{(n+1)\times q}(\mathbf{0}, \boldsymbol{\Sigma}; g)$, *so that (see Fang and Zhang, 1990)* $\boldsymbol{Y}_*|\boldsymbol{X}_*, \boldsymbol{\beta}, \boldsymbol{\Sigma} \sim EL_{(n+1)\times q}(\boldsymbol{X}_*\boldsymbol{\beta}, \boldsymbol{\Sigma}; g)$ *and has density function given by*

$$p(\boldsymbol{Y}_*|\boldsymbol{X}_*, \boldsymbol{\beta}, \boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}|^{-\frac{n+1}{2}} g(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{Y}_* - \boldsymbol{X}_*\boldsymbol{\beta})'(\boldsymbol{Y}_* - \boldsymbol{X}_*\boldsymbol{\beta})\boldsymbol{\Sigma}^{-\frac{1}{2}}),$$

and considering the prior distribution

$$p(\boldsymbol{x}_0, \boldsymbol{\beta}, \boldsymbol{\Sigma}|\boldsymbol{X}) \propto |\boldsymbol{\Sigma}|^{-\frac{q+1}{2}} p(\boldsymbol{x}_0, \boldsymbol{\beta}|\boldsymbol{X}),$$

we have that the posterior distribution $p(\boldsymbol{x}_0|\boldsymbol{Y}_*, \boldsymbol{X})$ *is independent of the generator density function* g *and is such that*

$$p(\boldsymbol{x}_0|\boldsymbol{Y}_*, \boldsymbol{X}) \propto p(\boldsymbol{x}_0|\boldsymbol{X}) \int_{\boldsymbol{\beta}} |(n-q)\mathbf{S}_* + (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_*)'\boldsymbol{X}_*'\boldsymbol{X}_*(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_*)|^{-\frac{n+q}{2}} p(\boldsymbol{\beta}|\boldsymbol{X}_*)\mathbf{d}\boldsymbol{\beta},$$

where

$$\hat{\boldsymbol{\beta}}_* = (\boldsymbol{X}_*'\boldsymbol{X}_*)^{-1}\boldsymbol{X}_*'\boldsymbol{Y}_* \textit{ and } \mathbf{S}_* = (\boldsymbol{Y}_* - \boldsymbol{X}_*\hat{\boldsymbol{\beta}}_*)'(\boldsymbol{Y}_* - \boldsymbol{X}_*\hat{\boldsymbol{\beta}}_*)/(n-q), \quad n > q.$$

This result is a multivariate extension of Theorem 2.1 and their proof follows by considering the transformation $\boldsymbol{W} = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{Y}_* - \boldsymbol{X}_*\boldsymbol{\beta})'(\boldsymbol{Y}_* - \boldsymbol{X}_*\boldsymbol{\beta})\boldsymbol{\Sigma}^{-\frac{1}{2}}$ *and using properties of matrix derivatives (see Fang and Zhang, 1990). Brown (1982), consider the above prior distribution with* $p(\boldsymbol{x}_0, \boldsymbol{\beta}|\boldsymbol{X}) = p(\boldsymbol{x}_0|\boldsymbol{X})p(\boldsymbol{\beta}|\boldsymbol{X})$. *If* $p(\boldsymbol{x}_0|\boldsymbol{X})$ *is taken as the Student-t multivariate density and* $p(\boldsymbol{\beta}|\boldsymbol{X})$ *is taken to be noninformative, then the above result proves that the results given in Brown (1982) (see also Brown, 1993) holds also for the left elliptical matrix distribution.*

The result presented in Theorem 2.1 depends on the choice of an improper prior distribution for the error variance. In the following we investigate the possibility of considering only proper conditional conjugate type prior distributions. Similarly to the normal model, we obtain conditional conjugation type results for the elliptical model. In Theorem 2.2 we present the conditional conjugate type elliptical family extended to a general linear regression model with k regressors and $n+1$ observations. A general expression for the posterior distribution in the calibration problem is considered in Corollary 2.3 for the conditional conjugate type family. For this, a m-dimensional generator density function g is denoted by $g^{(m)}$. Thus, for each $1 \leq k < m$, the corresponding k-dimensional marginal and conditional generator density functions are given by (see Fang et al., 1990)

$$g^{(k)}(u) = \frac{\pi^{\frac{m-k}{2}}}{\Gamma(\frac{m-k}{2})} \int_0^\infty r^{\frac{m-k}{2}-1} g^{(m)}(r+u)dr \quad \text{and} \quad g_v^{(k)}(u) = \frac{g^{(m)}(u+v)}{g^{(m-k)}(v)}, \tag{2.6}$$

$u, v \geq 0$, respectively.

Theorem 2.2 *Consider the dependent elliptical model given by*

$$\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{\beta}, \sigma^2 \sim El_{n+1}(\boldsymbol{X}_*\boldsymbol{\beta}, \sigma^2\boldsymbol{I}_{n+1}; g^{(n+1)}_{q(\beta)/\sigma^2}) \quad \textit{and} \quad \boldsymbol{\beta}|\boldsymbol{X}_*, \sigma^2 \sim El_k(\boldsymbol{b}_0, \sigma^2\boldsymbol{B}_0; g^{(k)}),$$

where $q(\boldsymbol{\beta}) = (\boldsymbol{\beta} - \boldsymbol{b}_0)'\boldsymbol{B}_0^{-1}(\boldsymbol{\beta} - \boldsymbol{b}_0)$ *and* $g_v^{(n+1)}$ *and* $g^{(k)}$ *are related through the generator density function* $g^{(n+1+k)}$ *as in (2.6). Then,*

$$\boldsymbol{\beta}|\boldsymbol{X}_*, \boldsymbol{y}_*, \sigma^2 \sim El_k(\boldsymbol{b}_*, \sigma^2\boldsymbol{B}_*; g^{(k)}_{q(\mathbf{y}_*)/\sigma^2}) \tag{2.7}$$

and

$$\boldsymbol{y}_*|\boldsymbol{X}_*, \sigma^2 \sim El_{n+1}(\boldsymbol{X}_*\boldsymbol{b}_0, \sigma^2(\boldsymbol{I}_{n+1} + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}_*'); g^{(n+1)}), \tag{2.8}$$

where

$$\boldsymbol{b}_* = \boldsymbol{A}_*\hat{\boldsymbol{\beta}}_* + (\boldsymbol{I}_{n+1} - \boldsymbol{A}_*)\boldsymbol{b}_0, \quad \boldsymbol{B}_* = \boldsymbol{B}_0 - \boldsymbol{B}_0\boldsymbol{X}_*'(\boldsymbol{I}_{n+1} + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}_*')^{-1}\boldsymbol{X}_*\boldsymbol{B}_0), \tag{2.9}$$

