APPROXIMATE BAYESIAN METHODS FOR LOGISTIC REGRESSION MODEL

Marcos Alves dos SANTOS[1]
Fernando Antonio MOALA[1]
Vilma Mayumi TACHIBANA[1]

▪ *ABSTRACT: In this paper we consider the Bayesian inference for the estimation of logistic regression model. The paper presents two approximate methods, Laplace's method (Tierney and Kadane, 1986) and Markov chain Monte Carlo (MCMC) to obtain the marginal posterior density for the parameters. A comparison of these methods is carried out for the prediction of the in-hospital death in patients with acute myocardial infarction in a particular hospital. Besides, we identify the risk factors that characterize the myocardial infarction. The Bayesian results are also compared with maximum likelihood estimation.*

▪ *KEYWORDS: Logistic regression; Bayesian inference; noninformative prior; Laplace's method; MCMC.*

1 Introduction

Logistic regression analysis is a standard method for building prediction models for a binary outcome and has been especially popular with medical research in which the dependent variable is whether or not a patient has a disease.

The Bayesian inference is an attractive framework for estimation of logistic regression parameters and has grown in popularity in recent years. Besides offering easier interpretability of parameter estimation it gives more reliable results for smaller samples. However, to obtain the posterior distributions for the logistic

---

[1]Departamento de Matemática e Estatística, Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista – UNESP, CEP:19060-900, Presidente Prudente, SP, Brasil. E-mail: *socramaz@yahoo.com.br* / *femoala@fct.unesp.br* / *vilma@fct.unesp.br*

regression parameters the approximation methods are indispensable in the analysis because we are not able to analytically specify the marginal posterior densities. In this case, we can use Markov chain Monte Carlo (MCMC) approaches or Laplace's method (Tierney and Kadane, 1986) to make the Bayesian posterior inference. In this paper a comparison of these methods is carried out for the prediction of the in-hospital death in patients with acute myocardial infarction.

To progress with the Bayesian analysis it is necessary to specify a joint prior distribution over the parameter space. Thus, we assume independent normal prior distribution with mean zero and low precision for the parameters avoiding any complaint about the specification of subjective beliefs about the death prediction in hospital.

The paper is organized as follows. Section 2 introduces the logistic regression model and on how Bayesian framework can be used to address the logistic regression. Section 3 briefly presents an overview of the two approximate approaches used in this paper: Laplace's method and MCMC. The inference is illustrated for the in-hospital death in patients with acute myocardial infarction data set and the results are presented in Section 4. Finally, the conclusions are reported in Section 5.

2 The Bayesian logistic regression model

In logistic regression, a single outcome variable y_i ($i = 1, \ldots, n$) follows a Bernoulli probability function that takes on the value 1 with probability π_i and 0 with probability $1 - \pi_i$. Thus π_i varies over the observations as an inverse logistic function of a vector $\boldsymbol{x}_i$ which includes a constant and $k-1$ explanatory variables. For a sample of n observations y_i, $i = 1, \ldots, n$ the logistic regression model for binary data is specified by

$$y_i \,|\, \pi_i \sim Bernoulli(\pi_i), \tag{1}$$

$$\pi_i = P\{y_i = 1\} = \frac{exp(\boldsymbol{x}_i^t \beta)}{1 + exp(\boldsymbol{x}_i^t \beta)}, i = 1, \ldots, n, \tag{2}$$

where $y_i = 1$ if the interest response is observed for the i-th individual and $y_i = 0$ otherwise, $\beta = (\beta_1, \beta_2, \ldots, \beta_k)^t$ is the vector of unknown parameters and $\boldsymbol{x}_i^t = (1, x_{i1}, x_{i2}, \ldots, x_{ik})^t$ is the vector of covariable values for the i-th individual. The likelihood function is given by

$$L(\beta| \,\boldsymbol{y}, \boldsymbol{X}) = \prod_{i=1}^{n} \Big(\frac{exp(\boldsymbol{x}_i^t \beta)}{1 + exp(\boldsymbol{x}_i^t \beta)} \Big)^{y_i} \Big(\frac{1}{1 + exp(\boldsymbol{x}_i^t \beta)} \Big)^{1-y_i}. \tag{3}$$

