RT-MAE 9623

BAYESIAN CALIBRATION UNDER A STUDENT-t MODEL

by

Márcia Branco, Heleno Bolfarine and Pilar Iglesias

Palavras-Chave: (Key words) **Calibration, Student-t model, regression model, Gibbs sampler, Bayesian inference, likelihood.**

Classificação AMS: 62F15, 62F35, 62J05, 62E17. (AMS Classification)

- Outubro de 1996 -

Bayesian Calibration under a Student-t model

Márcia Branco* Heleno Bolfarine† Pilar Iglesias‡

October 18, 1996

Summary

In this paper we consider linear calibration problems in regressions models with independent errors distributed according to the Student-t distribution. The approach followed is Bayesian, thus, involving the need for the specification of prior distributions for the model parameters. It is shown that the problem is equivalent to considering an heterochedastic regression model with an appropriate prior distributions on the model variances. By considering this alternative construction for the Student-t calibration model it is possible to use the Gibbs sampler to estimate the marginal posterior distributions. Simulation studies are reported which illustrate the performance of the approach proposed. An application to a data set analysed by Smith and Corbett (1987) on measuring marathon courses is reanalyzed by using the approach developed in the paper.

Key words: Calibration, Student-t model, regression model, Gibbs sampler, Bayesian inference, likelihood.

*Instituto de Matemática e Estatística, Universidade de São Paulo,CP 66281, 05389-970, São Paulo, SP, Brasil, mbranco@ime.usp.br

†Instituto de Matemática e Estatística, Universidade de São Paulo,CP 66281, 05389-970, São Paulo, SP, Brasil, hbolfar@ime.usp.br

‡Universidad catolica de Chile,Santiago,Chile,pliz@mat.puc.cl

1.Introduction

Controlled calibration problems can be divided in two steps. The first step is designed to the study of the relationship between the variables X and Y, leading to which typically is called the calibration curve. This step is based on a sample of n pairs $(x_1, y_1), \ldots, (x_n, y_n)$ from the distribution of Y given X. The second step is based on the prediction of an X value, denoted by x_0 given an observed value of Y, y_0, and the calibration curve. When the values $x_1, \ldots, x_n$ are fixed hand before , the problem is called controlled calibration. Otherwise, if they are a sample from a random variable, then the problem is called natural calibration. There exist a vast literature in calibration problems (Brown 1993), where the following assumptions are typically made: (i) linearity; (ii) normality; iii) independence and (iii) homocedasticity. Thus, the calibration model is given by

$$\begin{aligned} Y_i &= \alpha + \beta x_i + e_i, \quad i = 1, 2, ..., n \\ Y_0 &= \alpha + \beta x_0 + e_0, \end{aligned} \tag{1}$$

where, $e_0, e_1, \ldots e_n$ are independent and identically distributed according the $N(0, \sigma^2)$, that is, the normal distribution with mean 0 and variance σ^2.

Two estimators of x_0 are usually considered in the classical literature. The first is the maximum likelihood estimator, which we denote as $\hat{x}_c = \bar{x} + (y_0 - \bar{y})/\hat{\beta}$, where $\bar{x} = \sum_{i=1}^{n} x_i/n$, $\bar{y} = \sum_{i=1}^{n} y_i/n$ and $\hat{\beta}$ is the maximum likelihood estimator of β under the calibration model above. The second estimator, usually known as the inverse estimator, is given by $\hat{x}_I = \bar{x} + (y_0 - \bar{y})\hat{\phi}$, where $\hat{\phi}$ is the maximum likelihood estimator of ϕ, the regression coefficient considering the inverse regression of X in Y. Some problems with the classical estimator include infinite mean squared error and unlimited confidence intervals. On the other hand, the inverse estimator has no formal justification from the classical point of view, specially in the case of controlled calibration where X a constant, how to justify the regression model for X on Y?

A Bayesian solution to the calibration problem includes the specification of a model and, moreover, the specification of the prior distributions for the unknown parameters in the specified model. Within this formulation, the known quantities are $(x_1, y_1), \ldots, (x_n, y_n), y_0$ and the unknown quantities (parameters) are $\alpha, \beta, \sigma^2, x_0$. The assigment of prior distributions to the unknown quantities is justified within the subjectivistic view of probability, emphasizing that randomness is related to our uncertainty in the presence of the unknown.

One of the first Bayesian studies in calibration is the work of Hoadley (1970) . In that paper, it is shown by considering an improper prior distribution for $(\alpha, \beta, \sigma^2)$ and an appropriate Student-t prior for x_0, that an Student-t posterior distribution

is obtained for x_0, with posterior mean given by the inverse estimator, providing thus a Bayesian justification for the inverse estimator.

