Brazilian Journal of Probability and Statistics
Volume 22, Number 2, Pages 183–205, December 2008
ISSN 0103-0752

STATISTICAL ANALYSIS FOR LONGITUDINAL COUNTING DATA IN THE PRESENCE OF A COVARIATE CONSIDERING DIFFERENT "FRAILTY" MODELS

JORGE A. ACHCAR, EMÍLIO A. COELHO-BARROS, AND EDSON Z. MARTINEZ

ABSTRACT. In this paper, we present different "frailty" models to analyze longitudinal Poisson data in the presence of a covariate. These models incorporate the extra-Poisson variability and the possible correlation among the repeated counting data for each individual. A hierarchical Bayesian analysis is introduced for each different model considering usual MCMC (Markov Chain Monte Carlo) methods. Considering a real biological data set, we also discuss some Bayesian discrimination aspects for the choice of the best model.

1. INTRODUCTION

Longitudinal Poisson data is very common in many applications, especially with medical data, where the counting are measured for each unit or individual in different times. Usually, we have the presence of one or more covariates associated to each individual.

As an illustrative example and motivation for this paper, let us consider the data set of Table 1, where we have the grooming counting of 8 males rats of the Wistar species in different times and receiving saline and oxytocin (data set obtained from CEMEQ, Medical School of Ribeirão Preto, University of São Paulo, Brazil). In this experiment, realized in the Neurophysiology and Neuroethology Experimental Laboratory (Medical School of Ribeirão Preto, University of São Paulo, Brazil), the responses of interest were measured 24 times every 5 minutes after application of the two treatments in the following order: first the animal received saline (treatment 1) where it was measured 12 grooming counting (every 5 minutes); then, the experiment was repeated with the animals receiving oxytocin (Treatment 2), also measuring the grooming counting every 5 minutes, which totalizes 24 counting for each rat. In the same way, the experiment was repeated with 6 rats of the War species (data set in Table 2). The main interest of this research was to verify if the species have different effects in the grooming counting of the rats.

In Tables 1 and 2, we also have the sample means and the sample variances for each combination time × treatment. From the results of Tables 1 and 2, we observe that the

Key words and phrases. Biological data, "frailty" models, hierarchical Bayesian analysis, longitudinal Poisson data.

Received April/2007, accepted November/2007.

TABLE 1. Grooming counting in male rats with treatment order (saline/oxytocin)-Wistar.

	Rats										
	1	2	3	4	5	6	7	8		Standard	
t(min)	Saline								Mean	Deviation	Variance
5	0	3	2	2	4	0	5	8	3.000	2.673	7.143
10	9	3	9	6	6	0	9	0	5.250	3.845	14.786
15	6	10	8	8	8	7	8	0	6.875	2.997	8.982
20	0	0	0	5	5	1	7	6	3.000	3.024	9.143
25	9	0	6	2	14	4	0	1	4.500	4.957	24.571
30	11	0	0	0	4	0	0	0	1.875	3.944	15.554
35	16	10	4	0	2	0	0	0	4.000	5.952	35.429
40	0	0	10	2	5	0	0	0	2.125	3.643	13.268
45	0	0	0	0	1	0	0	0	0.125	0.354	0.125
50	0	0	4	0	0	0	0	0	0.500	1.414	2.000
55	0	0	12	0	4	0	0	0	2.000	4.276	18.286
60	0	0	2	0	0	0	0	0	0.250	0.707	0.500
	Oxytocin										
5	9	0	15	0	2	2	1	5	4.250	5.285	27.929
10	16	13	19	0	7	0	8	6	8.625	6.968	48.554
15	16	9	18	2	17	3	11	0	9.500	7.191	51.714
20	13	0	14	11	18	10	14	1	10.125	6.402	40.982
25	17	1	9	14	19	11	19	9	12.375	6.163	37.982
30	18	5	13	16	6	14	15	11	12.250	4.652	21.643
35	18	6	1	7	16	9	14	10	10.125	5.643	31.839
40	15	2	16	11	13	0	9	8	9.250	5.800	33.643
45	13	3	15	0	7	1	19	1	7.375	7.367	54.268
50	11	0	4	0	8	8	8	10	6.125	4.291	18.411
55	8	0	13	7	9	3	7	0	5.875	4.549	20.696
60	5	0	0	0	0	7	1	7	2.500	3.251	10.571

sample means are different of the sample variances for almost all combinations time × treatment, which is an indication of extra-Poisson variability.

To analyze the data of Tables 1 and 2, we assume a Poisson distribution for the counting data in the presence of a covariate. To incorporate the dependence among the counting data and the extra-Poisson variability, we introduce a random effect or "frailty" in different regression models for the parameter of the Poisson distribution. The use of a random effect or a "frailty" to analyze longitudinal discrete data is considered by many authors Albert & Chib (1993); Crouchley & Davies (1999); Dunson (2000, 2003); Jorgensen et al. (1999); Henderson & Shimakura (2003); Dunson & Herring (2005). Generalized linear mixed models with normally distributed random effects are considered by many authors Moustaki (1996); Sammel et al. (1997); Moustaki & Knott (2000); Dunson (2000, 2003).

In biostatistics applications, alternative Poisson latent variable models have been proposed in the literature, motivated by applications to studies of malformation (Legler &

TABLE 2. Grooming counting in male rats with treatment order (saline/oxytocin)-War.

	Rats								
	1	2	3	4	5	6		Standard	
t(min)	Saline						Mean	Deviation	Variance
5	2	2	6	2	6	4	3.667	1.966	3.867
10	3	3	0	2	1	2	1.833	1.169	1.367
15	3	0	6	10	7	0	4.333	4.033	16.267
20	11	7	8	7	0	9	7.000	3.742	14.000
25	1	0	13	0	8	1	3.833	5.419	29.367
30	3	0	1	3	0	0	1.167	1.472	2.167
35	2	1	0	3	0	0	1.000	1.265	1.600
40	0	0	0	0	0	0	0.000	0.000	0.000
45	5	6	0	0	0	0	1.833	2.858	8.167
50	0	1	0	0	0	0	0.167	0.408	0.167
55	0	0	0	0	3	0	0.500	1.225	1.500
60	0	0	4	11	10	0	4.167	5.154	26.567
	Oxytocin								
5	4	6	1	9	0	7	4.500	3.507	12.300
10	1	13	0	13	5	8	6.667	5.680	32.267
15	9	19	6	1	10	10	9.167	5.913	34.967
20	11	12	6	14	3	5	8.500	4.416	19.500
25	4	15	8	16	13	14	11.667	4.676	21.867
30	17	15	2	12	0	12	9.667	7.005	49.067
35	20	8	12	16	0	5	10.167	7.333	53.767
40	1	10	7	6	16	8	8.000	4.940	24.400
45	0	6	0	0	5	9	3.333	3.882	15.067
50	1	17	2	12	0	8	6.667	6.861	47.067
55	0	5	5	0	8	14	5.333	5.279	27.867
60	1	14	3	8	12	4	7.000	5.215	27.200