The maximum likelihood estimate $\widehat{\beta}$ is obtained by maximizing the logarithm of the likelihood function (3). This estimate is derived by Newton-Raphson algorithm. When using maximum likelihood estimation, inferences about the model are usually based on asymptotic theory. The asymptotic distribution of $\widehat{\beta}$ is given by

$$\widehat{\beta} \sim N_{k+1}(\beta, I(\widehat{\beta})^{-1}) \tag{4}$$

where $I(\widehat{\beta})$ is the Fisher information matrix evaluated for $\widehat{\beta}$. Jennings (1986) points out that confidence regions can be found through likelihood function instead of asymptotic distribution of $\widehat{\beta}$.

Methods based on likelihood estimation are often used in practical researches, however an argument against the likelihood estimation refers to its properties in small samples. It is well known that the maximum-likelihood estimate for logistic regression model is biased for small samples (see Griffiths, 1973), which is the case for many epidemiological and clinical studies, hence the bias could be relevant for parameter estimation.

In a Bayesian framework the inference is based on the information provided by the posterior distribution of parameters β given by

$$p(\beta|\, \boldsymbol{y}, \boldsymbol{X}) = cf(\beta) \prod_{i=1}^{n} \Big(\frac{exp(\boldsymbol{x}_i^t \beta)}{1 + exp(\boldsymbol{x}_i^t \beta)} \Big)^{y_i} \Big(\frac{1}{1 + exp(\boldsymbol{x}_i^t \beta)} \Big)^{1-y_i} \tag{5}$$

where $f(\beta)$ is a prior distribution and c is the normalizing constant.

If no prior information on the model parameters exists or it is difficult to elicit or formalize, then initial uncertainty about the parameters can be quantified with a noninformative prior distribution. This is the same to include in the analysis just the information provided by the data.

Souza and Migon (2004) and Migon and Tachibana (1997) proposed independent normal priors, with extremely small precisions, for the components of β. These priors are equivalent to noninformative priors on these parameter, with all parameter values treated as equally plausible, and they are given by

$$\beta_j \sim N(0, 10000)\ , j = 1, ..., k. \tag{6}$$

3 Approximate methods in Bayesian inference

Clearly (5) is a complex function of the parameters and numerical methods are needed to compute marginal posterior, posterior moments and predictive densities for each of the model parameters. Approximations can be obtained via Laplace methods (Tierney and Kadane, 1986) or methods based on Markov Chain Monte Carlo (MCMC). For high-dimensional parametrical spaces, the application of numerical integration can be impracticable or can perform poorly.

3.1 The Laplace approximation

Tierney and Kadane (1986) presented a simple second-order approximation for posterior expectations of positive functions. It is based on the 2^{nd} order Taylor expansion applied to the exponential term of the integral. We describe here Laplace's method to approximate marginal posterior densities of individual parameters in multiparameter settings. Suppose $\Theta = \Theta_1 \times \Theta_2$ and we are interested in the marginal density of β_1,

$$p(\beta_1 \,|\, \boldsymbol{y}, \boldsymbol{X}) = \frac{\int_{\Theta_2} e^{nh^*(\beta_1,\, \boldsymbol{\beta}_2)} d\boldsymbol{\beta_2}}{\int_{\Theta} e^{h(\boldsymbol{\beta})} d\boldsymbol{\beta}} \tag{7}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k) = (\beta_1, \boldsymbol{\beta}_2)$ can be partitioned into two components, the first a scalar, in that $\beta_1 = \beta_i$, $i = 1, \ldots, k$, and the second a $k-1$ vector of the remaining components, $\boldsymbol{\beta}_2 = (\beta_1, \ldots, \beta_{i-1}, \beta_{i+1}, \ldots, \beta_k)$.