Nonnormal calibration models are not common in the literature, specially due to the fact that classical analysis of such models are very complex. In regression models, the Student-t distribution has been considered as a way of accomodating extreme (outlying) observations. Lange et al. (1989), propose using maximum likelihood estimation in such models. In this paper, a generalization is proposed to the normal calibration model by considering that

$$Y_i \mid x_i, \alpha, \beta, \sigma^2 \sim T_\nu(\alpha + \beta x_i, \sigma^2) \ , \ i = 0, 1, ..., n,$$

with $Y_0, Y_i, ..., Y_n$ conditionally independent given $x_0, x_1, ..., x_n$. The approach followed is Bayesian and the main interest is to estimate x_0 with α, β, σ^2 as nuisance parameters. In Section 2, the equivalence between the Student-t model and the normal heterochedastic calibration model is established. A brief review of the Gibbs sampler approach to posterior distribution estimation is considered in Section 3. In Section 4 a simulation study is performed to illustrate the performance of the approach with simulated populations. Finally, Section 5 is dedicated to the application of the approach to the marathon courses data considered in Smith and Corbett (1987) .

2. The approach

In this section it is considered the model described by equation (1) under the following assumptions:

1. The errors terms $e_0, e_1 \dots e_n$ given $\alpha, \beta, \sigma^2, x_0$ are independent and identically distributed according to the Student-t distribution with ν degrees of freedom and scale parameter σ^2.

2. The parameters $\beta^* = (\alpha, \beta, \sigma^2)$ are independent of x_0 with a joint prior distribution $\pi(\beta^*, \sigma^2)\pi(x_0)$.

The observed data is written as $y = (y_1, \dots, y_n)$, $y^* = (y_0, \dots, y_n)$ and $x = (x_1, \dots, x_n)$. Under the above asumptions it follows that the likelihood function corresponding to the observed sample is given by

$$L(\beta^*, \sigma^2, x_0) \propto \sigma^{-(n+1)} \prod_{i=1}^{n} \{1 + \frac{(y_i - \beta^* x_i^*)^2}{\nu\sigma^2}\}^{\frac{-(\nu+1)}{2}} \{1 + \frac{(y_0 - \beta^* x_0^*)^2}{\nu\sigma^2}\}^{\frac{-(\nu+1)}{2}}, \quad (2)$$

where $x_i^* = (1, x_i)^t$, $i = 0, 1, ..., n$. The above expression is associated with a series of dificulties in obtaining estimators or marginal posterior densities, which usually involve the use of complex numerical procedures. The derivation of a posterior distribution for x_0 involve, for example, integrating the above likelihood with respect to β^* and σ^2, which can be quite complex. In fact,

$$p(x_0 \mid y^*, x) \propto p(y^* \mid x, x_0)\pi(x_0 \mid x). \tag{3}$$

Now assumptions 1 and 2 imply that the predictive distribution in (3) is given by

$$\begin{aligned} p(y^* \mid x, x_0) &= \int_{\beta^*,\sigma^2} p(y^*, \beta^*, \sigma^2) \mid x, x_0) d\beta^* \sigma^2 \\ &= \int_{\beta^*,\sigma^2} p(y_0 \mid \beta^*, \sigma^2, x_0) p(y \mid \beta^*, \sigma^2, x) \pi(\beta^*, \sigma^2) d\beta^* \sigma^2 \\ &= \int_{\beta^*,\sigma^2} L(\beta^*, \sigma^2, x_0) \pi(\beta^*, \sigma^2) d\beta^* \sigma^2 . . \end{aligned} \tag{4}$$

The next section presents an alternative analysis to the calibration model with Student t-erros, by considering an heterochedastic variances model. The approach allows deriving explict expressions for the conditional posterior distributions of the parameters involved, which makes it simple to consider the Gibbs sampler to estimate the posterior density function of x_0.

2.1. The Student-t model and the heterochedastic model

The approach presented in the following is somewhat related to the work of Geweke (1993) , where the main interest is on estimating the regression line parameters and improper priors distributions are considered for the parameters in the models. The problems related to the derivation and existence of the posterior distribution of x_0 will be considered later. The main object of this section is to establish an equivalency between the Student-t model and an heterochedastic normal model described next. Let

$$Y_i = \alpha + \beta x_i + e_i.$$

Such that given β^*, σ^2, x_0 and $w^* = (w_0, \ldots, w_n)$, e_i are independent and distributed as the $N(0, w_i\sigma^2)$, $i = 0, 1, \ldots, n$. Note that new parameters $w_0, w_1, \ldots, w_n$, are introduced in the model and are a priori considered to be indepedent with distribution proportional to a chi-square distribution, that is, $\nu/w_i \sim \chi^2_\nu$, $i = 1, \ldots, n$ and independent of (β^*, σ^2, x_0). The posterior distribution for x_0 is obtained by computing the predictive distribution $p(y_0, y/x_0, x)$, which is proportional to

$$\int_{\beta^*} \int_{\sigma^2} \int_{w^*} \pi(\beta^*, \sigma^2) \sigma^{-(n+1)} \prod_{i=0}^{n} w_i^{\frac{-(\nu+3)}{2}} e^{-\sum_{i=0}^{n} \frac{\sigma^{-2}[y_i - \beta^* x_i^*]^2 + \nu}{2w_i}} dw^* d\sigma^2 d\beta^*.$$