with $\boldsymbol{A}_* = \boldsymbol{B}_0\boldsymbol{X}_*'\boldsymbol{X}_*[\boldsymbol{I}_k - \boldsymbol{B}_0(\boldsymbol{I}_k + \boldsymbol{X}_*'\boldsymbol{X}_*)\boldsymbol{B}_0]^{-1}$, $\hat{\boldsymbol{\beta}}_* = (\boldsymbol{X}_*\boldsymbol{X}_*')^{-1}\boldsymbol{X}_*'\boldsymbol{y}_*$, *and*

$$q(\boldsymbol{y}_*) = (\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{b}_0)'(\boldsymbol{I}_{n+1} + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}_*')^{-1}(\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{b}_0). \tag{2.10}$$

Proof: Since (2.6) implies that $g_v^{(n+1)}(u)g^{(k)}(v) = g^{(n+1+k)}(u+v)$, we have that

$$p(\boldsymbol{y}_*, \boldsymbol{\beta}|\boldsymbol{X}_*, \sigma^2) = p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{\beta}, \sigma^2)p(\boldsymbol{\beta}|\boldsymbol{X}_*, \sigma^2) = \Big(\frac{1}{\sigma^2}\Big)^{\frac{n+1+k}{2}} |\boldsymbol{B}_0|^{-\frac{1}{2}} g^{(n+1+k)}\Big(\frac{q(\boldsymbol{y}_*, \boldsymbol{\beta})}{\sigma^2}\Big), \tag{2.11}$$

where $q(\boldsymbol{y}_*, \boldsymbol{\beta}) = (\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{\beta})'(\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{\beta}) + (\boldsymbol{\beta} - \boldsymbol{b}_0)'\boldsymbol{B}_0^{-1}(\boldsymbol{\beta} - \boldsymbol{b}_0)$. Now, after standard algebraic manipulations, we obtain

$$q(\boldsymbol{y}_*, \boldsymbol{\beta}) = \begin{pmatrix} \boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{b}_0 \\ \boldsymbol{\beta} - \boldsymbol{b}_0 \end{pmatrix}' \begin{bmatrix} \boldsymbol{I}_{n+1} + \boldsymbol{X}_* \boldsymbol{B}_0\boldsymbol{X}_*' & \boldsymbol{X}_*\boldsymbol{B}_0 \\ \boldsymbol{B}_0\boldsymbol{X}_*' & \boldsymbol{B}_0 \end{bmatrix}^{-1} \begin{pmatrix} \boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{b}_0 \\ \boldsymbol{\beta} - \boldsymbol{b}_0 \end{pmatrix}.$$

Thus, (2.11) implies that $(\boldsymbol{y}_*, \boldsymbol{\beta})|\boldsymbol{X}_*, \sigma^2 \sim El_{n+1+k}(\boldsymbol{\mu}, \sigma^2\boldsymbol{\Sigma}; g^{(n+1+k)})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{X}_*\boldsymbol{b}_0 \\ \boldsymbol{b}_0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{I}_{n+1} + \boldsymbol{X}_*\ \boldsymbol{B}_0\boldsymbol{X}'_* & \boldsymbol{X}_*\boldsymbol{B}_0 \\ \boldsymbol{B}_0\boldsymbol{X}'_* & \boldsymbol{B}_0 \end{bmatrix},$$

from where follows (2.7) and (2.8), after some algebraic manipulations. $\diamond$

Result (2.8) in Theorem 2.2 implies that if the variance is known then the predictive distribution is also in the elliptical family. This result is crucial in obtaining the posterior distribution for the calibration problem presented in Corollary 2.3. From (2.9) it follows also that $\boldsymbol{b}_*$ is a linear convex combination of $\boldsymbol{b}_0$ (prior mean of $\boldsymbol{\beta}$) and $\hat{\boldsymbol{\beta}}_*$ (the maximum likelihood estimator of $\boldsymbol{\beta}$). Moreover, $\boldsymbol{b}_*$ is independent of the generator density function $g^{(n+1+k)}$. This implies that under squared error loss, the Bayes estimator for $\boldsymbol{\beta}$ is $\boldsymbol{b}_*$ ($E[\boldsymbol{\beta}|\boldsymbol{X}_*, \boldsymbol{y}_*] = \boldsymbol{b}_*$) which also is independent of the elliptical model under consideration. Thus, the point estimator of $\boldsymbol{\beta}$ under the elliptical family is the point estimator under normality, a result which is independent of the prior distribution assigned to σ^2. However, the posterior variance of $\boldsymbol{\beta}$ depends on the particular elliptical model under consideration, as is noted, for example, in Zellner (1976). Another important result is given by the following relation:

$$Var[\beta_i|\boldsymbol{X}_*, \boldsymbol{y}_*] = E\{Var[\beta_i|\boldsymbol{X}_*, \boldsymbol{y}_*, \sigma^2]\} = b_{ii}E[\sigma^2 a_{g^{(k)}_{q(\mathbf{y}_*)/\sigma^2}}|\boldsymbol{X}_*, \boldsymbol{y}_*], \tag{2.12}$$

$i = 1, \ldots, k$, where $a_{g^{(k)}_{q(\mathbf{y}_*)/\sigma^2}}$ is the variance parameter associated with $g^{(k)}_{q(\mathbf{y}_*)/\sigma^2}$ and, as the notation shows, is a function of $\boldsymbol{y}_*$ and σ^2 through $q(\boldsymbol{y}_*)/\sigma^2$, where $q(\boldsymbol{y}_*)$ is defined in (2.10); b_{ij} is the component in the i-th row an j-th column of the matrix $\boldsymbol{B}_*$, and the last expectations in (2.12) are computed with respect to the posterior distribution of σ^2. The next Corollary follows as a consequence of this result.

Corollary 2.2 *If $p(\sigma^2|\boldsymbol{X}_*) \propto 1/\sigma^2$ and $g^{(n+1+k)}$ is independent of σ^2, then*

$$E[\sigma^2 a_{g^{(k)}_{q(\mathbf{y}_*)/\sigma^2}}|\boldsymbol{X}_*, \boldsymbol{y}_*] = b_{11}^{-1}Var_N[\beta_1|\boldsymbol{X}_*, \boldsymbol{y}_*],$$

where $Var_N(.)$ denote the variance under the normal model.