Suppose $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_1, \widehat{\boldsymbol{\beta}}_2)$ maximizes $h(\boldsymbol{\beta})$ and let Σ be minus the inverse of the Hessian of $h(\boldsymbol{\beta})$ at $\widehat{\boldsymbol{\beta}}$; thus Σ is a $k \times k$ matrix. For a given β_1, let the $(k-1)$ vector $\widehat{\boldsymbol{\beta}}_2^* = \widehat{\boldsymbol{\beta}}_2^*(\beta_1)$ maximize the function $h^*(\cdot) = h(\beta_1, \cdot)$, the function h with β_1 held fixed, and let $\Sigma^* = \Sigma^*(\beta_1)$ be minus the inverse of the Hessian of $h(\cdot)$, a $(k-1) \times (k-1)$ matrix. Applying Laplace's method to the integrals in the numerator and denominator of the marginal posterior density (7), we obtain the approximation

$$\widehat{p}(\beta_1 \,|\, \boldsymbol{y}, \boldsymbol{X}) = \left(\frac{det\, \Sigma^*(\beta_1)}{2\pi\, n\ det\, \Sigma}\right)^{\frac{1}{2}} exp\{n[h^*(\beta_1, \widehat{\boldsymbol{\beta}}_2^*) - h(\beta_1, \widehat{\boldsymbol{\beta}}_2)]\} \tag{8}$$

with

$$\Sigma* = -\left[\frac{\partial^2 h^*(\beta_2)}{\partial \beta_2^t \partial \beta_2}\right]^{-1}\Bigg|_{\beta_2 = \widehat{\beta}_2^*} \qquad and \qquad \Sigma = -\left[\frac{\partial^2 h(\beta)}{\partial \beta^t \partial \beta}\right]^{-1}\Bigg|_{\beta = \widehat{\beta}}. \tag{9}$$

Now, by assuming the logistic regression model (1) and a multivariate normal prior, with parameters μ and V, let X be the $n \times (p+1)$ input matrix, p is the n-vector of fitted probabilities with i-th element $\pi_i = \frac{exp(\boldsymbol{x}_i^t \beta)}{1+exp(\boldsymbol{x}_i^t \beta)}$ and W is an $n \times n$ diagonal matrix with i-th element $\pi_i(1-\pi_i)$ then

$$nh(\boldsymbol{\beta}) = -\frac{k}{2} log(2\pi) - \frac{1}{2} log|V| - \frac{1}{2}(\beta - \mu)^t V^{-1}(\beta - \mu) + \sum_{i=1}^{n} \boldsymbol{x}_i^t \beta y_i - \sum_{i=1}^{n} log(1 + e^{x_i^t \beta}), \tag{10}$$

$$\frac{\partial}{\partial \beta} h(\beta) = -V^{-1}(\beta - \mu) + X^t(y - p), \tag{11}$$

and

$$\frac{\partial^2}{\partial \beta^t \partial \beta} h(\beta) = -V^{-1} - X^t W X. \tag{12}$$

If we are interested in an approximation to evaluate the posterior mean of an arbitrary function $g(\beta)$, i.e,

$$E[g(\beta)|\, \boldsymbol{y}, \boldsymbol{X}] = \frac{\int g(\beta)p(\beta \,|\, D)d\beta}{\int p(\beta \,|\, D)d\beta} \tag{13}$$

that can be written as

$$E[g(\beta)|\, \boldsymbol{y}, \boldsymbol{X}] = \frac{\int b_N(\beta)e^{-nh_N(\beta)}d\beta}{\int b_D(\beta)e^{-nh_D(\beta)}d\beta}. \tag{14}$$

Tierney, Kass and Kadane (1989) proposed the approximation given by

$$E[g(\beta)|\, \boldsymbol{y}, \boldsymbol{X}] = \sqrt{\frac{det\,\Sigma_N}{det\,\Sigma_D}}\frac{b_N(\widehat{\beta}_N)}{b_D(\widehat{\beta}_D)}exp\{-n[h_N(\widehat{\beta}_N) - h_D(\widehat{\beta}_{\boldsymbol{D}})]\}, \tag{15}$$

where β_t is the maximum of $-h_t$ and $\Sigma_t = [D^2 h_t(\widehat{\beta}_t)]^{-1}$ for $t = N,\, D$.