Ryan, 1997) and tumorigenesis Yakovlev & Tsodikov (1996); Dunson & Baird (2002). Poisson-Gamma models for longitudinal counts are proposed by Crouchley & Davies (1999); Jorgensen et. al (1999); Henderson & Shimakura (2003) and gamma frailty models for survival data was introduced by Clayton (1991).

In this paper, we develop a comparative study for different "frailty" structures using hierarchical Bayesian methods based on Gibbs Sampling algorithm method (Gelfand & Smith, 1990; Chib & Greenberg, 1995). The paper is organized as follows: in Section 2, we present the formulation of the model considering three different "frailty" models to analyze longitudinal Poisson data; in Section 3, we introduce a Bayesian analysis for each model considering the use of MCMC (Markov Chain Monte Carlo) methods; in Section 4, we analyze the real data set introduced in Tables 1 and 2; and finally, in Section 5, we introduce a section of concluding remarks and some discussion of the obtained results.

2. Formulation of The Model

Let Y_{ij} be a random variable with a Poisson distribution,

$$P\left(Y_{ij} = y_{ij}\right) = \frac{e^{-\lambda_{ij}}\lambda_{ij}^{y_{ij}}}{y_{ij}!}, \tag{2.1}$$

where $y_{ij} = 0, 1, 2, \ldots$; $i = 1, \ldots, n$ (sample size) and $j = 1, \ldots, k$ (number of times). Associated to each combination time × individual, let us assume the presence of a covariate x_{ij}, $i = 1, \ldots, n$; $j = 1, \ldots, k$.

Since we have longitudinal data, we introduce a random effect or a "frailty" which captures the correlation among the repeated measures for each individual and the extra-Poisson variability. Different models are considered to analyze longitudinal counting data.

2.1. **Model 1.** Assuming that the counting data follows a Poisson distribution (2.1) with parameter λ_{ij}, let us assume the regression model,

$$\lambda_{ij} = \alpha_j \exp\left(\beta_j x_{ij} + w_i\right) \tag{2.2}$$

where $x_{ij} = 0$ indicates that the rat is from the Wistar group and $x_{ij} = 1$ indicates that the rat is from the War group. In this way, α_j measures the grooming mean in the j^{th} time for the rats of Wistar group; $\alpha_j e^{\beta_j}$ measures the grooming mean in the j^{th} time for the rats of War group; β_j is a regression parameter indicating the effect of the group species. In model (2.2), we also have the presence of a random effect or "frailty" w_i which captures the possible correlation among the repeated measures for each individual and extra-Poisson variability, assuming a normal distribution, that is,

$$w_i \overset{iid}{\sim} N\left(0, \tau^2\right) \tag{2.3}$$

for $i = 1, \ldots, n$.

Since Y_{ij} has a Poisson distribution, we have $E\left(Y_{ij}|\lambda_{ij}\right) = \lambda_{ij}$ and $Var\left(Y_{ij}|\lambda_{ij}\right) = \lambda_{ij}$. That is,

$$\begin{aligned} E\left(Y_{ij}|\alpha_j, \beta_j, w_i, x_{ij}\right) &= \alpha_j \exp\left(\beta_j x_{ij} + w_i\right); \\ Var\left(Y_{ij}|\alpha_j, \beta_j, w_i, x_{ij}\right) &= \alpha_j \exp\left(\beta_j x_{ij} + w_i\right). \end{aligned} \tag{2.4}$$

As,

$$E\left(Y_{ij}|\alpha_j, \beta_j, x_{ij}\right) = E\left[E\left(Y_{ij}|\alpha_j, \beta_j, w_i, x_{ij}\right)\right],$$

we have from (2.4),

$$E\left(Y_{ij}|\alpha_j, \beta_j, x_{ij}\right) = \alpha_j e^{\beta_j x_{ij}} E\left(e^{w_i}\right).$$

Also observe that from the normality of the random effects w_i (from (2.3)), e^{w_i} has a log-normal distribution with mean $E\left(e^{w_i}\right) = e^{\tau^2/2}$ and variance $Var\left(e^{w_i}\right) = \left(e^{\tau^2} - 1\right)e^{\tau^2}$. That is,

$$E\left(Y_{ij}|\alpha_j, \beta_j, x_{ij}\right) = \alpha_j e^{\beta_j x_{ij}} e^{\tau^2/2}. \tag{2.5}$$

As,

$$Var\left(Y_{ij}|\alpha_j, \beta_j, x_{ij}\right) = Var\left[E\left(Y_{ij}|\alpha_j, \beta_j, w_i, x_{ij}\right)\right] + E\left[Var\left(Y_{ij}|\alpha_j, \beta_j, w_i, x_{ij}\right)\right],$$

we have from (2.4),

$$Var\left(Y_{ij}|\alpha_j,\beta_j,x_{ij}\right) = \alpha_j^2 e^{2\beta_j x_{ij}} Var\left(e^{w_i}\right) + \alpha_j e^{\beta_j x_{ij}} E\left(e^{w_i}\right),$$

that is

$$Var\left(Y_{ij}|\alpha_j,\beta_j,x_{ij}\right) = \alpha_j^2 e^{2\beta_j x_{ij}} \left(e^{\tau^2}-1\right) e^{\tau^2} + \alpha_j e^{\beta_j x_{ij}} e^{\tau^2/2}. \tag{2.6}$$

From (2.5) and (2.6), we observe that the mean and the variance of Y_{ij} given α_j, β_j and x_{ij} are different, that is, we have the presence of extra-Poisson variability given by the term $\alpha_j^2 e^{2\beta_j x_{ij}} \left(e^{\tau^2}-1\right) e^{\tau^2}$, incorporated to the model (2.1) and (2.2).