With respect to the inner integral, it follows that

$$\int_{w_i} w_i^{\frac{-(\nu+3)}{2}} e^{-\sum_{i=0}^{n} \frac{\sigma^{-2}(y_i-\beta^* x_i^*)^2+\nu}{2w_i}} dw_i = \Gamma(\frac{\nu+1}{2})[\sigma^{-2}(y_i-\beta^* x_i^*)^2+\nu]^{\frac{-(\nu+1)}{2}},$$

$i = 0, 1, ..., n.$

Thus, the predictive distribution is proportional to

$$\int_{\beta^*}\int_{\sigma^2} \sigma^{-1}[\frac{(y_0-\beta^* x_0)^2}{\nu\sigma^2}+1]^{\frac{-(\nu+1)}{2}} \sigma^{-n}\prod_{i=1}^{n}[\frac{(y_i-\beta^* x_i^*)^2}{\nu\sigma^2}+1]^{\frac{-(\nu+1)}{2}}\pi(\sigma^2,\beta^*)d\sigma^2 d\beta^*.$$

Note that

$$\sigma^{-n}\prod_{i=1}^{n}[\frac{(y_i-\beta^* x_i^*)^2}{\nu\sigma^2}+1]^{\frac{-(\nu+1)}{2}}\pi(\sigma^2,\beta^*)$$

is the kernel of the posterior distribution of (β^*, σ^2) given $y_1, \ldots, y_n$ under a Student-t model. Moreover, simple algebraic manipulations show that $\sigma^{-1}[\frac{(y_0-\beta^* x_0)^2}{\nu\sigma^2}+1]^{\frac{-(\nu+1)}{2}}$ is the kernel of the conditional distribution of Y_0 given β^*, σ^2, x_0. Thus, the above expression is also the kernel of the predictive distribution if one considers the Student-t model, which shows the equivalence between the models.

The number of parameters increases substantially in the heterochedastic normal model case. In the original (Student-t) model the number of parameters is four. However, the number of parameters increases to $n+5$ in the heterochedastic model. This fact would make it more difficult to obtain the marginal posterior distribution by integrating on the nuisance parameters. However, since the approach to be considered is based on the Gibbs sampler, the heterochedastic model permits obtaining closed form expressions for the conditional distributions required for implementing the approach.

2.2. The conditional distributions

The conditional distributions obtained in the sequel are crucial for the implementation of the Gibbs sampler approach. The conditional posterior densities are denoted by

$$p(\beta^* \mid \sigma^2, w^*, x_0, y^*), \tag{5}$$

$$p(\sigma^2 \mid \beta^*, w^*, x_0, y^*), \tag{6}$$

$$p(w^* \mid \beta^*, \sigma^2, x_0, y^*), \tag{7}$$

and

$$p(x_0 \mid \beta^*, \sigma^2, w^*, y^*). \tag{8}$$

The first three can be obtained by using results in Geweke (1993) , who considered improper prior distributions for (β^*, σ^2). This follows since conditioning in x_0, a regression model in $n+1$ observations is obtained. We assume that the prior distribution of β^* is the bivariate normal distribution with mean vector b_0 and covariance matrix Σ_0 and, also that it ,is independent of σ^2 and x_0. Moreover, σ^2 and x_0 are independent with $\frac{\psi_0}{\sigma^2} \sim \chi^2_{r_0}$ and x_0 has a normal prior distribution with mean m_0 and variance v_0.

Thus, the kernel of the joint posterior density is given by

$$\prod_{i=0}^{n} w_i^{\frac{-(v+3)}{2}} \sigma^{\frac{-(n+r_0+3)}{2}} e^{\frac{-1}{2}\left[\frac{\psi_0}{\sigma^2} + \frac{(x_0-m_0)^2}{v_0} + \sum_{i=0}^{n} \frac{\sigma^{-2}(y_i-\beta^* x_i^*)^2+v}{w_i} + (\beta^*-b_0)\Sigma_0^{-1}(\beta^*-b_0)^t\right.}$$

The first conditional density considered is the density of β^* given σ^2, w, x_0, y, which is the bivariate normal distribution with mean b_1 and variance covariance matrix Σ_1, where

$$\begin{aligned} \Sigma_1 &= (\Sigma^{-1} + \Sigma_0^{-1})^{-1} \\ b_1 &= \Sigma_1^{-1}(\Sigma_0 b_0 + \Sigma b), \end{aligned}$$

$$\Sigma = \begin{bmatrix} w_h + \frac{\bar{x}(w^*)}{S_{xx}(w^*)} & \frac{-\bar{x}(w^*)}{S_{xx}(w^*} \\ \frac{-\bar{x}(w^*)}{S_{xx}(w^*)} & \frac{1}{S_{xx}(w^*)} \end{bmatrix},$$

with,

$S_{xx}(w^*) = \sum_{i=0}^{n} \frac{(x_i - \bar{x}(w^*))^2}{w_i}$, $S_{xy}(w^*) = \sum_{i=0}^{n} \frac{(x_i - \bar{x}(w^*))(y_i - \bar{y}(w^*))}{w_i}$

$\bar{x}(w^*) = (\sum_{i=0}^{n} \frac{x_i}{w_i}) w_h$, $\bar{y}(w^*) = (\sum_{i=0}^{n} \frac{y_i}{w_i}) w_h$ and $w_h = (\sum_{i=0}^{n} \frac{1}{w_i})^{-1}$.