Proof: Osiewalski and Steel (1993) show that under the non informative prior distribution $p(\sigma^2|\boldsymbol{X}_*) \propto 1/\sigma^2$, the posterior density $p(\boldsymbol{\beta}|\boldsymbol{X}_*, \boldsymbol{y}_*)$ is independent of the elliptical generator density function. Thus, $V[\boldsymbol{\beta}|\boldsymbol{X}_*, \boldsymbol{y}_*]$ is the same for each elliptical generator density function and hence,

$$Var[\beta_1|\boldsymbol{X}_*, \boldsymbol{y}_*] = Var_N[\beta_1|\boldsymbol{X}_*, \boldsymbol{y}_*].$$

Thus, the result follows by using the relation (2.7) . $\diamond$

The result presented in the above corollary is important because it allows to derive an analytical expression for the posterior mean of a function of σ^2. We call attention to the fact that in the dependent elliptical model the posterior distribution of σ^2 depends on the generator g and can be extremely difficult to obtain.

Remark 2.2 *Goel and DeGroot (1980) show that under the linear regression model, linearity in $\boldsymbol{y}_*$ of the posterior mean of $\boldsymbol{\beta}$ implies that $\boldsymbol{\beta}$ and $\boldsymbol{\epsilon}_*$ are normality distributed. Thus, we can conclude that only the normal distribution presents posterior mean linear in the data. It seems that result (2.7) contradicts this fact. However, Goel and DeGroot (1980) assume that $\boldsymbol{\beta}$ and $\boldsymbol{\epsilon}_*$ are independent. We conjecture that similar results holds for elliptical distributions, that is, if the independence assumptions between $\boldsymbol{\beta}$ and $\boldsymbol{\epsilon}_*$ is replaced by null correlation, then linearity of the posterior mean of $\boldsymbol{\beta}$ implies that $\boldsymbol{\beta}$ and $\boldsymbol{\epsilon}_*$ are elliptically distributed.*

The following result can be obtained directly from (2.8).

Corollary 2.3 *Consider the dependent elliptical calibration model defined in Theorem 2.2. Then,*

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(x_0|\boldsymbol{X})|\boldsymbol{I}_{n+1} + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}_*'|^{-\frac{1}{2}} \int_0^\infty (1/\sigma^2)^{\frac{n+1}{2}} g^{(n+1)}(q(\boldsymbol{y}_*)/\sigma^2) p(\sigma^2|\boldsymbol{X}_*) d\sigma^2,$$

where $q(\boldsymbol{y}_*)$ is defined as in (2.10).

We call attention to the fact that, in general, we need numerical methods to find the posterior distribution of x_0 given in the Corollary 2.3. However, there are simplifications for some special cases, as we show in the following examples.

Example 2.1 *The normal calibration model.* Let $g^{(m)}(u) = (2\pi)^{-m/2}e^{-u/2}$ be the generator density function of the m-dimensional normal distribution, namely, $N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Note that, the generator density function is such that $g^{(m)}(u+v) = g^{(m-k)}(u)g^{(k)}(v)$, so that $g_v^{(m-k)}(u) = g^{(m-k)}(u) = (2\pi)^{-k/2}e^{-u/2}$. Thus, letting $m = n+1+k$, with $k = 2$, in Theorem 2.2 we obtain the normal calibration model $\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{\beta}, \sigma^2 \sim N_{n+1}(\boldsymbol{X}_*\boldsymbol{\beta}, \sigma^2\boldsymbol{I}_{n+1})$, with conditional prior distribution $\boldsymbol{\beta}|\boldsymbol{X}_*, \sigma^2 \sim N_2(\boldsymbol{b}_0, \sigma^2\boldsymbol{B}_0)$ for $\boldsymbol{\beta}$. Assuming also that $\sigma^2|\boldsymbol{X}_* \sim IG(d_0/2, \lambda_0/2)$ (the inverse-gamma distribution with parameters $d_0/2$ and $\lambda_0/2$), we obtain the conjugate prior specification for $(\boldsymbol{\beta}, \sigma^2)$, that is, the normal-inverse-gamma prior distribution $(\boldsymbol{\beta}, \sigma^2)|\boldsymbol{X}_* \sim NIG(\boldsymbol{b}_0, \boldsymbol{B}_0, d_0, \lambda_0)$, having density of the form

$$p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}_*) \propto (1/\sigma^2)^{\frac{d_0}{2}+1}\exp\{-[\lambda_0 + (\boldsymbol{\beta} - \boldsymbol{b}_0)'\boldsymbol{B}_0^{-1}(\boldsymbol{\beta} - \boldsymbol{b}_0)]/2\sigma^2\}. \tag{2.13}$$

Moreover, from (2.8) we have that $\boldsymbol{y}_*|\boldsymbol{X}_*, \sigma^2 \sim N_{n+1}(\boldsymbol{X}_*\boldsymbol{b}_0, \sigma^2(\boldsymbol{I}_{n+1}+\boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}'_*))$. It follows, from Corollary 2.3, that

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(x_0|\boldsymbol{X})|\boldsymbol{I}_{n+1} + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}'_*|^{-\frac{1}{2}}\{\lambda_0 + q(\boldsymbol{y}_*)\}^{-\frac{n+1+d_0}{2}}.$$

Using also the partition

$$\boldsymbol{y}_* = \begin{pmatrix} y_0 \\ \boldsymbol{y} \end{pmatrix}, \quad \boldsymbol{X}_*\boldsymbol{b}_0 = \begin{pmatrix} \boldsymbol{x}'_0\boldsymbol{b}_0 \\ \boldsymbol{X}\boldsymbol{b}_0 \end{pmatrix}, \quad \boldsymbol{I}_{n+1} + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}'_* = \begin{pmatrix} 1 + \boldsymbol{x}'_0\boldsymbol{B}_0\boldsymbol{x}_0 & \boldsymbol{x}'_0\boldsymbol{B}_0\boldsymbol{X}' \\ \boldsymbol{X}\boldsymbol{B}_0\boldsymbol{x}_0 & \boldsymbol{I}_n + \boldsymbol{X}\boldsymbol{B}_0\boldsymbol{X}' \end{pmatrix},$$