The approximation will be very accurate when the posterior distribution is close to the normal. However, in case this does not occur, it will be necessary to choose a good parameterization, especially for small sample size, to obtain good accuracy.

3.2 The Metropolis Hastings algorithm

MCMC methods are classes of algorithms for sampling from probability distributions based on constructing a Markov chain. This MCMC sampling can be used to approximate the distribution (i.e, to generate a histogram), or to compute an integral (such as an expected value or normalization constant).

The MCMC approach draws samples from the required distribution by running a suitably constructed Markov chain for a long time.

Let $g(\phi)$ be the distribution of interest. Suppose at time t, ϕ_{t+1} is chosen by first sampling a candidate point η from a proposal distribution $q(\,\cdot \mid \phi_t)$. The candidate η is then accepted with probability

$$\alpha(\phi, \eta) = min\Big(1,\, \frac{g(\eta)q(\phi_t \mid \eta)}{g(\phi)q(\eta \mid \phi_t)}\Big). \tag{16}$$

If the candidate point is accepted, the next state becomes $\phi_{t+1} = \eta$. If it is rejected, then $\phi_{t+1} = \phi_t$ and the chain does not move. The proposal distribution is arbitrary, and, provided the chain is irreducible and aperiodic, the equilibrium distribution of the chain will be $g(\phi)$.

For more details of MCMC in a variety of models see, for example, Smith and Roberts (1993), Gelfand and Smith (1990) and Gilks et al. (1993).

We now describe our particular MCMC implementation used in the logistic regression framework. We generate a posterior Monte Carlo sample by simulating a Markov chain described as follows:

i) choose starting values $\beta^o = (\beta_0^o, \beta_1^o, \ldots \beta_k^o)$ a reasonable distance away from the posterior mode.

ii) sample a new value β^{i+1} from proposal normal distribution $N_{k+1}(\beta^i; I_F(\beta^i))$ where $I_F(\beta^i)$ is the inverse Fisher matrix of logarithm of posterior evaluate for β^i;

iii) thus, the chain is run for 25000 iterations and the first 5000 runs are discarded.

iv) after obtaining a random sample from the MCMC algorithm for each component of β, it is important to investigate issues such as convergence and mixing, to determine whether the sample can reasonably be treated as a set of random realizations from the target posterior distribution. Looking at marginal trace plots is the simplest way to examine the output besides formal procedures.

The MCMC output is shown in Figure 1 of Section 4.

After the convergence, all the values of the chain have marginal distribution given by the equilibrium distribution.

The Metropolis–Hastings method has the advantage of being easy to implement. An important characteristic of this algorithm is that, practically there are no restrictions on the posterior distribution. However, care is required when choosing a proposal distribution to ensure that the chain mixes well. Some tuning of the proposed distribution may be required.

4 Numerical illustration

In this section we present an application of Bayesian logistic regression in the prediction of in-hospital death for patients with myocardial infarction. Laplace and Metropolis-Hastings methods are considered in this application to specify the marginal posterior densities for the parameters of the logistics regression model.

The data set consist on the observations of clinical variables registered in the patients' admission with myocardial infarction, interned at the Pró-Cardíaco hospital of Rio de Janeiro, in the period of January 1991 to December 1996. The sample is composed by 599 patients, with 81 in hospital deaths. These data were used initially by Tachibana (1995) and updated by Souza (1999).

Five explanatory variables age, sex, has, iamp, kient and two interactions sex&has, iamp&kient were selected by forward stepwise regression to include in the model, representing known or potential risk factors (see Table 1 for details).

Thus, the Metropolis-Hastings sampling is running and Figure 1 illustrates the behavior of the chains along the iterations for the logistics regression parameters.

