Sampling-Based Inference for the Generalized Time-Dependent Logistic Hazard Model

Francisco Louzada-Neto and Caroline Pires Cremasco
Dept. of Statistics, Universidade Federal de São Carlos, Brazil
dfn@ufscar.br

Gilbert MacKenzie
Dept. of Mathematics & Statistics, University of Limerick, Ireland
gilbert.mackenzie@ul.ie

Abstract

Lifetimes which satisfy a non-proportional hazard model may arise in several areas, such as, Medicine, Biometrics, Criminology and Industrial Reliability. For these data it is reasonable to presume that the hazard function is time-dependent, thereby accommodating crossing hazards. Such dependency can be modelled directly by introducing a time-dependent term in the model for the hazard function. Accordingly, in this paper we utilize a generalized time-dependent logistic (GTDL) hazard model which can accommodate non-proportional hazards data. A sampling-based inference procedure based on Markov chain Monte Carlo Methods is developed and the methodology is used to investigate survival from advanced lung cancer in a well known dataset.

Keywords and Phrases: Bayesian Modelling, GTDL, Hazard Modelling, MCMC, Model Comparison, Non-proportional Hazard Function, Sampling-Based Inference, Time-Dependent Hazard Function.

2000 Mathematics Subject Classification: Statistics.
1 Introduction

In practice, it is common to find data with crossing hazards, which cannot be accommodated by the usual proportional hazards (PH) model of Cox (1972). In the United Kingdom it is well-known that survival from cancer of the female breast and cancer of the lung tend not to follow the proportional hazards assumption, especially in age (Blagojevic, MacKenzie & Ha, 2003; MacKenzie & Gillon, 2004). Accordingly, as an example, we consider data on the survival of males with advanced inoperable lung cancer (Kalbfleisch and Prentice, 1980, p.60).

MacKenzie (1996, 2002) proposed the generalized time-dependent logistic hazard regression (GTDL) model as a wholly parametric competitor for the PH mode. An advantage of the GTDL model is its generalization of the relative risk in the PH model of Cox (1972) to time-dependent form. The upper panel of Figure 1 shows the estimated cumulative hazard function, while the lower panel of Figure 1 shows the Kaplan-Meier survival estimate and the fit of the GTDL model to the data, which will be discussed further in Section 4. The model also inherits the tail-deficit property of the GTDL family, which implies that, for some values of the parameters, the data are a mixture of "mortals" and "immortals". That the GTDL model has such a frailty interpretation, is perhaps not surprising, since the density may be derived as a Gompertz ∼ Gamma mixture (MacKenzie, 1996).

To date inference for the GTDL family has been conducted, wholly, in the classical framework. However, in this paper we develop, for the first time, a Bayesian approach for analyzing the GTDL model in which inference for the model parameters is based on Markov Chain Monte Carlo (MCMC) methods. In Section 2, we describe the model in detail, while in Section 3 we present sampling-based inference for the model parameters. A real set of medical data is analyzed in Section 4 and in Section 5 we make some final remarks.

2 Model Formulation

2.1 Basic Properties

Let \( h(t|x) \) denote the hazard function at time \( t \) for an individual with covariate vector \( x \).

Then, the GTDL model (MacKenzie, 2002) assumes that

\[
h(t|x) = \frac{\lambda \exp(t\alpha + x'\beta)}{1 + \exp(t\alpha + x'\beta)}
\]  
(1)
where $\lambda > 0$ is a scalar, $\alpha$ is a scalar measuring the effect of time and $\beta' = (\beta_1, ..., \beta_p)'$ is a vector of $p$ parameters measuring the influence of the $p$ covariates $x' = (x_1, ..., x_p)'$. Consequently, the density function may be written as

$$f(t|x) = \lambda \{q(\alpha, \beta)g(\beta)\}^{\frac{1}{\alpha}}$$

where the individual components are simple functions of the time-dependent multiple-logistic function, are given by

$$p(\alpha, \beta) = \exp(\alpha t + x'\beta)/\{1 + \exp(\alpha t + x'\beta)\}$$

$$q(\alpha, \beta) = 1/\{1 + \exp(\alpha t + x'\beta)\}$$

$$g(\beta) = 1 + \exp(x'\beta).$$

Intrinsically, equation (1) is neither a proportional hazards model nor a accelerated life model, but it will approach a proportional hazards model when $q(\alpha, \beta) \approx 1$ and, when this condition holds, the estimates of the regression parameter $\beta$ should be similar in both models. Moreover, when $\alpha = 0$ the hazard reduces to a multiple of the usual linear logistic regression model whence the survival distribution is a type of PH model, i.e., an Exponential regression model with

$$E(T) = \{\lambda p(\alpha, \beta)\}^{-1}$$

$$V(T) = \{\lambda p(\alpha, \beta)\}^{-2}$$

Accordingly, the GTDL is flexible enough to model PH and non-PH survival data.