Finally, $b = (\hat{\alpha}, \hat{\beta})$, where $\hat{\alpha} = \bar{y}(w^*) - \hat{\beta}\bar{x}(w^*)$, $\hat{\beta} = \frac{S_{xy}(w^*)}{S_{xx}(w^*)}$, are the weigthed least squares estimators.

The second conditional posterior considered is the density of $\sigma^2 \mid \beta, w^*, x_0, y^*$, which is $\psi_1 \chi_{r_1}^{-2}$, where $\psi_1 = \sum_{i=0}^{n} \frac{[y_i - \beta^* x_i^*]^2}{w_i} + \psi_0, r_1 = r_0 + n$ and $\chi_{r_1}^{-2}$ denotes a random variable distributed according to the inverse chi-square distribution with r_1 degrees of freedom.

The third distribution is the joint posterior distribution of w^*, with density given by

$$p(w^* \mid \beta^*, \sigma^2, x_0, y^*) = \prod_{i=0}^{n} p(w_i \mid \beta, \sigma^2, x_0, y^*),$$

where the marginal conditional density of w_i is such that

$$w_i|\beta^*, \sigma^2, x_0, y \sim S_i \chi^{-2}_{\nu+1},$$

with $S_i = (y_i - \alpha - \beta x_i)^2/\sigma^2 + \nu$, $i = 0, 1, \ldots, n$. Finally, the last conditional distribution, that of x_0 given $\beta^*, \sigma^2, w^*, y^*$ is normal with mean m_1 and variance v_1, where $m_1 = v_1[\frac{m_0}{v_0} + \frac{\beta(y_0-\alpha)}{\sigma^2 w_0}]$ and $v_1 = [v_0^{-1} + \frac{\beta^2}{\sigma^2 w_0}]^{-1}$.

3.The Gibbs sampler

The main idea behind Markov Chain Monte Carlo (MCMC) is to build up a Markovian process whose stationary distribution (with density f) is the one of interest. The process is then iterated for a suficiently large time t and a sample of size m of this process is then a sample of f. Typically, two ways of selecting this sample can be considered. The first considers the last value generated form the Markovian process as the first element of the sample. The Markovian process is then (independently) reiterated until m such sequences are obtained and the last element of each sequence is considered, forming then a random (iid) sample of size m from f. Another approach consists in obtaining a large sequence from the Markovian process (say, 30000), disregarding part of this sequence (corresponding to process convergence) and for the remaining part, take one from each 10-th observation, to obtain an independent sample. Among the MCMC methods, the most well known is the Gibbs sampler, introduced in Bayesian inference by Geman and Geman (1984) , when studing problems related to image processing. It became more popular in Bayesian inference after the paper by Smith and Gelfand (1990) , where several applications of the approach are considered. For a posterior density f depending on k variables $X_1, \ldots, X_k$, the process is built up on the conditional densities

$$p(X_j|X_i; i \neq j = 1, \ldots, n),$$

according to the following scheme. Giving starting values $x_1^0, \ldots, x_k^0$, generate a value x_1^1 from the distribution with density $p(X_1|X_2 = x_2^0, \ldots, X_k = x_k^0)$, generate a value x_2^1 from the density $p(X_2|X_1 = x_1^1, X_3 = x_3^0, \ldots, X_n = x_k^0)$ and so on until a value x_k^1 is generated from the density $p(X_k|X_1 = x_1^1, \ldots, X_{k-1} = x_{k-1})$. This completes the first cycle of the sampler. The process then goes to a second cycle, with $(x_1^1, \ldots, x_k^1)$ as starting values. After t of such cycles, as sample $(x_1^t, \ldots, x_k^t)$ of the density f is obtained. Following this process, by reiterating the process m times a sample of size m is obtained from f. From this sample, the marginal densities of the variables X_j can be estimated by constructing histograms, for example. Moreover, estimates of the posterior mean, variance and quartils can be easily obtained by computing the

corresponding sample mean, variance and quartils, that is, all the inference can be based on descriptive aspects of the sample. Another possibilty is to compute the following estimate for the density of X_j:

$$\hat{f}_{X_j}(x_j) = \frac{1}{m}\sum_{t=1}^{m} p(x_j|x_1^t, \ldots, x_{j-1}^t, x_{j+1}^t, \ldots, x_k^t). \tag{9}$$

As emphasized in Gelfand and Smith (1990), this density typically presents better inferences then simply considering descriptive inferences from the marginal posterior densities. The estimators based on (9) is known as the Rao-Blackwellized version of the histogram, having smaller mean squared error than the histogram as an estimator of f. The above algorithm can be implemented in any computer language, as for example, FORTRAN or C. However, in some less complex situations, a software specific to Gibbs sampling implementation developed by Spieglharter et al. (1994) is available and can be obtained by using the appropriate *ftp* address.