where $\boldsymbol{x}_0 = (1, x_0)'$, it follows that

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(x_0|\boldsymbol{X})\Big\{\frac{1}{1 + \boldsymbol{x}'_0\boldsymbol{B}_1\boldsymbol{x}_0}\Big\}^{\frac{1}{2}}\Big\{\lambda_{q(\mathbf{y})} + \frac{(y_0 - \boldsymbol{x}'_0\boldsymbol{b}_1)^2}{1 + \boldsymbol{x}'_0\boldsymbol{B}_1\boldsymbol{x}_0}\Big\}^{-\frac{n+1+d_0}{2}}, \tag{2.14}$$

where $\lambda_{q(\mathbf{y})} = \lambda_0 + q(\boldsymbol{y})$, with $q(\boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}_0)'(\boldsymbol{I}_n + \boldsymbol{X}\boldsymbol{B}_0\boldsymbol{X}')^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}_0)$,

$$\boldsymbol{b}_1 = \boldsymbol{b}_0 + \boldsymbol{B}_0\boldsymbol{X}'(\boldsymbol{I}_n + \boldsymbol{X}\boldsymbol{B}_0\boldsymbol{X}')^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}_0) \quad \text{and} \quad \boldsymbol{B}_1 = \boldsymbol{B}_0 - \boldsymbol{B}_0\boldsymbol{X}'(\boldsymbol{I}_n + \boldsymbol{X}\boldsymbol{B}_0\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{B}_0.$$

The result in (2.14) presents an analytical expression for the posterior distribution of x_0. This extends the work of Hoadley (1970) to proper posterior distributions for $(\boldsymbol{\beta}, \sigma^2)$.

Example 2.2 *The generalized Student-t calibration model.* Let $g^{(m)}(u) = c(m, \nu)\lambda^{\nu/2}\{\lambda + u\}^{-(m+\nu)/2}$, $u \geq 0$, where $c(m, \nu) = \Gamma[(m + \nu)/2]/\Gamma[\nu/2]\pi^{m/2}$, be the generator density function of the m-dimensional generalized Student-t distribution, namely, $t_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \lambda, \nu)$, which reduces to the ordinary m-dimensional Student-t distribution with ν degrees of freedom when $\lambda = \nu$, (see Arellano-Valle and Bolfarine, 1995). Thus, from (2.6) we have that $g^{(k)}(u) = c(k, \nu)\lambda^{\nu/2}\{\lambda + u\}^{-(k+\nu)/2}$ and $g_v^{(k)}(u) = c(k, \nu_k)\lambda_v^{\nu_k/2}\{\lambda_v + u\}^{-(k+\nu_k)/2}$, where $\nu_k = \nu + m - k$ and $\lambda_v = \lambda + v$. Consequently, in Theorem 2.2 we have the generalized Student-t calibration model $\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{\beta}, \sigma^2 \sim t_{n+1}(\boldsymbol{X}_*\boldsymbol{\beta}, \sigma^2\boldsymbol{I}_{n+1}; \lambda + q(\boldsymbol{\beta})/\sigma^2, \nu)$ and the conditional prior distribution $\boldsymbol{\beta}|\boldsymbol{X}_*, \sigma^2 \sim t_2(\boldsymbol{b}_0, \sigma^2\boldsymbol{B}_0; \lambda, \nu)$. Therefore from (2.8), $\boldsymbol{y}_*|\boldsymbol{X}_*, \sigma^2 \sim t_{n+1}(\boldsymbol{X}_*\boldsymbol{b}_0, \sigma^2(\boldsymbol{I}_{n+1} + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}'_*); \lambda, \nu)$, thus, from Corollary 2.3, it follows that

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(x_0|\boldsymbol{X})|\boldsymbol{I}_{n+1} + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}'_*|^{-\frac{1}{2}} \int_0^\infty (\lambda\sigma^2)^{\nu/2}\{\lambda\sigma^2 + q(\boldsymbol{y}_*)\}^{-\frac{n+1+\nu}{2}} p(\sigma^2|\boldsymbol{X}_*)d\sigma^2. \tag{2.15}$$

Note that if $\lambda = \lambda(\sigma^2) = \lambda_0/\sigma^2$, the above model is independent of σ^2 and (2.15) reduces to (2.14) with $d_0 = \nu$. This same result is obtained if $\nu = \nu_0 + d_0$, $\lambda = \lambda_0 + 1/\sigma^2$ and $p(\sigma^2|\boldsymbol{X}_*) \propto (\sigma^2)^{\nu_0/2-1}\{1 + \lambda_0\sigma^2\}^{-(\nu_0+d_0)/2}$, which is a proper prior distribution. (Note that for $\lambda_0 = \nu_0/d_0$, $\sigma^2 \sim F_{\nu_0,d_0}$, that is, the Fisher distribution with ν_0 and d_0 degrees of freedom).

3 A Bayesian solution for the independent case

In this section, it is considered that the error terms in equations (1.1)-(1.2) are independent with common distribution $El_1(0, \sigma^2; g)$, which yields the joint density given in (1.5). From (1.5) it follows that the likelihood for the independent elliptical model is given by

$$p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{\beta}, \sigma^2) = \left(\frac{1}{\sigma^2}\right)^{\frac{n+1}{2}} \prod_{i=0}^{n} g\left[\frac{(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2}{\sigma^2}\right], \tag{3.1}$$

where $\boldsymbol{x}_=(1, x_i)'$, $i = 0, 1, \ldots, n$, and as before, $\boldsymbol{y}_* = (y_0, \boldsymbol{y}')'$, $\boldsymbol{X}_* = [\mathbf{1}_{n+1} \quad \boldsymbol{x}]$, with $\boldsymbol{y} = (y_1, \ldots, y_n)'$, $\boldsymbol{x} = (x_1, \ldots, x_n)'$, and $\boldsymbol{\beta} = (\beta_1, \beta_2)'$. In other words,

$$y_i|\boldsymbol{x}_i, \boldsymbol{\beta}, \sigma^2 \overset{ind}{\sim} El_1(\boldsymbol{x}_i'\boldsymbol{\beta}, \sigma^2, g), \tag{3.2}$$

$i = 0, 1, \ldots, n$, so that, $y_0, \ldots, y_n$ are conditionally independent given $\boldsymbol{X}_*$, $\boldsymbol{\beta}$ and σ^2. We consider here the case where the function g is representable, that is,

$$g(u) = \int_0^\infty (2\pi v)^{-1/2} \mathrm{e}^{-u/2v} dG(v), \tag{3.3}$$

for any distribution function G such that $G(0) = 0$. Representation (3.3) defines a subclass of the elliptical family, namely, the compound normal distributions. Under this subclass, the specification (3.2) can be represented equivalently by considering the following two steps:

(i) $y_i|\boldsymbol{x}_i, v_i, \boldsymbol{\beta}, \sigma^2 \overset{ind}{\sim} N(\boldsymbol{x}_i'\boldsymbol{\beta}, \sigma^2 v_i), \quad i = 1, \ldots, n,$

and

(ii) $v_i \overset{iid}{\sim} G, \quad i = 1, \ldots, n,$ and independent of $\boldsymbol{\beta}$, σ^2 and $\boldsymbol{X}_*$.