2.2. **Model 2.** Let us assume that the counting data follows a Poisson distribution (2.1) with λ_{ij} given by,

$$\lambda_{ij} = w_i \alpha_j \exp\left(\beta_j x_{ij}\right), \tag{2.7}$$

where x_{ij} is defined in model 1, w_i is a random effect or a "frailty" with gamma distribution, that is,

$$w_i \overset{iid}{\sim} Gamma\left(\phi^{-1};\phi^{-1}\right) \tag{2.8}$$

for $i = 1,\ldots,n$.

The random effect or "frailty" w_i is structured to accommodate the correlation among the repeated counting data and the extra-Poisson variability. Observe that $E(w_i) = 1$ and $Var(w_i) = \phi$. This "frailty" structure is related to additive gamma "frailty" models introduced in the literature by Korsgaard & Andersen (1998); Petersen (1998) & Li (2002).

As the result,

$$E\left(Y_{ij}|\alpha_j,\beta_j,x_{ij}\right) = E\left[E\left(Y_{ij}|\alpha_j,\beta_j,w_i,x_{ij}\right)\right],$$

we have,

$$E\left(Y_{ij}|\alpha_j,\beta_j,x_{ij}\right) = \alpha_j e^{\beta_j x_{ij}} E\left(e^{w_i}\right),$$

that is,

$$E\left(Y_{ij}|\alpha_j,\beta_j,x_{ij}\right) = \alpha_j e^{\beta_j x_{ij}}, \tag{2.9}$$

since $E(w_i) = 1$ for $i = 1,\ldots,n$. As the result,

$$Var\left(Y_{ij}|\alpha_j,\beta_j,x_{ij}\right) = Var\left[E\left(Y_{ij}|\alpha_j,\beta_j,w_i,x_{ij}\right)\right] + E\left[Var\left(Y_{ij}|\alpha_j,\beta_j,w_i,x_{ij}\right)\right],$$

we have,

$$Var\left(Y_{ij}|\alpha_j,\beta_j,x_{ij}\right) = \alpha_j^2 e^{2\beta_j x_{ij}} Var\left(w_i\right) + \alpha_j e^{\beta_j x_{ij}} E\left(w_i\right),$$

that is,

$$Var\left(Y_{ij}|\alpha_j,\beta_j,x_{ij}\right) = \phi\alpha_j^2 e^{2\beta_j x_{ij}} + \alpha_j e^{\beta_j x_{ij}}. \tag{2.10}$$

From (2.9) and (2.10), we observe that the extra-Poisson variability is given by

$$\phi\alpha_j^2 e^{2\beta_j x_{ij}},$$

where the parameter ϕ is related to the extra-Poisson variability.

Observe that model 2, a mixture of a Poisson with a gamma distribution, results in a generalization of a negative binomial distribution (see for example, Bernardo & Smith, 1995, p. 119) for the unconditional distribution for Y_{ij}, $i = 1,\ldots,n$; $j = 1,\ldots,k$.

For model 1, the unconditional distribution for Y_{ij} is obtained from a mixture of a Poisson with a log-normal distribution which is different of the obtained unconditional distribution for Y_{ij} assuming model 2.

2.3. **Model 3.** Another model which generalizes model (2.7) is given by

$$\lambda_{ij} = \left(\sum_{l=1}^{r} \alpha_l w_{li}\right) \exp\left(\beta_{1j} + \beta_{2j} x_{ij}\right), \tag{2.11}$$

where $i = 1, \ldots, n$; $j = 1, \ldots, k$. In this case we have that $\sum_{l=1}^{r} \alpha_l \exp\left(\beta_{1j}\right)$ measures the grooming mean in the j^{th} time for the rats of Wistar group; $\sum_{l=1}^{r} \alpha_l \exp\left(\beta_{1j} + \beta_{2j}\right)$ measures the grooming mean in the j^{th} time for the rats of War group; β_{2j} is a regression parameter indicating the effect of the group species, this model is an additive gamma "frailty" model.

A special case of model (2.11) is given considering $r = 2$, that is,

$$\lambda_{ij} = (\alpha_1 w_{1i} + \alpha_2 w_{2i}) \exp\left(\beta_{1j} + \beta_{2j} x_{ij}\right), \tag{2.12}$$

where w_{1i} and w_{2i} are random effects or "frailties" assumed to be independent with gamma distributions,

$$\begin{aligned} w_{1i} &\overset{iid}{\sim} Gamma\left(\phi_1^{-1}; \phi_1^{-1}\right) \\ w_{2i} &\overset{iid}{\sim} Gamma\left(\phi_2^{-1}; \phi_2^{-1}\right) \end{aligned} \tag{2.13}$$

for $i = 1, 2, \ldots, n$.

As,

$$E\left(Y_{ij}|\alpha_{1j}, \alpha_{2j}, \beta_j, x_{ij}\right) = E\left[E\left(Y_{ij}|\alpha_{1j}, \alpha_{2j}, \beta_j, x_{ij}, w_{1i}, w_{2i}\right)\right],$$

we have

$$E\left(Y_{ij}|\alpha_{1j}, \alpha_{2j}, \beta_j, x_{ij}\right) = \alpha_1 e^{\beta_{1j}+\beta_{2j}x_{ij}} E\left(w_{1i}\right) + \alpha_2 e^{\beta_{1j}+\beta_{2j}x_{ij}} E\left(w_{2i}\right).$$

Since $E\left(w_{1i}\right) = 1$ and $E\left(w_{2i}\right) = 1$, we have,

$$E\left(Y_{ij}|\alpha_{1j}, \alpha_{2j}, \beta_j, x_{ij}\right) = \left(\alpha_1 + \alpha_2\right) e^{\beta_{1j}+\beta_{2j}x_{ij}}. \tag{2.14}$$

Also as the result,

$$\begin{aligned} Var\left(Y_{ij}|\alpha_{1j}, \alpha_{2j}, \beta_j, x_{ij}\right) &= Var\left[E\left(Y_{ij}|\alpha_{1j}, \alpha_{2j}, \beta_j, x_{ij}, w_{1i}, w_{2i}\right)\right] + \\ &\quad + E\left[Var\left(Y_{ij}|\alpha_{1j}, \alpha_{2j}, \beta_j, x_{ij}, w_{1i}, w_{2i}\right)\right], \end{aligned}$$

we have,

$$\begin{aligned} Var\left(Y_{ij}|\alpha_{1j}, \alpha_{2j}, \beta_j, x_{ij}\right) &= Var\left(\alpha_1 e^{\beta_{1j}+\beta_{2j}x_{ij}} w_{1i} + \alpha_2 e^{\beta_{1j}+\beta_{2j}x_{ij}} w_{2i}\right) + \\ &\quad + E\left(\alpha_1 e^{\beta_{1j}+\beta_{2j}x_{ij}} w_{1i} + \alpha_2 e^{\beta_{1j}+\beta_{2j}x_{ij}} w_{2i}\right). \end{aligned}$$