4. A simulation study

In this section results of a simulation study are presented to illustrate the behavior of the approach considered in the previous sections to estimated x_0, recalling that no explicit expression is obtained for its posterior distribution. Fixed values were considered for $x = (x_1, \ldots, x_n)$ so that $|x_i - x_{i-1}| = 1$ and $\bar{x} = 0$ and a value of x_0 from which simulated samples were simulated for $(y_0, y_1, \ldots, y_n)$, considering a Student-t distribution with ν degrees of freedom with location and scale parameters given, respectively, by βx_i and τ. The values presented in Table 1 represent 0.90 credibility regions for the posterior mean of x_0.

TABLE 1 : Credibility regions for several values of β
($n = 21, x_0 = 0.5, \tau = 1$)

β	Degrees of freedom			
	2	5	10	30
0.5	(-8.657,17.68)	(-4.942,6.184)	(-3.049, 5.75)	(-2.929,4.174)
	4.461	0.6653	1.332	0.6398
1	(1.129,4.567)	(-3.03,2.23)	(-0.03616,4.16)	(-0.3528,2.5)
	2.851	-0.3512	2.014	1.09
3	(-0.1953,1.821)	(-0.1979,1.454)	(0.2835,1.656)	(-0.5092,0.7082)
	0.8388	0.6199	0.9702	0.09634

Table 1 illustrates the fact that by increasing the degrees of freedom parameter the credibility regions become more precise, a result somewhat expected. Moreover, the table also illustrates the fact that there is a strong influence of the values of β on the precision of the credibility regions. For example, for $\nu = 30$ the interval sizes go from 7 (for $\beta = 0.5$) to 1.2 (for $\beta = 3$). This result illustrates the fact that predicton becomes more precise as the regression line becomes steeper. Thus, uncertainty about x_0 decreases as β increases and some caution is recommended when the value of β is small.

TABLE 2 : Credibility regions for different sample sizes
($x_0 = 0.5, \beta = 1, \tau = 1$)

n	Degrees of Freedom			
	2	5	10	30
7	(-4.96,8.26)	(-1.57,5.39)	(-1.545,1.729)	(-3.48,2.86)
	0.8828	1.541	0.07227	1.200
21	(1.129,4.567)	(-3.03,2.23)	(-0.03616,4.16)	(-0.3528,2.5)
	2.851	-0.3512	2.014	1.09
51	(-1.961,3.884)	(-1.237,2.89)	(-1.739,1.682)	(-2.197,0.7855)
	0.8238	0.8112	-0.05913	-0.7048

TABLE 3: Credibility regions for different values of x_0
(n=51 , β=1 , var=1)

x_0	Degrees of Freedom			
	2	5	10	30
centre	(-1.961,3.884)	(-1.04,2.863)	(-1.739,1.682)	(-2.197,0.7855)
(0.5)	0.8238	0.8112	-0.05913	-0.7048
border	(22.43,30)	(17.83,22.11)	(20.41,24.01)	(19.87,22.95)
(22.5)	25.97	19.95	22.2	21.42
outside	(26.2,33.77)	(28.37,32.3)	(27.34,31.02)	(27.4,30.52)
(30)	29.72	30.3	29.13	28.98

Tables 2 and 3 illustrate, respectively, the behavior of the credibility regions for varying sample sizes (n) and different positions for the value x_0 being estimated. As expected, the precison of the estimates increase by increasing the sample sizes. Moreover, as indicated by Table 3, the position of the value being estimated has practically no influence on the precision of the estimates.

5. An application to measuring marathon courses

In this application, the data set on marathon courses studied by Smith and Corbett (1987) is reanalysed. The main part of the course of interest is divided into 25 intervals, including the baselines (standard distances), whith lengths precisely known. Ordinary bicycles with revolution counters which records accurately the number of revolutions of the front weel, are then ridden over the standard (baseline) distances to calibrate the counter. The standard distance is measured by surveyor's stell tape or by eletronic distance meter. The bicycles are then ridden over the route to be measured. The following assumptions are considered:

$$Y_{ij} = \beta_j Z_i + e_{ij1},$$

with $e_{ij1} \sim N(0, s_{ij}^{-1}\sigma^2)$, and

$$Y_{kj2} = \beta_j X_k + e_{kj2},$$

with $e_{kj2} \sim N(0, s_{kj}^{-1}\sigma^2)$, where Y_{ij1} and Y_{kj2} are the counts obtained by measurer j on the i-th calibration interval, the k-th course interval, respectively. The true (known) length of the i-th calibration interval is denoted by Z_i and the unknown length of the k-th interval is denoted by X_k. The main interest is on estimating the total distance $T = \sum_{k=1}^{13} X_k = 13\bar{X}$, by considering that

$$\bar{Y}_{j2} = \beta_j \bar{X} + e_{j2}^*,$$

where $e_{j2}^* \sim N(0, \bar{s_j}^{-1}\sigma^2)$, with $s_{ij}^{-1} = w_i$ (unknown) and $\bar{s}_j = w^*$, which also is unknown, and $w = (w_1, \ldots, w_n)'$.