Note that the correspondence between the class of the density generator functions g which are representable and the class of the distribution functions G such that $G(0) = 0$ is one to one, so that the Bayesian analysis of the calibration problem under specification (3.1) is equivalent to the Bayesian analysis under specifications (i) and (ii), that is, under the heteroscedastic normal model. One advantage of this last specification is the fact that the derivation of the posterior distributions can be based on properties of the normal model, according to step (i). For step (ii) the main problem is the choice of the distribution function $G_*(\boldsymbol{v}_*) = \Pi_{i=0}^n G(v_i)$ for the weight vector $\boldsymbol{v}_* = (v_0, v_1, \ldots, v_n)'$ and the numerical integration problem that results. Thus, under conditions (i) and (ii) in relation (2.6), we have that

$$p(\boldsymbol{y}_*|\boldsymbol{X}_*) = \int_{\boldsymbol{v}_*} p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{v}_*) dG_*(\boldsymbol{v}_*), \tag{3.4}$$

where

$$(3.5) \qquad p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{v}_*) = \int_{\boldsymbol{\beta}} \int_0^\infty p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{v}_*, \boldsymbol{\beta}, \sigma^2) p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}_*) d\sigma^2 \mathbf{d}\boldsymbol{\beta},$$

with

$$(3.6) \quad p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{v}_*, \boldsymbol{\beta}, \sigma^2) = (1/2\pi\sigma^2)^{\frac{n+1}{2}} |\boldsymbol{V}_*|^{-1/2} \exp\{-(\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{\beta})' \boldsymbol{V}_*^{-1} (\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{\beta})/2\sigma^2\},$$

and $\boldsymbol{V}_* = \operatorname{diag}\{v_0, v_1, \ldots, v_n\}$.

Notice that when the conditional posterior distribution $p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}_*, \boldsymbol{y}_*, \boldsymbol{v}_*)$ is available, (3.5) can be computed from the relation

$$p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{v}_*) = \frac{p(\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{v}_*, \boldsymbol{\beta}, \sigma^2) p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}_*)}{p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}_*, \boldsymbol{y}_*, \boldsymbol{v}_*)}.$$

If $p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}_*)$ is an improper prior distribution, the posterior distribution $p(x_0|\boldsymbol{X}, \boldsymbol{y}_*)$ may not be a proper distribution. In this case it is necessary to study conditions under which such posterior is a proper distribution.

Two specifications for the prior distribution $p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}_*)$ are considered next. Both are motivated by the Bayesian formulation of the usual regression problem. The first is the usual improper reference prior distribution:

$$(3.7) \qquad \pi(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}_*) \propto \frac{1}{\sigma^2}.$$

The second is the normal-inverse-gamma prior distribution $(\boldsymbol{\beta}, \sigma^2)|\boldsymbol{X}_* \sim NIG(\boldsymbol{b}_0, \boldsymbol{B}_0, d_0, \lambda_0)$ (see Example 2.1), which has density as in (2.13). As usually, $\boldsymbol{b}_0$ represents a prior assignment to $\boldsymbol{\beta}$ with $\boldsymbol{B}_0$ expressing uncertain in that assignment.

In both cases, a proper prior distribution, $p(x_0|\boldsymbol{X})$, is considered for x_0, which however does not guarantee a proper posterior distribution for x_0 in the first case.

Theorem 3.1 *Let consider the model defined by (3.1) and (3.3) and the prior distribution given in (3.7). If $p(x_0|\boldsymbol{X})$ is a proper prior density and $n \geq 3$, then the posterior distribution of x_0 is also proper.*

Proof: First we note that

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(x_0|\boldsymbol{X}) p(y_0|\boldsymbol{X}_*, \boldsymbol{y}),$$

where

$$(3.9) \qquad p(y_0|\boldsymbol{X}_*, \boldsymbol{y}) = \int_0^\infty \int_{\boldsymbol{\beta}} p(\boldsymbol{y}_0|\boldsymbol{X}_*, \boldsymbol{y}, \boldsymbol{\beta}, \sigma^2) p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}, \boldsymbol{y}) \boldsymbol{\beta} d\sigma^2.$$

Now, as is showed in Fernandez and Steel (1999), $p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{X}, \boldsymbol{y})$ is proper if $n \geq 3$, so that $p(y_0|\boldsymbol{X}_*, \boldsymbol{y})$ in (3.9) is also proper. The proof follows then from the fact that $p(x_0|\boldsymbol{X})$ is proper. $\diamond$

The next result is related to second case with prior distribution given by (2.13). We use the notation $t_p(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda, \nu)$ to denote the density of the p-dimensional generalized Student-t distribution $\boldsymbol{z} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda, \nu)$ (see Example 2.2).

Theorem 3.2 *Let $(\boldsymbol{\beta}, \sigma^2)|\boldsymbol{X}_* \sim NIG(\boldsymbol{b}_0, \boldsymbol{B}_0, d_0, \lambda_0)$ and $p(x_0|\boldsymbol{X})$ a proper prior density for x_0. Then, the posterior density that follows is given by*

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(x_0|\boldsymbol{X}) \int_{\boldsymbol{\mathcal{V}}_*} t_{n+1}(\boldsymbol{y}_*; \boldsymbol{X}_*\boldsymbol{b}_0, \boldsymbol{V}_* + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}'_*, \lambda_0, d_0) dG_*(\boldsymbol{v}_*).$$