Since $E\left(w_{1i}\right) = E\left(w_{2i}\right) = 1$, $Var\left(w_{1i}\right) = \phi_1$ and $Var\left(w_{2i}\right) = \phi_2$, we have,

$$Var\left(Y_{ij}|\alpha_{1j}, \alpha_{2j}, \beta_j, x_{ij}\right) = \phi_1 \alpha_1^2 e^{2\left(\beta_{1j}+\beta_{2j}x_{ij}\right)} + \phi_2 \alpha_2^2 e^{2\left(\beta_{1j}+\beta_{2j}x_{ij}\right)} + \left(\alpha_1 + \alpha_2\right) e^{\beta_{1j}+\beta_{2j}x_{ij}}. \tag{2.15}$$

From (2.14) and (2.15), we observe that the extra-Poisson variability is given by

$$\phi_1\alpha_1^2 e^{2(\beta_{1j}+\beta_{2j}x_{ij})} + \phi_2\alpha_2^2 e^{2(\beta_{1j}+\beta_{2j}x_{ij})},$$

for $i = 1,\ldots,n$ and $j = 1,\ldots,k$.

Assuming model 3, the unconditional distribution for Y_{ij} results from a mixture of a Poisson distribution with a linear combination of gamma distributions. This is a new modeling approach and could give more flexibility of fit for longitudinal counting data in the presence of covariates.

The use of a linear combination of latent variables also has been considered in Tg. AC transgenic mouse bioassays (see for example, Dunson & Herring, 2005), where the mice have an oncogene inserted and the susceptibility to tumorigenesis is studied. In these studies it is well known that there is extra-Poisson variability in the number of tumors per animal, and the latent variables have transgene interpretations.

Observe that to have identifiability in model 3, some of the loading parameters α_l should be constrained; an alternative is to use appropriate informative prior distributions. Usually, the choice for the frailty multipliers α_l in some applications as genetic studies have biological interpretations (see, Dunson & Herring, 2005); in other applications, these choices are not so simple.

Models 1 and 2 do not have identifiability problems and are reasonable simple to get the posterior summaries of interest using standard MCMC methods. However, we explore the use of model 3 as an alternative to get better fit for the counting data set introduced in Tables 1 and 2.

Other existing more sophisticated modeling for the latent variables could be used to analyze the counting data set (see for example, Chib et al., 1998).

3. A Bayesian Analysis

Assuming the Poisson model (2.1), the likelihood function for $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_k)$ and $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_k)$ given the observed data Y_{ij}, the non-observed variables w_i and the covariates x_{ij}, $i = 1,\ldots,n$; $j = 1,\ldots,k$, is given by,

$$L(\boldsymbol{\alpha},\boldsymbol{\beta}) = \prod_{i=1}^{n}\prod_{j=1}^{k}\frac{e^{-\lambda_{ij}}\lambda_{ij}^{y_{ij}}}{y_{ij}!} \tag{3.1}$$

where λ_{ij}, depending on the model, is given in (2.2), (2.7) or (2.12). That is,

$$L(\boldsymbol{\alpha},\boldsymbol{\beta}) \propto \exp\left(-\sum_{i=1}^{n}\sum_{j=1}^{k}\lambda_{ij}\right)\prod_{i=1}^{n}\prod_{j=1}^{k}\lambda_{ij}^{y_{ij}}. \tag{3.2}$$

3.1. A Bayesian Analysis For Model 1.

For the first stage of a hierarchical Bayesian analysis for the model 1, let us assume the following prior distributions for the parameters α_j and β_j

$$\begin{aligned}&\alpha_j \sim Gamma(a;b);\ a,\ b \text{ known};\\ &\beta_j \sim N\left(c;d^2\right);\ c,\ d \text{ known};\end{aligned} \tag{3.3}$$

for $j = 1, \ldots, k$; also observe that the "frailties" w_i are assumed to be independent random variables with normal distribution $N\left(0, \tau^2\right)$. For a second stage of the hierarchical Bayesian analysis, let us assume an inverse gamma distribution for τ^2, that is,

$$\tau^2 \sim IG\,(f, g)\,;\ f,\ g \text{ known.} \tag{3.4}$$

We further assume prior independence among the parameters. The joint posterior distribution for $\boldsymbol{\alpha}, \boldsymbol{\beta}$, $\mathbf{w}$ and τ^2 is proportional to,

$$\begin{aligned}\pi\left(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{w}, \tau^2 | \mathbf{y}, \mathbf{x}\right) \propto & \prod_{j=1}^{k} \alpha_j^{a-1} e^{-b\alpha_j} \times \prod_{j=1}^{k} \exp\left[-\tfrac{1}{2d^2}\left(\beta_j - c\right)^2\right] \times \\ & \times \prod_{i=1}^{n} \tfrac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\tfrac{w_i^2}{2\tau^2}\right) \times \left(\tau^2\right)^{-(f+1)} \exp\left(-\tfrac{g}{\tau^2}\right) \times \\ & \times \exp\left(-\sum_{i=1}^{n}\sum_{j=1}^{k} \lambda_{ij}\right) \prod_{i=1}^{n}\prod_{j=1}^{k} \lambda_{ij}^{y_{ij}}\end{aligned} \tag{3.5}$$

were λ_{ij} is as in (2.2), $\mathbf{w} = (w_1, \ldots, w_n)$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)$, $\mathbf{y}$ is the vector of counting data and $\mathbf{x}$ is the vector of covariates.

Posterior summaries of interest are obtained through MCMC methods Gelfand & Smith (1990); Smith & Roberts (1993). The conditional posterior distributions needed for the Gibbs sampling algorithm are given in Appendix A. A great simplification is obtained using the software Winbugs (Spiegelhalter et al., 1995) which requires only the specification of the distribution for the data and the prior distributions for the parameters.

It is important to point out that we could incorporate the dependence of β_j assuming other prior distributions; in this way, we could model β_j by a time series process.