### 2.2 Genesis

Since its original development more has become known about alternative derivations of the GTDL model. The form of the density (2) may be derived in a variety of different ways. We have already noted one route above, namely as a Gompertz $\sim$ Gamma mixture (MacKenzie, 1996), a result which places the models in the family described by Aalen (1988). MacKenzie (2002) also showed that the density could be derived as the modulus of a special case of Fisher’s Z distribution. Alternatively we may begin with a Weibull $\sim$ Gamma mixture leading to Pareto distribution of the second kind and by applying a non-linear transformation of the time scale, reach the GTDL form (Blagojevic, MacKenzie & Ha, 2004).
Figure 1: High panel: cumulative hazard plot. Lower panel: survival plot and the fits of the GTDL model (curve adjusted).

2.3 Relative Risk

The time dependent relative risk (RR) function, defined as the ratio of hazard functions at time $t$, for two subjects with different covariate vectors $x_1$ and $x_0$, is given by

$$\rho(t|x_1, x_0) = h(t|x_1)/h(t|x_0) = \exp\{(x_1 - x_0)^T \beta\} \psi(t|x_1, x_0)$$

(2)

where

$$\psi(t|x_1, x_0) = \frac{q(t|\alpha, \beta, x_1)}{q(t|\alpha, \beta, x_0)}.$$  

The leading term on the right hand side of (5) is Cox’s constant of proportionality (the RR in a PH model) and in the GTDL model this constant is moderated by $\psi(\cdot)$, a function of both time and the covariates, thus demonstrating again that the model is non-PH. Moreover, it should be
noted that (5) does not depend on the parameter $\lambda$. In general the rate of change of the relative risk with time is

$$\frac{\partial \rho(t|x_1, x_0)}{\partial t} = -\alpha \exp\{(x_1 - x_0)^T \beta\} \psi(t|x_1, x_0)[p(t|\alpha, \beta, x_1) - p(t|\alpha, \beta, x_0)]$$

(3)
a function which is zero when $\alpha = 0$ and when $x_1 = x_0$, where $x_0$ is a reference value.

3 Inference

3.1 Likelihood

Consider a sample of independent random variables $T_1, \ldots, T_n$ denoting the lifetimes of $n$ units. Assume that $T_i$ has associated an indicator variable defined by $\delta_i = 1$ if $T_i = t_i$ is an observed failure time and $\delta_i = 0$ if it is a right-censored observation. When the censoring process is non-informative, the likelihood function for the generic parameter $\theta' = (\lambda, \alpha, \beta')$ is given in general by

$$L(\theta) = \prod_{i=1}^{n} h(t_i|\theta)^{\delta_i} S(t_i|\theta)$$

(Lawless, 1982). For the regression model defined by (1), the survivor function is

$$S(t_i|\theta) = \left\{ \frac{1 + \exp(x_i' \beta)}{1 + \exp(t_i \alpha + x_i' \beta)} \right\}^{\lambda/\alpha}$$

(4)
whence the likelihood function becomes

$$L(\lambda, \alpha, \beta) = \prod_{i=1}^{n} \left\{ \frac{\exp(t_i \alpha + x_i' \beta)}{1 + \exp(t_i \alpha + x_i' \beta)} \right\}^{\delta_i} \left\{ \frac{1 + \exp(x_i' \beta)}{1 + \exp(t_i \alpha + x_i' \beta)} \right\}^{\lambda/\alpha}.$$ 

(5)

3.2 Sampling-Based Inference

For inference we adopt a fully Bayesian approach. The prior distributions for the parameters in the model, details of the MCMC algorithm and the comparison of models, via Bayes’ factors, are described below. The target distribution for inference is the posterior, $\pi(\theta|data) \propto \pi(\theta)L(data|\theta)$, where $\pi(\theta)$ is the prior for $\theta$.