The logistic regression model was fitted by using the software R with

Table 1 - Characterization of the clinical variables in study

Variable	Identification	Codification
Hospital death	obith	no= 0, yes= 1
Age	age	years
Sex	sex	female= 0, male= 1
Arterial hypertension	has	absence= 0, presence= 1
Previous myocardial infarction	iamp	no= 0, yes= 1
Diabetes	diab	absence= 0, presence= 1
Smoke	smoke	no smoker= 0, smoker= 1
	parea	no= 0, yes= 1
Location	parei	no= 0, yes= 1
	pare1	no= 0, yes= 1
	pares	presence= 0, absence= 1
Killip	kient	1, 2, 3, 4

the MCMC algorithm generated by the function "MCMClogit". The candidate generating density is a multivariate Normal density centered at the current value of beta with variance-covariance matrix that is an approximation of the posterior based on the maximum likelihood estimates and the prior precision multiplied by the tuning parameter squared. The chains are run for 25000 iterations and the first 5000 runs are discarded as a burn-in.

An informal analysis for the convergence diagnostic of the chains is through the trace plot in Figure 1. Note that the time series look like stationarity and the chain can be considered as well mixing indicating that the convergence is reached. The acceptance rates for the chains were around 35%. Another intuitive and easily implemented diagnostic tool is an autocorrelation plot shown in Figure 2.

There are also many methods of formal convergence diagnostic based on statistical properties of Markov chains. One of the most popular methods was proposed by Geweke (1992) based on a test for equality of the means of the first and last part of a Markov chain (by default the first 10% and the last 50%). If the samples are drawn from the stationary distribution of the chain, the two means are equal and Geweke's statistic has an asymptotically standard normal distribution. Hence values of Z which fall in the extreme tails of a standard normal distribution suggest that the chain was not fully converged. This approach is available in R software with Coda package (Plummer et al., 2005). Selecting Geweke option from the CODA Diagnostics Menu produces the output for the chains given in Table 2. The results for the chains provide no evidence against convergence for each variable

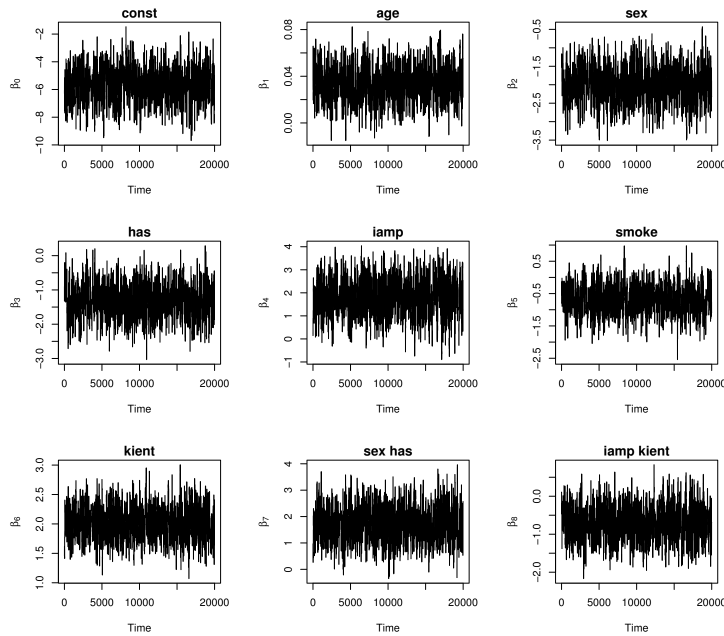

Figure 1 - Time series plot of MCMC output.

for a z-value of 1.96. Therefore we conclude that the chain converges fairly well. Hence, we can reasonably assume that our simulations are drawn from the desired posterior distribution $p(\beta\,|\,\boldsymbol{y}, X)$ and we use a kernel smoothing density estimate to represent the marginal distributions from posterior samples.