3.2. **A Bayesian Analysis For Model 2.** For the first stage of a hierarchical Bayesian analysis under model 2, let us assume the same prior distributions (3.3) for α_j and β_j, $j = 1, \ldots, k$. Under model 2, the "frailties" w_i are assumed to be independent random variables with gamma distribution $Gamma\left(\phi^{-1}; \phi^{-1}\right)$. For a second stage of the hierarchical Bayesian analysis, let us assume a gamma distribution for ϕ, that is,

$$\phi \sim Gamma\,(f; g)\,;\ f,\ g \text{ known.} \tag{3.6}$$

We also assume prior independence among the parameters. The joint posterior distribution for $\boldsymbol{\alpha}, \boldsymbol{\beta}$, $\mathbf{w}$ and ϕ is given by,

$$\begin{aligned}\pi\left(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{w}, \phi | \mathbf{y}, \mathbf{x}\right) \propto & \prod_{j=1}^{k} \alpha_j^{a-1} e^{-b\alpha_j} \times \prod_{j=1}^{k} \exp\left[-\tfrac{1}{2d^2}\left(\beta_j - c\right)^2\right] \times \\ & \times \prod_{i=1}^{n} \tfrac{\phi^{-\phi^{-1}}}{\Gamma\left(\phi^{-1}\right)} w_i^{\phi^{-1}-1} \exp\left(-\phi^{-1} w_i\right) \times \phi^{f-1} \exp\left(-g\phi\right) \times \\ & \times \exp\left(-\sum_{i=1}^{n}\sum_{j=1}^{k} \lambda_{ij}\right) \prod_{i=1}^{n}\prod_{j=1}^{k} \lambda_{ij}^{y_{ij}}\end{aligned} \tag{3.7}$$

were λ_{ij} is given in (2.7) for $i = 1, \ldots, n$; $j = 1, \ldots, k$. The conditional posterior distributions needed for the Gibbs sampling algorithm are given in Appendix A.

3.3. **A Bayesian Analysis For Model** 3 **Assuming** $r = 2$**.** Assuming a special case of model 3 with $r = 2$, where λ_{ij} is given in (2.12), let us assume normal prior distribution for β_{1j} and β_{2j}, that is,

$$\begin{aligned} \beta_{1j} &\sim N\left(\beta_{1j}^{*}, d_1^2\right) \\ \beta_{2j} &\sim N\left(\beta_{2j}^{*}, d_2^2\right) \end{aligned} \tag{3.8}$$

with d_1 and d_2 known.

Since we have identifiability problems for the parameters α_l, $l = 1, 2$ in model 3, we consider a Bayesian analysis in two steps: in the first step, we assume model 2 for the counting data. In the second step of the Bayesian analysis, we assume as prior information (an empirical Bayesian approach) for the choice of the hyperparameters β_{1j}^{*} and β_{2j}^{*} in (3.8), the estimated posterior means for α_j and β_j in model 2 obtained from the first step. That is, denoting by $\hat{\alpha}_j$ and $\hat{\beta}_j$ the estimated posterior means for α_j and β_j in model 2, we assume $\beta_{1j}^{*} = e^{\hat{\alpha}_j}$ and $\beta_{2j}^{*} = \hat{\beta}_j$.

Different prior distributions could be assumed for the loading parameters α_l, $l = 1, 2$. We assume a beta prior distribution, that is,

$$\begin{aligned} \alpha_1 &\sim Beta\,(a_1; b_1)\,;\, a_1,\ b_1 \text{ known;} \\ \alpha_2 &\sim Beta\,(a_2; b_2)\,;\, a_2,\ b_2 \text{ known;} \end{aligned} \tag{3.9}$$

for $j = 1, \ldots, k$; for model 3, the "frailties" w_{1i} and w_{2i} are assumed independent random variables with $Gamma\left(\phi_l^{-1}; \phi_l^{-1}\right)$ distributions (2.13) for $l = 1, 2$. For a second stage of the hierarchical Bayesian analysis, let us assume gamma prior distributions for ϕ_l, that is,

$$\phi_l \sim Gamma\,(f_l; g_l)\,;\, f_l,\ g_l \text{ known; } l = 1, 2. \tag{3.10}$$

We also assume prior independence among the parameters. The joint posterior distribution for α_1, α_2, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, ϕ_1, ϕ_2, $\mathbf{w}_1$ and $\mathbf{w}_2$ is given by,

$$\begin{aligned} \pi\left(\alpha_1,\alpha_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\mathbf{w}_1,\mathbf{w}_2,\phi_1,\phi_2|\mathbf{y},\mathbf{x}\right) \propto\ & \tfrac{\Gamma(a_1+b_1)}{\Gamma(a_1)\Gamma(b_1)}\alpha_1^{a_1-1}\left(1-\alpha_1\right)^{b_1-1} \\ & \times \tfrac{\Gamma(a_2+b_2)}{\Gamma(a_2)\Gamma(b_2)}\alpha_2^{a_2-1}\left(1-\alpha_2\right)^{b_2-1} \times \prod_{j=1}^{k}\exp\left[-\tfrac{1}{2d_1^2}\left(\beta_{1j}-\beta_{1j}^{*}\right)^2\right] \\ & \times \prod_{j=1}^{k}\exp\left[-\tfrac{1}{2d_2^2}\left(\beta_{2j}-\beta_{2j}^{*}\right)^2\right] \times \prod_{i=1}^{n}\tfrac{\phi_1^{-\phi_1^{-1}}}{\Gamma\left(\phi_1^{-1}\right)}w_{1i}^{\phi_1^{-1}-1}\exp\left(-\phi_1^{-1}w_{1i}\right) \\ & \times \prod_{i=1}^{n}\tfrac{\phi_2^{-\phi_2^{-1}}}{\Gamma\left(\phi_2^{-1}\right)}w_{2i}^{\phi_2^{-1}-1}\exp\left(-\phi_2^{-1}w_{2i}\right) \times \phi_1^{f_1-1}\exp\left(-g_1\phi_1\right) \\ & \times \phi_2^{f_2-1}\exp\left(-g_2\phi_2\right) \times \exp\left(-\sum_{i=1}^{n}\sum_{j=1}^{k}\lambda_{ij}\right)\prod_{i=1}^{n}\prod_{j=1}^{k}\lambda_{ij}^{y_{ij}} \end{aligned} \tag{3.11}$$

where λ_{ij} is given in (2.12) for $i = 1, \ldots, n$; $j = 1, \ldots, k$. The conditional posterior distributions needed for the Gibbs sampling algorithm are given in Appendix A.