3.2.1 Priors

Since each parameter of the model (1) has a direct interpretation in the context of the time-to-event data, available expert opinions may be expressed in terms of a prior distribution for
each parameter separately. Thus, one approach, is to encapsulate expert opinion, on the model parameters, as a set independent marginal distributions. This is by no means the only approach available in this setting, but it is a natural first step which has the advantage of simplifying the resulting computations. Thus, we assume that the joint density of $\lambda$, $\alpha$ and $\beta$ is given by

$$
\pi(\lambda, \alpha, \beta) = \pi(\lambda) \pi(\alpha) \pi(\beta).
$$

In order to have unbounded parameters we consider the reparametrization $\lambda = \exp \varphi$ and $\alpha = \exp \phi$ and assume normal prior for $\varphi, \phi$ and $\beta$. That is, $\varphi \sim N(\varphi_0, \kappa_0)$, $\phi \sim N(\phi_0, \kappa_0)$ and $\beta \sim N_p(\mu_0, \Sigma_0)$. These arrangements ensure that $\lambda > 0$, a global condition, and also that $\alpha > 0$, a local condition tailored to the hazard form encountered in the lung cancer data analyzed. The hazard is a increasing function of time. More generally, one may adopt a Normal prior for $\alpha$, rather than the log-Normal prior implemented here.

### 3.2.2 MCMC

Irrespective of the form of the priors considered, the joint posterior distributions of the parameters are analytically intractable. We overcome this computational difficulty by using the Metropolis-Hastings algorithm (Hastings, 1970; Chib and Greenberg, 1995), which allows us to simulate observations from complicated joint distributions by generating random samples successively from the full conditional distributions for the unknown parameters. The full conditional posterior densities for $\lambda$, $\alpha$ and $\beta$ that are used in each step of the iterative sampling-based algorithms are given in Appendix A.

To generate samples of $\lambda$, $\alpha$ and $\beta$ from their conditional distributions we first start with the values $\theta^{(0)} = (\lambda^{(0)}, \alpha^{(0)}, \beta^{(0)})$, then we generate $\theta^*$ from the the proposal distribution $q(\theta, \theta^*)$, which is assumed to be the prior distribution as presented in the section above, and later generate the value $u$ from the uniform distribution $U(0, 1)$ in order to test the acceptability of the proposal. Thus, if

$$
u \leq \min \left(1, \frac{\pi(\theta^*) q(\theta^*, \theta^{(j)})}{\pi(\theta^{(j)}) q(\theta^{(j)}, \theta^*)} \right),$$

we accept $\theta^{(j+1)} = \theta^*$ else we set $\theta^{(j+1)} = \theta^{(j)}$. We then repeat the process, but now using $\theta^{(1)} = (\lambda^{(1)}, \alpha^{(1)}, \beta^{(1)})$ as the starting values and so on until to obtain the desired sample.
3.2.3 Model Uncertainty

We now consider the problem of accounting for uncertainty about model form. We are faced with models that involve time-dependent terms or sets of covariates or other model parameters. In order to decide for the best model to be fitted we can use the Bayes factor which is the relative weight of evidence for model $M_1$ compared to model $M_2$ given by

$$B_{12} = \frac{f(t_{\text{obs}}|M_1)}{f(t_{\text{obs}}|M_2)},$$

where $t_{\text{obs}}$ denotes the actual observations and $f(t_{\text{obs}}|M_k)$ denotes the marginal density under model $M_k$, $k = 1, 2$ (Gelfand, 1996). The model $M_1$ is preferred over $M_2$ when $B_{12} > 1$; see Kass and Raftery (1995) for more details. It can be useful to consider twice the logarithm of (6), which is on the same scale as the deviance and the likelihood ratio test statistics. According to the rough classification of Kass and Raftery (1995), there is no evidence of difference when twice the logarithm of the Bayes factors lies between 0 and 1/2. However, when it lies between 1/2 and 1 there is positive evidence, between 1 and 2 strong evidence, and when it is greater than 2 there is very strong evidence against model $M_1$. Of course this evidence is crude, but more detailed calculations depend crucially on the choice of prior. We approximate the marginal densities in (6) by their Monte Carlo estimates, obtained from the S generated samples, given by

$$\hat{f}(t_{\text{obs}}|M_k) = (1/s) \sum_{s=1}^{S} f(t_{\text{obs}}|\theta_i^{(s)}, M_i).$$

In addition to the Bayes factor above we use the conditional predictive ordinate (CPO) statistics (see Chen and Chang, 1997), a cross-validated predictive approach for model diagnostic.