Table 2 - Geweke convergence diagnostic

Variable	Z-score
const	-0.08493
age	0.2189
sex	0.008531
has	-0.6933
iamp	-0.2467
smoke	1.695
kient	0.4666
sex has	0.9111
iamp kient	0.2234

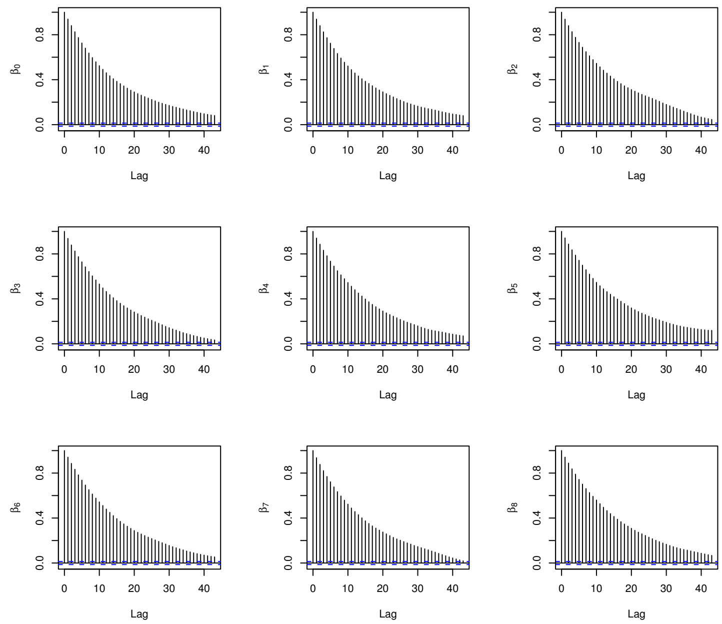

Figure 2 - Autocorrelation plot of MCMC chains for the parameters.

Figure 3 presents the estimate posterior distributions via Metropolis-Hastings and Laplace's method. The marginal posterior distributions obtained by both approaches are practically indistinguishable and seem to have a normal shape mainly due to large n value ($n = 599$).

In Table 3 it is presented posterior mean and standard deviation (se) via MCMC and Laplace's method, whose results are similar to the maximum likelihood estimate (MLE). Also, Table 4 presents a comparison of 95% intervals in Bayesian (BI) and frequentist (FI) approaches for each parameter β in the model. There are no practically differences between Bayesian and MLE estimates certainly because the sample size is large enough and the prior is normal but with very small precision. We also see that there is no strong difference between the bayesian interval (BI) and frequentist interval (FI) for the same reason.

It is important to emphasize that although the null value falls into the 95% interval for some covariates, they are considered clinically relevant by the physicians and, therefore, maintained in the model.

In order to evaluate the predictive quality of the logistic regression model, two common diagnoses measures of classification are: the c-index and the predictive rate of correct classification. The c-index represents the proportion of times in

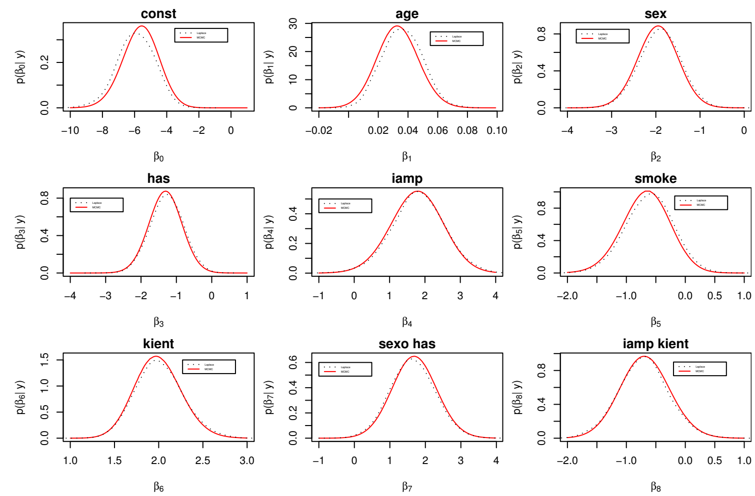

Figure 3 - Estimate posterior densities for the parameters via Metropolis-Hastings (dotted line) and Laplace (continuous line).