Proof: Let $\boldsymbol{z}_* = \boldsymbol{V}_*^{-1/2}\boldsymbol{y}_*$, $\mathbf{W}_* = \boldsymbol{V}_*^{-1/2}\boldsymbol{X}_*$, where as before $\boldsymbol{X}_* = [\mathbf{1}_{n+1} \quad \boldsymbol{x}_*]$, which is of dimension $(n+1) \times 2$. Thus, it follows that $\mathbf{Z}_*|\boldsymbol{X}_*, \boldsymbol{v}_*, \boldsymbol{\beta}, \sigma^2, \sim N_{n+1}(\mathbf{W}_*\boldsymbol{\beta}, \sigma^2\boldsymbol{I}_{n+1})$. Now, since $(\boldsymbol{\beta}, \sigma^2)|\mathbf{X}_* \sim NIG(\boldsymbol{b}_0, \boldsymbol{B}_0, d_0, \lambda_0)$, straightforward algebraic manipulations yields

$$\boldsymbol{z}_*|\boldsymbol{X}_*, \boldsymbol{v}_* \sim t_{n+1}(\mathbf{W}_*\boldsymbol{b}_0, \boldsymbol{I}_{n+1} + \mathbf{W}_*\boldsymbol{B}_0\mathbf{W}'_*, \lambda_0, d_0).$$

Hence, simple transformation yields

$$\boldsymbol{y}_*|\boldsymbol{X}_*, \boldsymbol{v}_* \sim t_{n+1}(\boldsymbol{X}_*\boldsymbol{b}_0, \boldsymbol{V}_* + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}'_*, \lambda_0, d_0),$$

from where it follows that

$$p(\boldsymbol{y}_*|\boldsymbol{X}_*) = \int_{\boldsymbol{\mathcal{V}}_*} t_{n+1}(\boldsymbol{y}_*; \boldsymbol{X}_*\boldsymbol{b}_0, \boldsymbol{V}_* + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}'_*, \lambda_0, d_0) dG_*(\boldsymbol{v}_*),$$

which concludes the proof. $\diamond$

In the special case of G_* being degenerate at the point $\boldsymbol{v}_* = (1, \ldots, 1)'$, the posterior density of x_0 can be written as

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(x_0|\boldsymbol{X}) t_{n+1}(\boldsymbol{y}_*; \boldsymbol{X}_*\boldsymbol{b}_0, \boldsymbol{I}_{n+1} + \boldsymbol{X}_*\boldsymbol{B}_0\boldsymbol{X}'_*, \lambda_0, d_0),$$

from where follows the expression (2.14), by noting first that

$$p(x_0|\boldsymbol{X}, \boldsymbol{y}_*) \propto p(x_0|\boldsymbol{X}) p(y_0|\boldsymbol{X}_*, \boldsymbol{y}) p(\boldsymbol{y}|\boldsymbol{X})$$

and using the properties of the generalized Student-t distribution (see Arellano-Valle and Bolfarine, 1995), we can write

$$y_0|\boldsymbol{X}_*, \boldsymbol{y} \sim t_1(\boldsymbol{x}_0'\boldsymbol{b}_1, 1 + \boldsymbol{x}_0'\boldsymbol{B}_1\boldsymbol{x}_0; \lambda_0 + q(\boldsymbol{y}), d_0 + n).$$

The two theorems above imply that the posterior distribution of x_0 in the calibration problem depends on the generating density g only through G_*. However, it is worth noticing that it is not possible to obtain closed analytical expressions for the posterior distribution of x_0, except in the case of the normal distribution. One alternative in such situations is to employ Markov Chain Monte Carlo (MCMC) based approaches. The approach allows the derivation of precise estimates of the posterior density of interest. To implement the Gibbs sampling algorithm (Gelfand and Smith, 1990), it is required to derive the conditional posterior distribution of one parameter given the others. Considering proper priors in the conditional conjugate class, the conditional posterior distributions involve well known and easy to simulate distributions.

We suppose that $x_0|\boldsymbol{X} \sim N(m_0, w_0)$ and $(\boldsymbol{\beta}, \sigma^2)|\boldsymbol{X} \sim NGI(\boldsymbol{b}_0, \boldsymbol{B}_0, d_0, \lambda_0)$ are independent. Then,

(i) $\boldsymbol{\beta}|\boldsymbol{X}_*, \boldsymbol{y}_*, \boldsymbol{v}_*, \sigma^2 \sim N_2(\hat{\boldsymbol{\beta}}_*(\boldsymbol{v}_*), \sigma^2 \boldsymbol{B}_*(\boldsymbol{v}_*))$, where $\hat{\boldsymbol{\beta}}_*(\boldsymbol{v}_*)$ is given in (3.12) and

$$\boldsymbol{B}_*(\boldsymbol{v}_*) = (\boldsymbol{B}_0^{-1} + [\boldsymbol{X}_*' \boldsymbol{V}_*^{-1} \boldsymbol{X}_*]^{-1})^{-1}.$$

(ii) $\sigma^2|\boldsymbol{X}_*, \boldsymbol{y}_*, \boldsymbol{v}_*, \boldsymbol{\beta} \sim IG(d_1/2, \lambda_1/2)$, where

$$d_1 = d_0 + n \quad \text{and} \quad \lambda_1 = \lambda_0 + (\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{\beta})' \boldsymbol{V}_*^{-1} (\boldsymbol{y}_* - \boldsymbol{X}_*\boldsymbol{\beta});$$

(iii) $x_0|\boldsymbol{y}_*, \boldsymbol{v}_*, \boldsymbol{\beta}, \sigma^2 \sim N(m_1, w_1)$, where

$$m_1 = \Big(\frac{v_0\sigma^2}{v_0\sigma^2 + w_0\beta_2^2}\Big)m_0 + \Big(\frac{w_0\beta_2^2}{v_0\sigma^2 + w_0\beta_2^2}\Big)\Big(\frac{y_0 - \beta_1}{\beta_2}\Big) \quad \text{and} \quad w_1 = \frac{w_0 v_0 \sigma^2}{v_0\sigma^2 + w_0\beta_2^2}.$$

The conditional posterior to $\boldsymbol{v}_*$ given the other parameters depends of the form of the generator function g. Some examples are present in the following.