Considering the $i$-th lifetime, its CPO statistics over model $M_k$ can be defined as,

$$CPO_{i,k} = f_k \left( t_i|D^{(-i)} \right) = f_k \left( t_i|\Theta, x_i \right) \pi_k \left( \Theta|D^{(-i)} \right) d\Theta,$$

where $t_i$ denotes the $i$-th lifetime, $x_i$ is its covariate vector, $D^{(-i)}$ denotes the data excluding the $i$-th lifetime and $\pi_k \left( \Theta|D^{(-i)} \right)$ denotes the posterior density for $\Theta$ given the data $D^{(-i)}$ and the model $M_k$. We approximate the $CPO_{i,k}$ statistics in (7) by their Monte Carlos estimates, obtained from the S generated samples. We can visualize the better fitting over competing models by plotting the logarithm of the odds between the $CPO_{i,k}$s over the both models against the number of observations with positive evidence to the model with biggest $CPO_{i,k}$ odds logarithm.
A summary of the \( CPO_{i,k} \) plot result is the mean of the logarithm of the \( CPO_i \)'s given by,

\[
MLCPO_k = \frac{\sum_{i=1}^{n} \log (CPO_{i,k})}{n}.
\]

(8)

High \( MLCPO_k \) values are positive evidence to a model \( M_k \) in comparison with another one.

## 4 Reanalysis of Lung Cancer Data

### 4.1 Data Description

Survival time in days and several covariates were available for the 137 patients with inoperable lung cancer; 9 patients were right-censored. After preliminary investigations, it was discovered that initial performance status exerted a strong prognostic effect. The Karnofsky score is measured on a scale 0-100, with high values implying improved performance, typically among patients who are less-ill. Although the original objective of this trial was to assess chemotherapy, we focus on how the Karnofsky score (henceforth performance) influences survival. Patients with performance up to 50 were assigned to Group 1 while patients with performance greater than 50 were assigned as Group 2. The left panel of Figure 1 shows the cumulative hazard plot for each group. As the plots cross, the PH model does not seem to be appropriate and we therefore adopt the non-PH GTDL model form and apply the methods described above.

### 4.2 Bayesian Analysis

We first set the hyperparameters to the following values: \( \varphi_0 = \alpha_0 = \mu_0 = 0 \) and \( \kappa_0 = 10 \) and \( \Sigma_0 = 10 \). The hyperparameter values were chosen subjectively. However, subsequently, a small sensitivity analysis was conducted by choosing different hyperparameter values. Large variances were considered to ensure non-informativeness. The sensitivity analysis did not modify, substantially, the results presented below and accordingly details (available from the authors) are omitted here.

Summaries of the posterior distributions of the parameters \( \lambda, \alpha \) and \( \beta \) were calculated from the samples generated by the Metropolis-Hastings technique. To generate 3 chains of 53,000 iterations for the parameters. The first ones 3,000 were unknown and of the remaining we selected samples with jump of 5, resulting in a final sample of 10,000 iterations.
For using of the algorithm Metropolis-Hastings, it was used to generate values, candidates the own priori densities. Although this procedure is artificial, it is possible to monitoring through the acceptance rate (Bessag et al., 1995 and Geyes, 1992). In addition, the Gelman Rubin convergence approach was applied (Gelman et al., 1995) with 3 chains for it assures the homogeneity of the generated sequences. The number of iteration was considered enough for the approximate convergence when the reduction of esteemed potential scale was $R < 1.1$.

Also, the convergence of the chains were assessed according to three convergence diagnostics implemented in CODA (Best et al., 1997). The graphical output and kernel density estimation for each parameter showed that there were no convergence problems (see Figure 2 in Appendix B). Interested readers can obtain the computational codes used to generate the chains by writing directly to the first author.

The results are summarized in Table 1, which shows the posterior means, the posterior standard deviations and the 95% credible intervals for the parameters of interest.

It is important to note that there is significant difference in survival between Groups 1 and 2, as indicated by the estimated $\beta_1$. This result is corroborated by the value of the twice the logarithm of (6) and by the difference between the means of the logarithm of the $CPO_i$'s (8) of the full model ($M_2$) with respect to the model without the binary covariate ($M_1$), which are 3.02 and 0.92, respectively. Also, there is significant time-dependent effect. The twice the logarithm of (6) and the difference between the means of the logarithm of the $CPO_i$'s (8) of the full model ($M_2$) with respect to the model without time-dependent term ($M_1$) are equal 2.54 and 0.97, respectively. The lower panel of Figure 1 shows the Kaplan-Meier survival estimate and the fits of the GTDL model to the data. Figure 3 in Appendix C show the logarithm of the odds between the $CPO_{i,k}$'s over the both models against the number of observations, giving evidence to the models $M_2$.