Table 3 - Punctual estimates and standard deviations for the regression model

Parameter	Metropolis-Hastings	Laplace	MLE
const	-5,602 (1,1258)	-5,6236 (1,1367)	-5,4512 (1,1091)
age	0,03297(0,0138)	0,0333 (0,0100)	0,0322 (0,0137)
sex	-2,0026 (0,45463)	-1,9526 (0,5198)	-1,9087 (0,4516)
has	-1,3276 (0,4587)	-1,3109 (0,4525)	-1,2748 (0,4569)
iamp	1,7877 (0,7450)	1,7738 (0,7509)	1,7736 (0,7222)
smoke	-0,6752 (0,4195)	-0,6636 (0,3973)	-0,6287 (0,3948)
kient	2,001 (0,2689)	1,9929 (0,2179)	1,9246 (0,2577)
sex has	1,7307 (0,6261)	1,6913 (0,6481)	1,6532 (0,6155)
iamp kient	-0,6905 (0,4292)	-0,6829 (0,3879)	-0,6880 (0,4120)

which the probability of in-hospital death in the survivors group is smaller than in the deaths group. The predictive rate of correct classification is the proportion of patients correctly allocated to the death and survival groups. These measures allow to evaluate the predictive ability of the model jointly. This fitting checking for the model proposed in this paper are presented in Santos (2007) with excelent results providing rate of correct classification 88% and c-index 90%. Souza and Migon (2004) also perform a Bayesian analysis but using Gibbs sampler algorithm to make the inference and present diagnostic checking and residuals analysis for

Table 4 - Bayesian interval vs frequentist interval for parameters

Parameter	FI	BI
const	(-7,62; -3,28)	(-7,88; -3,46)
age	(0,005; 0,059)	(0,006; 0,060)
sex	(-2,79; -1,02)	(-2,88; -1,09)
has	(-2,17; -0,38)	(-2,21; -0,40)
iamp	(0,36; 3,19)	(0,33; 3,25)
smoke	(-1,40; 0,14)	(-1,51; 0,15)
kient	(1,42; 2,43)	(1,49; 2,56)
sex has	(0,45; 2,86)	(0,55; 2,95)
iamp kient	(-1,49; 0,12)	(-1,55; 0,12)

this data set. Their Bayesian residual analysis is based on an approach proposed by Albert and Chib (1995).

Conclusions

In this paper we shown that the Bayesian logistic regression model is very useful and performs very well to predict, based on some relevant covariates, the in-hospital death in patients with acute myocardial infarction. The use of approximation approaches is justified because the analytical integration of joint distribution for the regression parameter β is intractable. We show that, by using the Laplace approximation we can directly compute very accurate approximations to the posterior marginals. The main benefit of this approximation is computational: where MCMC algorithms need hours to run for a large number of covariates, Laplace approximation provides precise estimates in few minutes although the Laplace's approximation requires quite a bit more analytical work (calculation of gradients and Hessians). The methods based on MCMC are easier to apply however for high-dimension parametric spaces the convergence can be difficult and bit long to achieve.

Besides, in our practical application it was shown that the Laplace's method and MCMC algorithm perform as well as classical procedures by considering a weak prior distribution. The Bayesian framework also allows the use of addictional information that can provide more accurate estimate.

SANTOS, M. A.; MAOLA, F. A.; TACHIBANA, V. M. Métodos aproximados para análise bayesiana em regressão logística. *Rev. Bras. Biom.,* São Paulo, v.27, n.2, p.288-300, 2009.

▪ RESUMO: Neste artigo consideramos a inferência Bayesiana para a estimação do modelo de regressão logística. O artigo apresenta dois métodos aproximados, o método de Laplace (Tierney e Kadane, 1986) e cadeias de Markov Monte Carlo (MCMC) para obter a densidade a posteriori marginal para os parâmetros. Uma comparação destes métodos é realizada para a predição de óbito hospitalar em pacientes internados com infarto agudo do miocárdio em um particular hospital. Além disso, identificamos os fatores de risco que caracterizam o infarto do miocárdio. Os resultados Bayesianos também são comparados com a estimação de máxima verossimilhança.