Similar results are obtained considering $r > 2$ in (2.11); in this way, we could consider different fixed values for r and to chose the best model using some existing Bayesian discrimination criterion.

4. Analysis of the Grooming Counting Data

For a Bayesian analysis of the data set of Tables 1 and 2, we assume for models 1 and 2, the following hyperparameters values for the prior distributions (3.3), (3.4) and (3.6): $a = b = 0.01$; $c = 0$; $d^2 = 1000$; $f = g = 0.1$. For model 3, we assume $a_1 = a_2 = b_1 = b_2 = 1.0$; $d_1^2 = d_2^2 = 1.0$; $f_1 = f_2 = g_1 = g_2 = 1.0$ for the hyperparameters of the prior distributions (3.8), (3.9) and (3.10). This choice for the hyperparameters values was motivated to have approximately noninformative prior distributions and also to have convergence for the MCMC algorithm used for simulation of Gibbs sample for the joint posterior distribution of interest using the Winbugs software Spiegelhalter et al. (1995). The codes of the Winbugs program are given in Appendix B. For models 1, 2, and 3, we simulated 1005000 samples, where the first 5000 samples ("burn-in-samples") were discarded to eliminate the effect of the initial values for the Gibbs sampling algorithm. To have approximately uncorrelated Gibbs sample, we considered the samples $100^a, 200^a, 300^a, ...$, which results in final samples of size 10000 for each parameter. Convergence of the Gibbs sampling algorithm was monitored by usual time series plots for the simulated samples and also using some existing convergence methods Gelman & Rubin (1992).

The posterior summaries for the parameters of the models 1, 2, and 3 are given in Figures 1 to 3. In Figure 4, we have the sample variances for the counting data in each combination time × treatment and Monte Carlo estimates for the posterior means of the variances in each combination time × treatment considering models 1, 2 and 3. From the results of Figure 4, we observe, in general, a better fit of model 3 for the data set of Tables 1 and 2, since the estimated variances are more close to the obtained sample variances.

For model selection, we can use some existing adequacy measures as the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2000). Smaller values of DIC indicates better models. In Table 3, we have the estimated DIC for each model obtained using the Winbugs software. We observe that model 3 is better fitted by the counting data of Tables 1 and 2 (smaller values of DIC). We also observe that the differences among the estimated variance and the sample variances are, in general, smaller considering model 3. In Table 4, we have the sum of squares of these differences assuming each proposed model. From the results of Table 4, we observed smaller sum of squares of these differences for model 3, especially for the Wistar group ($x = 0$); when compared to models 1 and 2.

Table 3. DIC Criterion.

Model	DIC	Number of parameters
Model 1	2016.790	49
Model 2	2016.360	49
Model 3	2011.480	52

In this way, we assume model 3 to get other Bayesian inferences of interest for the counting data of Tables 1 and 2. To verify treatment effect for the Wistar group of rats, in each time, we considered the parameters,

$$\theta_k = (\alpha_1 + \alpha_2)\, e^{\beta_{1,k+12}} - (\alpha_1 + \alpha_2)\, e^{\beta_{1,k}},$$

TABLE 4. Sum of squares for the differences between the estimated variances and the sample variances.

X	Model 1	Model 2	Model 3
0	5073.575	5907.703	4982.026
1	3354.934	4398.451	3452.206

were $k = 1, 2, \ldots, 12$; we also considered the model 3, to verify treatment effect for the War group of rats in each time, considering the parameters

$$\eta_k = (\alpha_1 + \alpha_2)\, e^{\beta_{1,k+12}+\beta_{2,k+12}} - (\alpha_1 + \alpha_2)\, e^{\beta_{1,k}+\beta_{2,k}},$$

were $k = 1, 2, \ldots, 12$.

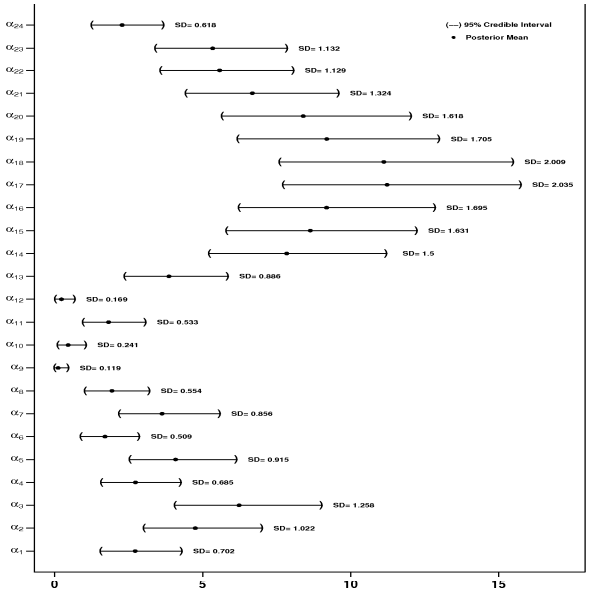
Monte Carlo estimates for θ_k and η_k, $k = 1, 2, \ldots, 12$ considering the 10000 generated Gibbs samples are given in Table 5. We observe that we have significant treatment effects for times $k = 2, 4, 5, 6, 7, 8, 9, 10, 11, 12$ for the Wistar group of rats since zero is not included in the 95% credible intervals for each θ_k. Considering the War group of rats, we observe significant treatment effects for times $k = 2, 3, 5, 6, 7, 8, 10, 11, 12$.