Although, the PH and GTDL models have different physical interpretations and should be treated separately, as a last comparison, we fitted a PH model to the data. The value of the twice the logarithm of (6) of the PH model versus the GTDL model is 1.96, giving strong evidence in favour of the GTDL model.
Table 1: Posterior summaries for the model parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>95% Credible</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varphi$</td>
<td>0.959</td>
<td>0.024</td>
<td>[0.91 ; 1.00]</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.609</td>
<td>0.084</td>
<td>[0.56 ; 0.77]</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>−0.968</td>
<td>0.015</td>
<td>[−0.99 ; −0.06]</td>
</tr>
</tbody>
</table>

4.3 Coverage Probabilities

Inference for the parameter vector $\theta = (\lambda, \alpha, \beta)$ can be based on the well-known large sample properties of the MLEs where $\hat{\theta}_{\text{aaysm}} \sim N_p(\theta, \Sigma(\theta))$, where $p$ is the dimension of $\theta$.

However, in reliability and survival studies, it is common to find small or moderate datasets. In order to check the behavior of the asymptotic theory for small and moderate sized samples, we performed a small-scale simulation study for examining the coverage probabilities of the asymptotical confidence intervals for the parameters. The study was based on generating 1,000 samples according to the following scheme. Each lifetime $t_i$ was given by $t_i = \min(y_i, c_i)$, for $i = 1, ..., n$, where $y$ and $c$ were two independent random variables representing the lifetimes and the censoring times, respectively. Both were generated according to (1) with $\lambda = \alpha = \beta = 0.5$ and $x_i$ was generated according to a Bernoulli distribution with probability of success equal to 0.5. The censoring variable was given by $\delta_i = 1$, for $y_i < c_i$ and $\delta_i = 0$, otherwise, characterizing a Type I censoring scheme. We considered sample sizes in the range: $n = 15, 30, 50, 100, 300$ and 1000.

Table 2 shows the variation in coverage of nominal 90% confidence intervals by sample size. For example, the 90% confidence interval based on $n=1000$ is given by (0.884, 0.9156). If a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverage should be inside these bounds. Clear under-coverage of the confidence intervals for small and moderate sized samples is present in the table. Such findings are evidence for the need of a more adequate procedure for small or moderate sized samples such as the Bayesian procedures developed here.
Table 2. Coverage probabilities of the 90% asymptotic confidence intervals.

<table>
<thead>
<tr>
<th>n</th>
<th>φ</th>
<th>ϕ</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>0.781</td>
<td>0.762</td>
<td>0.833</td>
</tr>
<tr>
<td>30</td>
<td>0.887</td>
<td>0.874</td>
<td>0.898</td>
</tr>
<tr>
<td>50</td>
<td>0.891</td>
<td>0.889</td>
<td>0.872</td>
</tr>
<tr>
<td>100</td>
<td>0.913</td>
<td>0.921</td>
<td>0.893</td>
</tr>
<tr>
<td>300</td>
<td>0.899</td>
<td>0.881</td>
<td>0.911</td>
</tr>
<tr>
<td>1000</td>
<td>0.933</td>
<td>0.905</td>
<td>0.900</td>
</tr>
</tbody>
</table>

Table 3. Slope of $\log\{\text{var}(\cdot)\}$ on $\log n$: expressing the relation between the variance and $n$.

<table>
<thead>
<tr>
<th>n</th>
<th>var(φ)</th>
<th>var(ϕ)</th>
<th>var(β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[15, 30]</td>
<td>-1.401</td>
<td>-1.965</td>
<td>-1.468</td>
</tr>
<tr>
<td>[30, 50]</td>
<td>-1.131</td>
<td>-1.096</td>
<td>-1.305</td>
</tr>
<tr>
<td>[50, 100]</td>
<td>-0.974</td>
<td>-1.000</td>
<td>-1.160</td>
</tr>
<tr>
<td>[100, 300]</td>
<td>-0.946</td>
<td>-0.939</td>
<td>-1.22</td>
</tr>
<tr>
<td>[300, 1000]</td>
<td>-0.992</td>
<td>-0.998</td>
<td>-0.938</td>
</tr>
</tbody>
</table>

Table 3 shows the slopes obtained by regressing $\log\{\text{var}(\cdot)\}$ on $\log n$. That is, the first entry of Table 3 means that, for $15 \leq n \leq 30$, $\text{var}(\hat{\beta}) \propto n^{-1.401}$, which correspond to a difference in slope of 40.1% in comparison with the asymptotic slopes which are equal to 1. Overall, the asymptotic slopes are well approached only for $n \geq 100$, corroborating again the need for the type of statistical methodology developed above.