Example 3.1 *The Student-t calibration model.* Considering that

$$y_i|\boldsymbol{x}_i, \boldsymbol{\beta}, \sigma^2, \nu \overset{ind.}{\sim} t_1(\boldsymbol{x}_i'\boldsymbol{\beta}, \sigma^2; \nu),$$

$i = 0, \ldots, n$, it follows that the density function $p(y_i|\boldsymbol{x}_i, \boldsymbol{\beta}, \sigma^2, \nu)$ is representable with the mixing measure given by the inverse gamma distribution, that is, $IG(\nu/2, \lambda/2)$. Thus, the generating function can be written as

$$g(u) = \int_0^\infty (2\pi s)^{-n/2} \mathrm{e}^{-1/2s} h(s) ds,$$

where $h(s) = (\nu/2)^{\nu/2}(1/s)^{\nu/2+1}\mathrm{e}^{-\nu/2s}/\Gamma[\nu/2]$, which is the density function of the inverted gamma distribution. This shows that the assumption of independent generalized Student-t errors is equivalent to considering a hierarchical formulation with heteroscedastic errors, by assuming that the weights $v_0, \ldots, v_n$ are independent with $v_i \sim \nu/\chi^2_\nu$. The heteroscedastic approach is considered in Geweke (1993) for estimating the regression line parameters. The posterior distribution of $\boldsymbol{v}_*$, which is required in the implementation of the Gibbs sampling algorithm in the Student-t model is given by

$$p(\boldsymbol{v}_*|\boldsymbol{X}_*, \boldsymbol{y}_*, \boldsymbol{\beta}, \sigma^2) = \prod_{i=0}^{n} p(v_i|\boldsymbol{X}_*, \boldsymbol{y}_*, \beta, \sigma^2),$$

where

$$v_i|\boldsymbol{X}_*, \boldsymbol{y}_*, \boldsymbol{\beta}, \sigma^2 \overset{ind}{\sim} \nu_i/\chi^2_{\nu+1},$$

$i = 0, \ldots, n$ with $\nu_i = \nu + (y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2/\sigma^2$.

Example 3.2 *The contaminated normal model.* We consider now that

$$p(y_i|\boldsymbol{x}_i, \boldsymbol{\beta}, \sigma^2, \tau^2, \varepsilon) = \varepsilon f_N(y_i|\boldsymbol{x}_i'\boldsymbol{\beta}, \sigma^2) + (1-\varepsilon) f_N(y_i|\boldsymbol{x}_i'\boldsymbol{\beta}, (\tau\sigma)^2),$$

$i = 0, 1, .., n$, where $f_N(.|\mu, \psi)$ denotes the normal density with mean μ and variance ψ. This model is representable with the mixing measure given by

$$G(v) = \begin{cases} \varepsilon, & \text{if } v = 1 \\ 1-\varepsilon, & \text{if } v = \tau^2. \end{cases}$$

Thus, the assumption of contaminated normal error is equivalent to the assumption of heteroscedastic normal errors with weights $v_0, \ldots, v_n$ i.i.d. G. In this case, $G_*(\boldsymbol{v}_*) = \prod_{i=0}^{n} G(v_i) = \varepsilon^r(1-\varepsilon)^{n+1-r}$, where r is the number of 1's in $\boldsymbol{v}_*$. To implement the Gibbs sampler, we have to specify a distribution for $p(\boldsymbol{v}_*|\boldsymbol{X}_*, \boldsymbol{y}_*, \boldsymbol{\beta}, \sigma^2)$ assuming that τ^2 and ε are known. It follows that

$$p(\boldsymbol{v}_*|\boldsymbol{X}_*, \boldsymbol{y}_*, \boldsymbol{\beta}, \sigma^2) \propto \prod_{i=0}^{n} \left\{ \varepsilon \exp\Big\{-\frac{(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2}{2\sigma^2}\Big\} \boldsymbol{I}_{(v_i=1)} + \frac{1-\varepsilon}{\tau} \exp\Big\{-\frac{(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2}{2(\tau\sigma)^2}\Big\} \boldsymbol{I}_{(v_i=\tau^2)} \right\},$$

so that $v_0, \ldots, v_n$ are conditionally independent with marginal density given by

$$p(v_i|\boldsymbol{X}_*, \boldsymbol{y}_*, \boldsymbol{\beta}, \sigma^2) = \frac{\varepsilon}{K}\phi\Big(\frac{y_i - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\Big)\boldsymbol{I}_{(v_i=1)} + \frac{1-\varepsilon}{\tau K}\phi\Big(\frac{y_i - \boldsymbol{x}_i'\boldsymbol{\beta}}{\tau\sigma}\Big)\boldsymbol{I}_{(v_i=\tau^2)},$$

where ϕ denotes the standard normal density and

$$K = \varepsilon\phi\Big(\frac{y_i - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\Big) + \frac{1-\varepsilon}{\tau}\phi\Big(\frac{y_i - \boldsymbol{x}_i'\boldsymbol{\beta}}{\tau\sigma}\Big).$$

4 An Application

The data set considered next was obtained at the Pediatric Department of the Pontificia Universidad Catolica de Chile. Two measures of body fatness are considered for evaluating the percentage of body fatness: DEXA (dual-energy X-ray absortiometry) and the equation of Durnin, relating skinfold thickness and density. The Table 4.1. presents the measurements for a group of n=24 women aged 16 - 19 years. The main objective of the study is to predict the body fatness by DEXA, denoted by x_0, from a measure obtained by using the equation of Durnin, denoted by y_0. We consider first the classical analysis of the calibration problem where the point estimate is the classical estimator and confidence intervals are based on the F statistics. Next, we consider a Bayesian approach following results from Hoadley (1970), where the point estimator is the inverse estimator and credibility intervals follow from Corollary 2.1. Finally, we consider the Bayesian analysis based on the independent Student-t model with ν degrees of freedoms. The new step here is to consider a priori distribution for ν. Similarly as considered in Geweke (1993) and recently in Fernandez and Steel (1999), we assume an exponential distribution for the degrees of freedom parameter, namely, $p(\nu) = d\exp\{-d\nu\}$, with $d = 0.10$, thus obtaining mean 10 and variance 100 which spreads the prior mass over a wide variety of values of ν. Completing the prior specification, let $\beta_1 \sim N(9.8, 100), \beta_2 \sim N(0.6, 1)$ and $\sigma^2 \sim IG(0.1, 0.1)$. The prior distributions for the regression parameters are based in the maximum likelihood estimates under the normal model. Two different prior distributions are considered for the parameter of interest x_0, one less informative (Prior I) and other more informative (Prior II). The first, $N(31.5, 52.94)$, does not depend of the new observation y_0 and is considered for cross-validation analysis. This prior distribution is an approximation of the prior considered in Hoadley (1970), that is, in this case, a Student-t distribution with 21 degree of freedom, mean 31.5 and scale parameter 52.94. The second prior distribution will be more informative and given next.

Two new women where observed with Durnim measures given by 26.98 and 33.25. The first one is considered thin and the second fat. The new prior distributions for x_0 are $N(22.57, 16)$ and $N(39.69, 14.52)$, respectively, and were constructed jointly with the medical researcher in charge of the project. The values of the mean and variance for the first [second] normal distribution, produce a 4σ-range of $(18.50, 26.65)[(35.88, 43.50)]$, where the limits of the intervals are the minimum and the first quartile [the third quartile and the maximum] of the 24 DEXA measures. Estimates for the DEXA measure were obtained using the three approaches and the results are given in the Table 4.2. The posterior mean is used as point estimates in the Bayesian approach with independent Student-t errors and the intervals derived are equal tail intervals.