TABLE 5. Posterior summaries considering model 3

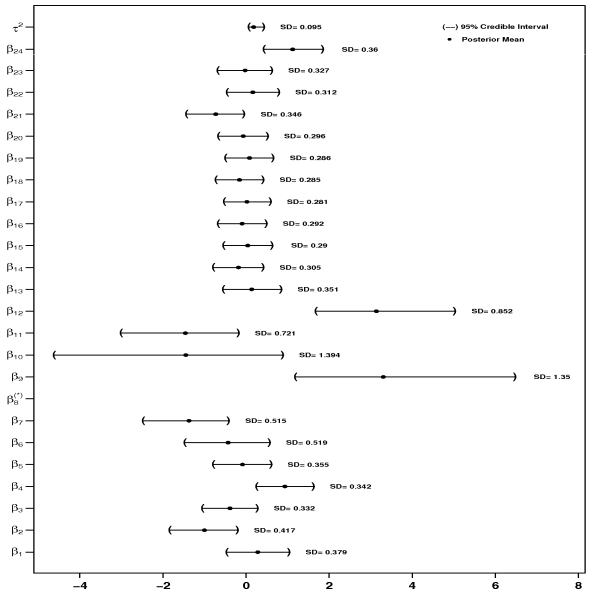
Parameter	Mean	SD	95% Credible Interval	Parameter	Mean	SD	95% Credible Interval
θ_1	1.300	0.964	(−0.524; 3.302)	η_1	0.889	1.252	(−1.506; 3.448)
θ_2	3.486	1.423	(0.824; 6.472)	η_2	5.171	1.507	(2.604; 8.513)
θ_3	2.713	1.528	(−0.160; 5.841)	η_3	5.166	1.818	(2.041; 9.158)
θ_4	7.347	1.677	(4.455; 11.050)	η_4	1.601	1.726	(−1.671; 5.126)
θ_5	8.078	1.868	(4.836; 12.160)	η_5	8.384	2.196	(4.690; 13.320)
θ_6	10.660	2.017	(7.274; 15.150)	η_6	9.076	2.048	(5.764; 13.700)
θ_7	6.287	1.587	(3.511; 9.704)	η_7	9.794	2.146	(6.319; 14.630)
θ_8	7.343	1.582	(4.620; 10.880)	η_8	8.531	1.827	(5.588; 12.680)
θ_9	7.456	1.421	(5.105; 10.550)	η_9	1.546	1.014	(−0.313; 3.745)
θ_{10}	5.762	1.214	(3.742; 8.451)	η_{10}	6.945	1.589	(4.379; 10.510)
θ_{11}	3.975	1.143	(1.973; 6.419)	η_{11}	5.135	1.312	(2.986; 8.113)
θ_{12}	2.316	0.658	(1.234; 3.802)	η_{12}	3.016	1.551	(0.237; 6.405)

5. CONCLUDING REMARKS AND DISCUSSION

Longitudinal counting data in the presence of one or more covariates are very common, especially in medical studies. Usually we need models to capture the correlation among the counting data and the presence of superdispersion. Different "frailty" models are introduced in the literature to analyze counting Poisson data. The use of hierarchical Bayesian methods is a suitable way to analyze Poisson longitudinal data, especially using recent software to simulate samples for the joint posterior distribution of interest. In this way, the use of Winbugs software gives a great simplification to obtain the posterior summaries of interest.



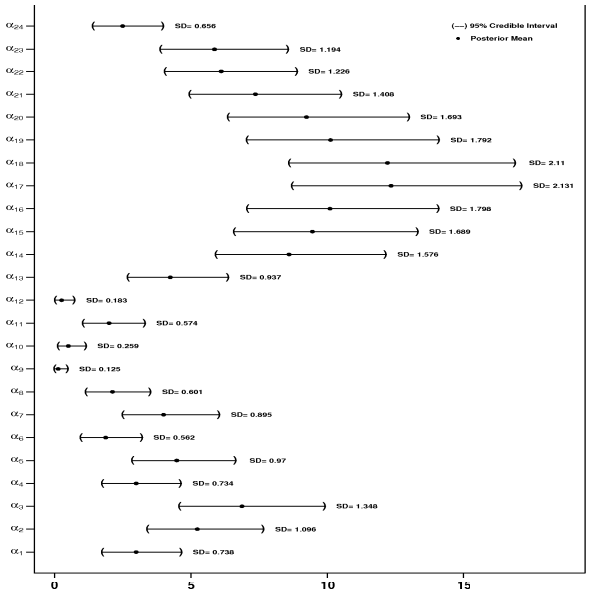
(a) α_j



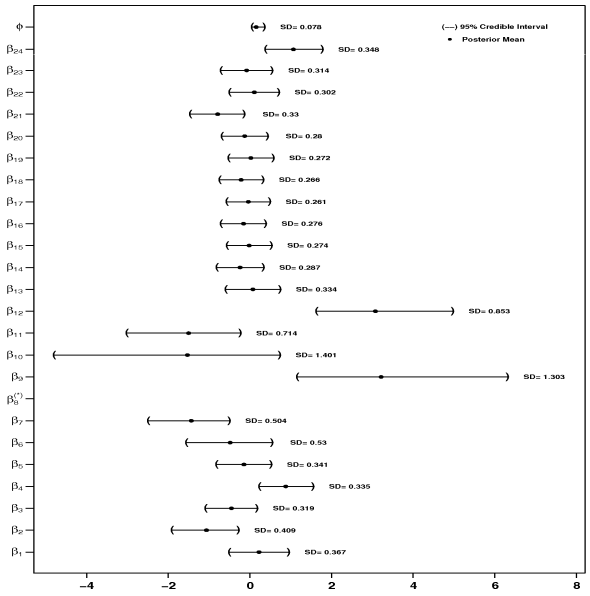
(b) β_j and τ^2

*Posterior mean, 95% credible interval and SD for β_8 are given, respectively, by −27.280, (−73.69; −3.539) and 18.78

FIGURE 1. Graphics for the posterior summaries assuming model 1.