5 Final Remarks

We have developed a Bayesian approach for the analysis of the GTDL survival regression model. Although, we have adopted Normal priors throughout, the methodology employed is quite general, and other prior specifications could be adopted relatively easily. Inference for the
model parameters is based on Markov Chain Monte Carlo Methods which worked well in this application. Our comparative study of model forms, revealed that it is possible to distinguish between competing models for the lung cancer data, using Bayes factors. In this case the evidence favoured the GTDL model over the PH model, in the presence of a binary covariate based on performance status.

Acknowledgments

This work has been partially supported by the Brazilian Organization FAPESP and CNPq. Francisco Louzada-Neto was supported by CNPq grant number 300045/2003-4. Caroline Pires Cremasco was supported by FAPESP schoolarship number 03/05132-0. Gilbert MacKenzie was supported by the SFI-funded research project BIO-SI (SFI 07/MI/012).

Appendix A: The conditional posteriors for the model parameters

Considering the joint prior distribution for $\varphi, \phi$ and $\beta$ given by $\pi(\varphi, \phi, \beta) = \pi(\varphi) \pi(\phi) \pi(\beta)$ and combining this joint prior distribution with the likelihood function (5) we obtain the joint posterior distribution for $\varphi, \phi$ and $\beta$ is given by

$$\pi(\varphi, \phi, \beta | \text{Data}) \propto \pi(\varphi) \pi(\phi) \pi(\beta) \, L(\varphi, \phi, \beta).$$

After some algebraic manipulations, the full conditional distributions for $\varphi, \phi$ and $\beta$ are given by as following.

The conditional posteriori of $\varphi$ given $\phi$ and $\beta$ is

$$\pi(\varphi | \phi, \beta, \text{Dados}) \propto \prod_{i=1}^{n} \left\{ \exp(\delta_i \varphi) \left[ \frac{1 + \exp(t_i \exp(\phi) + x'_i \beta)}{1 + \exp(x'_i \beta)} \right]^{-\exp(\phi)} \right\} \pi(\varphi).$$

The conditional posteriori of $\phi$ given $\varphi$ and $\beta$ is

$$\pi(\phi | \varphi, \beta, \text{Dados}) \propto \prod_{i=1}^{n} \left\{ \frac{\exp(t_i \exp(\phi))}{1 + \exp(t_i \exp(\phi) + x'_i \beta)} \right\}^{\delta_i} \left[ \frac{1 + \exp(t_i \log(\phi) + x'_i \beta)}{1 + \exp(x'_i \beta)} \right]^{-\exp(\phi)} \pi(\phi).$$

The conditional posteriori of $\beta_j$ ($j = 1, ..., p$) given $\varphi$ and $\phi$ is
\[ \pi(\beta_j|Dados, \phi, \beta_{-j}) \propto \prod_{i=1}^{n_i} \left\{ \left[ \frac{\exp(x_i'\beta_j)}{1 + \exp(t_i \exp(\phi) + x_i'\beta_j)} \right]^{\delta_i} \left[ \frac{1 + \exp(t_i \exp(\phi) + x_i'\beta_j)}{1 + \exp(x_i'\beta_j)} \right]^{-\frac{\exp(\phi)}{\exp(\phi)}} \right\} \pi(\beta_j). \]

**Appendix B: Plots of the generated samples and empirical marginal posterior densities**

Considering the GTDL model, Figure 2 shows the plots of the generated samples and the empirical marginal posterior densities based on the generated chains.

**Appendix C: Plots of the \( CPO_{i,k} \)'s odds logarithm over the competing models against the number of observations**

Figure 3 shows the \( CPO_{i,k} \)'s odds logarithm over the competing models against the number of observations.

**References**


Figure 2: Left panels: plots of the generated samples. Right panels: empirical marginal posterior densities.
Figure 3: High panel: plots of the logarithm of the odds between the CPO\textsubscript{i,k}ts of the full model (\(M_2\)) with respect to the model without the binary covariate (\(M_1\)). Lower panel: plots of the logarithm of the odds between the CPO\textsubscript{i,k}ts of the full model (\(M_2\)) with respect to the model without time-dependent term (\(M_1\)).