The Bayesian analysis of the above model was also implemented by using the BUGS program. The main difference between this situation and the one considered in the previous sections is that $\beta = (\beta_1, \ldots, \beta_{13})^t$ where β_j is associated to the j-th cyclist. The parameters $\beta_1, \ldots, \beta_{13}$ are considered to be a priori indepedendent with $N(0; 10^4)$ distribution, which will have little effect on the posterior distribution. Similarly, we consider $\bar{X} \sim N(1000; 10^6)$. Moreover, independent inverse gamma prior distributions with parameter ν_1 are considered for the parameters $w_1, \ldots, w_{12}$, and an independent inverse gamma prior distribution with parameter ν_2 is considered for w^*.

Figure 1, which represents a typicall output from the BUGS program, presents posterior densities for the parameters of the model, where *beta.bar* denotes the mean of $\beta_1, \ldots, \beta_{13}$, τ denotes $1/\sigma^2$, $wN = w^*$ and $xN = \bar{X}$. Further, we took $\nu = 30$, $t = 500$ and $m = 1000$ (the size of the chain).

Figure 1. Posterior (Gibbs) densities

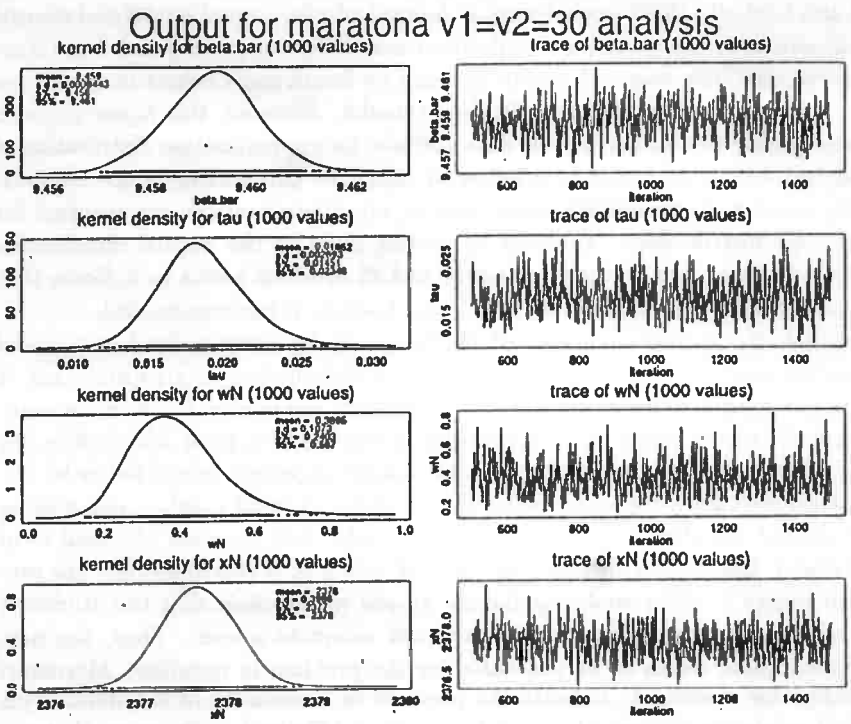


Table 4 shown below presents values of the posterior mean and standard deviation of $\hat{T}_N$, the total (unknown) length of the course to be estimated for several values of ν, with $\nu = \nu_1 = \nu_2$.

Table 4. Posterior mean and standard deviation for T_N

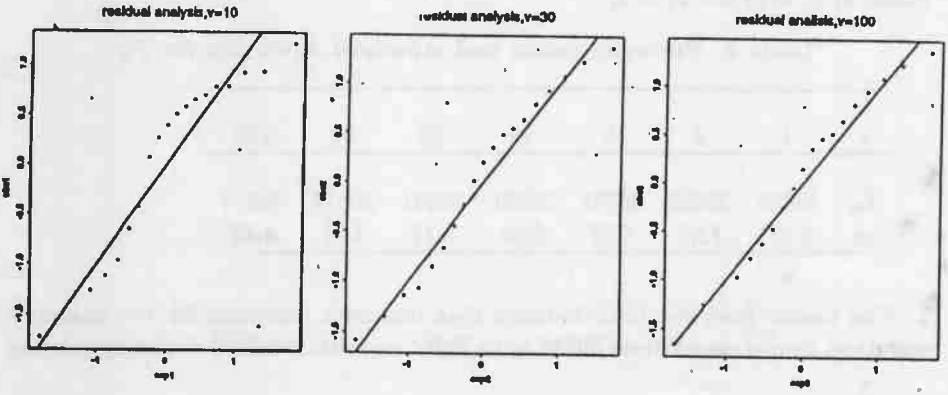
ν	3	4	5	10	30	60	100
$\hat{T}_N$	30888	30888	30901	30901	30901	30914	30914
sd	8.31	7.96	7.57	6.94	5.17	4.81	4.42