▪ PALAVRAS-CHAVE: Regressão logística; inferência bayesiana; prioris não-informativas; aproximação de Laplace; MCMC.

References

ALBERT, J. H.; CHIB, S. Bayesian residual analysis for binary response regression models. *Biometrika*, London, v.82, n.4, p.747-759, 1995.

GELFAND, A. E.; SMITH, A. F. M. Sampling-based approaches to calculating marginal densities. *J. Am. Stat. Assoc.*, Washington, v.85, n.410, p.398-409, 1990

GEWEKE, J. Evaluating the accurary of sampling-based approaches to the calculation of posterior moments. In: BERNARDO, J. M.; BERGER, J. O.; DAWID, A. P.; SMITH, A. F. M. (Ed.) emphBayesian Statistics 4., New York: Oxford University Press, 1992. p.625-631.

GILKS, W. R.; CLAYTON, D. G.; SPIEGELHALTER, D. J.; BEST, N. G.; McNEIL, A. J.; SHARPLES, L. D.; KIRBY, A. J. Modelling complexity: applications of Gibbs sampling in medicine, *J. R. Stat. Soc., Ser. B*, New York, v.55, p.39-52, 1993.

GRIFFITHS, D. A. Maximum likelihood estimation for the beta-binomial distribution and an application to the household distribution of the total number of cases of disease. *Biometrics*, Washington, v.29, p.637-649, 1973.

HOSMER Jr., D. W.; LEMESHOW, S. *Applied logistic regression.* New York: John Wiley and Sons, 1989.

JENNINGS, D. E. Judging inference adequacy in logistic regression. *J. Am. Stat. Assoc,*, New York, v.81, p.471-476, 1986.

MIGON, H. S.; TACHIBANA, V. M. Bayesian approximations in randomized response models. *Comp. Stat. Data Anal.*, Amsterdam, v.24, p.401-409, 1997.

O'HAGAN, A.; WOODWARD, E. G.; MOODALEY, L. C. Practical Bayesian analysis of simple logistic regression: predicting corneal transplants. *Stat. Med.*, New York, v.9, p.1091-1101, 1991.

PLUMMER, M.; BEST, N.; COWLES, K.; VINES, K. Output analysis and diagnostics for MCMC. The Coda Package. Software R. Disponível em: <http://www-fis.iarc.fr/coda/>. Acesso em: jan. 2009.

SANTOS, M. A. *Métodos Bayesianos aproximados para o modelo de regressão logística*. 2007. (Trabalho de graduação em Estatística) - Faculdade de Ciência e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, 2007.

SMITH, A. F. M.; ROBERTS, G. O. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods (with discussion). *J. R. Stat. Soc. Ser. B*, New York, v.55, p.3-24, 1993.

SOUZA, A. D. P. *Métodos aproximados em modelos hierárquicos diâmicos Bayesianos*. 1999. Tese (Doutorado em Engenharia de Produão) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1999.

SOUZA, A. D. P.; MIGON, H. S. Bayesian binary regression model: An aplication to in-hospital death after AMI prediction. *Pesq. Oper.*, São José dos Campos, v.24, n.2, p.253-267, 2004.

TACHIBANA, V. M. *Métodos aproximados em modelos Bayesianos de resposta aleatorizada e regressão logística*. 1995. Tese (Doutorado em Engenharia de Produção) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1995.

TIERNEY, L.; KADANE, J. B. Accurate approximations for the posterior moments and marginal densities. *J. Am. Stat. Assoc.*, Washington, v.81, n.393, p.82-86, 1986.

TIRNEY, L.; KASS, R. E.; KADANE, J. B. Fully exponential Laplace approximations to expectations and variances of non-positive functions. *J. Am. Stat. Assoc.*, Washington, v.84, n.407, p.710-716, 1989.

Received in 17.11.2008.

Approved after revised in 06.07.2009.