Table 4.1: Body fatness in a group of 24 women aged 16-19 years.

Woman	DEXA(x)	Durnim(y)	Woman	DEXA(x)	Durnim(y)
1	28.6	28.34	13	24.7	27.51
2	24.8	25.68	14	32.7	31.25
3	30.9	25.26	15	34.2	31.16
4	32.2	29.12	16	28.9	32.86
5	35.6	35.56	17	23.8	25.96
6	18.4	18.86	18	25.6	22.60
7	37.9	35.06	19	34.9	24.53
8	27.0	32.77	20	30.7	26.88
9	37.6	34.84	21	36.7	31.44
10	40.8	33.84	22	42.4	33.84
11	33.7	30.87	23	43.5	35.56
12	20.6	18.63	24	29.7	25.12

The Gibbs samples were generated by using a program implemented in the software WinBUGS (see Spiegelhalter et al., 1994), with an average time of 46 seconds to generate a sample of size 70,000, disregarding the 1000 initial iterations. A cross-validation analysis was done, considering the three proposals, where the following values for mean squared deviation ($MSD = \sum_{i=1}^{24}(x_i - \hat{x}_i)^2/24$) were obtained: 27.19, 17.65 and 16.96, respectively, for classical, Hoadley's inverse estimator and Student-t model. The results seem to indicate that the Bayesian analysis with independent Student-t errors presents much better results than the classical analysis and a little better than Hoadley's inverse estimator.

Table 4.2: Point and Interval (0.95) estimates for DEXA (x_0).

y_0	Classical	Hoadley	Prior I	Prior II
26.98	28.11	29.25	28.92	24.22
(thin)	(21.22, 41.23)	(19.92, 38.58)	(20.12, 38.12)	(17.49, 30.46)
33.25	38.36	36.04	35.48	39.09
(fat)	(29.04, 56.17)	(24.58, 47.50)	(25.83, 44.25)	(33.09, 45.32)

Generally, our credibility intervals are more precise than classical and Hoadley's intervals. The point estimates are very similar in the first case (thin) and in the second case (fat) the classical estimate is somewhat different. For practical purposes, Prior II is more interesting because the estimative belong to the same category. That is, if a woman is considered thin (fat) using the Durnim measure she is also considered thin (fat) using the DEXA measure. In the DEXA scale a woman is thin (fat) if $X < 26$ ($X > 36$), therefore the estimate 24.22

(39.09) is in that category. The same does not happen with the other estimates in the Table 4.2., except for the classical estimate in the second case. This is associated with the fact that Hoadley's (or inverse) estimator and the Bayes estimator with Prior I can be at some extent influenced by a prior distribution putting mass around $\bar{x}$ (the mean of the fixed x).

5 Concluding remarks

The present paper considers Bayesian solutions to the linear calibration problem under two types of elliptical models. Under non informative priors for the scale parameter, it is shown that the results obtained under the normal model remain valid under dependent elliptical models. The robustness results derived for the dependent elliptical model do not hold under independent elliptical models. Moreover, as in the case of a single covariate, closed form expressions for the posterior distribution of x_0 are difficult to obtain. However, in the represented case, the Gibbs sampler, provides a natural framework for obtaining approximations to the posterior distribution of x_0. As in the dependent case, these results can be extended to the multivariate calibration problem. As suggested by a referee an interesting case between the dependent and independent cases would follow by considering that $(\epsilon_1, \ldots, \epsilon_n)$ is from a n-dimensional elliptical distribution and ϵ_0 is from another independent elliptical distribution. The implementation of the approach developed in the paper for this more general situation is also similar to the one described in the section 3.

Acknowledgments: The authors acknowledge Doctor Gloria Icaza of the INTA, University of Chile, and Doctor Andreina Cattani of the Pediatric Department of Catholic University of Chile for providing the data set used in Section 5 and also for helpful discussions on the medical aspects of the problem. Partial support from Fondecyt-Chile under grants 1960937, 1990431, 7990041 and 1971128, Andes-Chile under grant C/127779 and CNPq-Brasil is also acknowledged.

References

Arellano-Valle, R.B. and Bolfarine, H. (1995). On some characterizations of the t-distribution. *Statistics and Probability Letters,* **25**, 79-85.

Brown, P.J. (1982). Multivariate calibration.*Journal of the Royal Statistical Society,* B, **44**, 287-321.

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

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

Fang, K.T. and Kotz, S. and Ng, K.W. (1990). *Symmetric multivariate and related distributions.* London-New York:Chapman and Hall.

Fernandez, C. and Steel, M. (1999). Multivariate Student-t regression models: pitfalls and inference. *Biometrika,* **86**, 153-167.

Gelfand, A.E. and Smith, A.F. (1990). Sampling-Based Approaches to Calculating Marginal Densities.*Journal of the American Statistical Association,* **vol.85, No.410**, 398-409.

Geweke, J. (1993). Bayesian treatment of the independent Student-t linear model. *Journal of Applied Econometrics,* **8**, 519-540.

Hoadley, B. (1970). A Bayesian look at inverse linear regression, *Journal of the American Statistician,* **65**, 356-369.

Hunter, W.G. and Lamboy, W.F. (1981). A Bayesian analysis of the linear calibration. *Technometrics*, **23**, 323-350.

Kelker,D. (1970). Distribution theory of special distributions and a location-scale parameter. *Sankhyā,* **A, 32,** 319-430.

Kubokawa, T. and Robert, C.P. (1994). New perspectives on linear calibration. *Journal of Multivariate Analysis*, **51**, 178-200.

Lange, K.L. and Little, R.J.A. and Taylor, J.M.G. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistician*, **84**, 881-896.

Osiewalski, J. and Steel, M.F.J. (1993). Robust Bayesian inference in elliptical regression models. *Journal of Econometrics*, **57**, 345-363.

Smith, A.F.M. and Roberts, G.O. (1993). Bayesian computation via Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*, B, **55**, 3-23.

Spiegelhalter, D.J. and Thomas, A. and Best, N.G. and Gilks,W.R. (1994). BUGS: Bayesian Inference Using Gibbs Sampling. version 0.30 *manual.*

Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error term. *Journal of the American Statistician*, **71**, 400-405.