The results from the table indicate that the point estimates for the unknown marathon course ranges from 30888 m to 30914 m, with standard deviation ranging

from 8.31 m ($\nu = 3$) to 4.42 m ($\nu = 100$). The corresponding values obtained in Smith and Corbett (1987) considering an homochedastic normal model and normal approximations to the posterior distributions were 30904 m (mean) and 2 m (standard deviation). The standard deviation found by Smith and Corbett is small than the one obtained by considering a Student-t model. However, this figure depends on the normality assumption for the data and also for approximation distribution of T_N. We recall that the values that follow by using the Gibbs sampler are obtained from the exact posterior distributions, that is, no approximations are required for the posterior distributions. However by getting close so the normal distribution ($\nu = 100$), the standard deviation observed is 4.42 m which seems to indicate that the standard deviation obtained by Smith and Corbett is underestimated.

Although the approach considered in this paper by using a Student-t model is somewhat general since it includes the normal distribution as a special case, it requires the specification of the degrees of freedom parameter ν. A more complete analysis would follow by incorporating in the model a prior distribution for this parameter. However, an immediate and simpler approach would follow by doing exploratory data analysis based on $Q - Q$-plots. Several graphis representing $Q - Q$-plots of the observed residual from the model with the ones obtained form the Student-t distribution theoreeal for several values of ν (ν=10,30,100) are presented in Figure 2. After analysing the the graphs we conclude that the Student-t model with the largest value for ν is the most adequate model. Thus, the normality assumption seems to be justifiable for the problem in question. Moreover, small values for ν seems to indicate the presence of bimodality in the data. This somewhat reinforces the remarks made by Smith and Corbett in the sense that the measurements obtained by the cyclists in the morning and in the afternoon could be analyzed separately.

Figure 2. $Q - Q$-plots (t-model x residual)



6. Conclusions

One consequence of considering de Student-t model is the specification of the number of degrees of freedom, ν. It is often the case that an exact value for this quantity is not available. One way to contouring this difficulty is to assigning a prior distribution for ν as, for example, the exponential distribution. As is well known with this distribution, most of its mass is concentraded close to the orign, which would be adequate in situations where the normality assumption is questionable. Considering the exponential prior distribution with parameter λ for the degrees of freedom ν, it is necessary to specify the conditional density $p(\nu \mid \beta^*, \sigma^2, x_0, y)$. It follows that this distributions is proportional to $(\frac{\nu}{2})^{\frac{n\nu}{2}} e^{-\eta\nu}$, where $\eta = \frac{1}{2}\sum_{i=0}^{n}(log w_i + w_i^{-1}) + \lambda$. Since there is no closed form for the condicional density of ν given $\beta^*, \sigma^2, x_0, y$, the Gibbs sampler is not adequate and another algorithm should be used. Geweke (1993) considers the problem of the existence of the joint posterior density of (β^*, σ^2). This is due to consideration of a uniform priori distribuitions for those parameters. We recall that uniform priors are used to express the intention of impartial preference over the parameter space, making the Bayesian inference essentialy objetive rather than subjective. The difficulty in specifying proper prior distributions for β^* is another argument in favor of noinformative priors. This fact is evident with the maraton data. However even in this example, informative prior distributions can be considered with larger variances, denoting uncertainty in the especification of prior values for the parameters of the model. Analytically, it may be interesting to study the limiting case, that is, the prior variances go to infinity, leading to improper priors. The study reported by Geweke (1993) on the existence of the posterior distibutions for the parameters of the line with improper priors seems to be incorret and we haven't yet found ways of correcting those results, which will be the subject of incoming research.

Calibracion Bayesiana con errores t-Student

Resumen

En este trabajo consideramos el problema de calibración en el modelo de regresión suponiendo que los terminos de error son independientes con distribución comun t-Student. El problema es abordado dentro de la perspectiva Bayesiana. Se muestra que el problema puede ser formulado de manera equivalente, utilizando un modelo de regresion heterocedastico con distribucion a priori apropiada para las varianzas. Esta formulación alternativa facilita la aplicación del muestreo de Gibbs para obtener aproximaciones de las distribuciones a posteriori. Los resultados son ilustrados con datos obtenidos de simulaciones y los analizados en Smith y Corbett (1987), para estimar el recorrido de una Maraton.

References

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

Geman, S. and Geman, D. (1984). Stochastic relaxation Gibbs distributions, and the bayesian restoration of imagens, *IEEE* **PAMI-6**: 721–741.

Geweke, J. (1993). Bayesian treatment of the independent student-t linear model, *Journal Appl.Econometrics* **8**: 519–540.

Hoadley, B. (1970). A bayesian look at inverse linear regression, *JASA* **65**: 356–369.

Smith, A. and Gelfand, A. (1990). Sampling-based aproaches to calculating marginal densities, *JASA* **85**: 398–409.

Smith, R. and Corbett, M. (1987). Measuring marathon courses:an application of statistical calibration theory, *Appl.Stat.* **36**: 283–295.

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

9601 - MENTZ, R.P.; MORETTIN, P.A. and TOLOI, C.M.C. Bias correction for estimators of the residual variance in the ARMA (1,1) Model. São Paulo, IME-USP, 1996. 21p. (RT-MAE-9601)

9602 - SINGER, J.M.; PERES, C.A.; HARLE, C.E. Performance of Wald's test for the Hardy-Weiberg equilibrium with fixed sample sizes. São Paulo, IME-USP, 1996.15p. (RT-MAE-9602).

9603 - SINGER, J.M. and E. SUYAMA, E. Dispersion structure, Hierarchical models, Random effects models, Repeated measures. São Paulo, IME-USP, 1996. 21p. (RT-MAE-9603).

9604 - LIMA, A.C.P. and SEN, P.K. A Matrix-Valued Counting Process with First-Order Interactive Intensities. São Paulo, IME-USP, 1996, 25p. (RT-MAE-9604)

9605 - BOTTER, D.A. and SINGER, J.M. Experimentos com Intercâmbio de Dois Tratamentos e Dois Períodos: Estratégias para Análise e Aspectos Computacionais, São Paulo, IME-USP, 1996. 18p. (RT-MAE-9605)

9606- MORETTIN, P.A. From Fourier to Wavelet Analysis of Time Series. São Paulo, IME-USP, 1996. 11p. (RT-MAE-9606)

9607 - SINGER, J.M.; HO, L.L. Regression Models for Bivariate Counts. São Paulo, IME-USP, 1996. 20p. (RT-MAE-9607)

9608 - CORDEIRO, G.M. and FERRARI, S.L.P. A method of mements for finding Bartlett-type corrections. São Paulo, IME-USP, 1996. 11p. (RT-MAE-9608)

9609 - VILCA-LABRA, F.; ARELLANO-VALLE, R.B.; BOLFARINE, H. Elliptical functional models. São Paulo, IME-USP, 1996. 20p. (RT-MAE-9609)

9610 - BELITSKY, V.; FERRARI, P.A.; KONNO, N. A Refinement of Harris-FKG Inequality for Oriented Percolation. São Paulo, IME-USP, 1996. 13p. (RT-MAE-9610)

9611 - GIMENEZ, P.; BOLFARINE, H. Unbiased Score Functions in Error-In-Variables Models. São Paulo, IME-USP, 1996. 23p. (RT-MAE-9611)

9612 - BUENO, V.C. Comparing component redundancy allocation in K-out-of-n system. IME-USP, 1996. 10p. (RT-MAE-9612)

9613 - GALEA, M.; PAULA, G.A.; BOLFARINE, H. Local influence in elliptical linear regression models. IME-USP, 1996. 14p. (RT-MAE-9613)

9614 - LIMA, C.R.; BOLFARINE, H.; SANDOVAL, M.C. Linear calibration in multiplicative measurement error models. IME-USP, 1996. 12p. (RT-MAE-9614)

9615 - AVERBACH, M.; CUTAIT, R.; WECHSLER, S.; CORRÊA, P.; BORGES, J.L.A. Aplicação da inferência bayesiana no estudo das probabilidades de diagnóstico, por colonoscopia, das afecções colorretais em portadores da síndrome da imunodeficiência adquirida com diarréia. IME-USP, 1996. 10p. (RT-MAE-9615)

9616 - YOSHIDA, O.S.; LEITE, J.G.; BOLFARINE, H. Inferência bayesiana do número de espécies de uma população. IME-USP, 1996. 33p. (RT-MAE-9616)

9617 - CARMONA, S.C.; TANAKA, N.I. Exponential estimates for "Not Very Large Deviations" and ware front propagation for a class of reaction-diffusion equations. IME-USP, 1996. 30p. (RT-MAE-9617)

9618 - IRONY, T.Z.; PEREIRA, C.A.B.; TIWARI, R.C. On The Comparison Between Two Correlated Proportions in 2 x 2 Tables. IME-USP, 1996. 15p. (RT-MAE-9618)

9619 - ANDRÉ, C.D.S.; ELIAN, S.N.; NARULA, S.C.; AUBIN, E.C. Stepwise procedures for selecting variables in the minimum sum of absolute errors regression. IME-USP, 1996.9p. (RT-MAE-9619)

9620 - ANDRÉ-C.D.S.; NARULA, S.C.; PERES, C.A.P.; VENTURA.G.A. Asymptotic properties of the MSAE estimators in the dose-response model. IME-USP, 1996. 9p. (RT-MAE-9620)

9621 - FERRARI, S.L.P.; CRIBARI-NETO, F. On bootstrap and analytical bias corrections. IME-USP, 1996. 9p. (RT-MAE-9621)

9622 - NARULA, S.C.; SALDIVA, P.H.N.; ANDRÉ, C.D.S.; ELIAN, S.N.; AUBIN, E.C.Q. MSAE regression: a more robust alternative to the least squares regression. IME-USP, 1996. 10p. (RT-MAE-9622)

The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05389-970 - São Paulo, Brasil