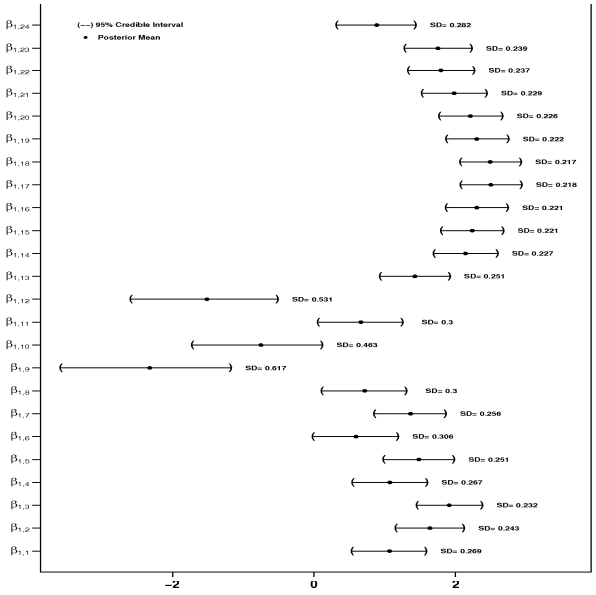
(a) α_j



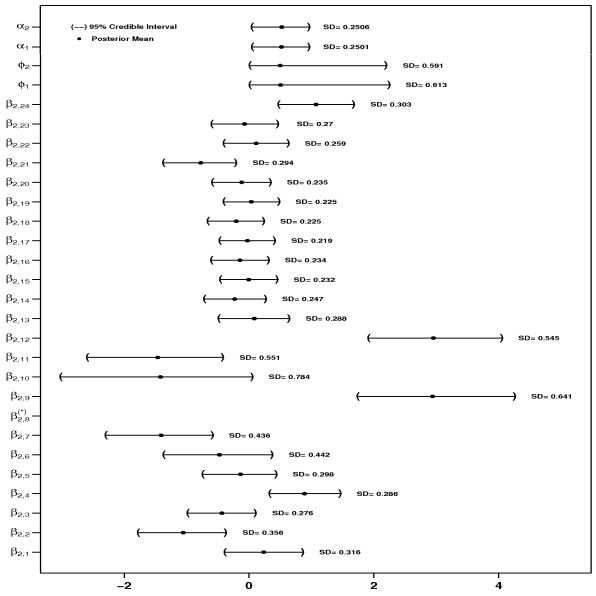
(b) β_j and ϕ

*Posterior mean, 95% credible interval and SD for β_8 are given, respectively, by −27.15, (−72.16; −3.688) and 18.51

FIGURE 2. Graphics for the posterior summaries assuming model 2.



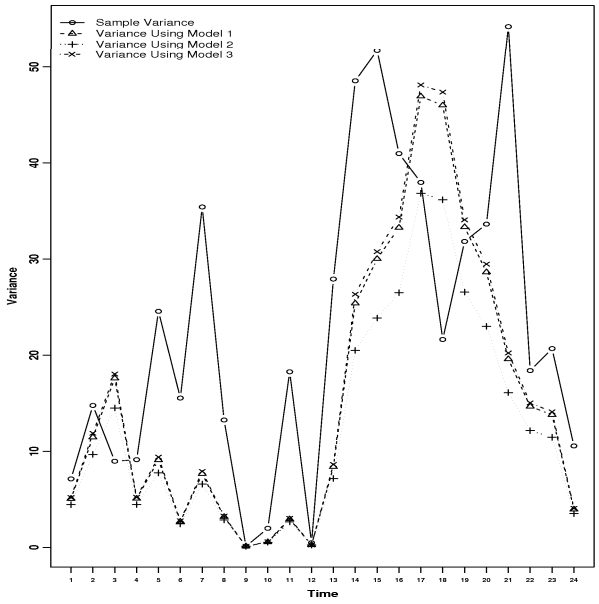
(a) $\beta_{1,j}$



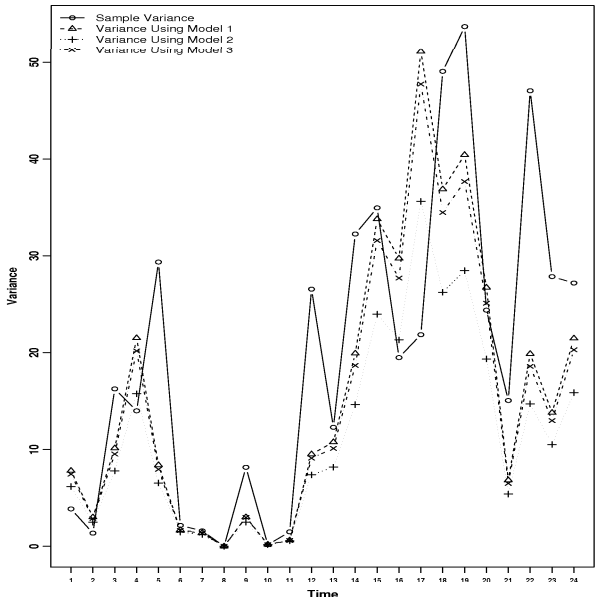
(b) $\beta_{2,j}$; ϕ_l and α_l

*Posterior mean, 95% credible interval and SD for $\beta_{2,8}$ are given, respectively, by -27.15, $(-29.09; -25.20)$ and 0.989

FIGURE 3. Graphics for the posterior summaries assuming model 2.



(a) $x = 0$



(b) $x = 1$

FIGURE 4. Bayesian estimators for the variance of counting data in each time

To analyze the counting data set introduced in Tables 1 and 2, we considered three models with different "frailty" structures. Assuming models 1 and 2, given in sections 2.1 and 2.2, respectively, we observed very similar DIC values (close to 2016). That is, using this discrimination criterion, we can not say that one of these two models is better fitted by the data. However, if our goal is related to model the variances of the counting data, model 1 is better fitted by the data (see Table 4).

Under the DIC criterion, we observe that model 3 is better fitted by the data, since the estimated DIC is given by 2011.48 (smaller than the DIC values for model 1 and 2). We also observe that in terms of estimated variances, model 3 gives similar results as compared to model 1 (see Table 4), considering $X = 0$ or $X = 1$. Other discrimination criterion also could be used to compare the three models (see for example, Gelfand & Ghosh, 1998).

In the data analysis considered as an example, we observed better fit for model 3 (an additive "frailty" model) introduced in section 2 assuming $r = 2$. Possibly, better fit could be obtained considering model 3 with $r > 2$. It is important to point out that expert opinion could be considered for the choice of the hyperparameters of the prior distributions (3.8) and (3.9), usually assuming biological interpretation Dunson & Herring (2005). Other possibility is to use prior information to fix one of the α_l, $l = 1, 2$ (see (3.8)) in the Bayesian analysis (or one of the ϕ_l, $l = 1, 2$ given in (3.9)). In this way, we could obtain better inference results. Further research should be done in this direction.

Other important positive aspect of the Bayesian methodology is related to the discrimination of the proposed models, and the possible use of informative prior distributions considering the opinion of experts, which is common in medical studies.

Appendix

Appendix A: Conditional Posterior Distributions Needed For The Gibbs Sampling Algorithm .

Model 1.

(*i*)**:**

$$\alpha_j|\boldsymbol{\alpha}_{(j)},\boldsymbol{\beta},\mathbf{w},\tau^2,\mathbf{y},\mathbf{x} \sim Gamma\left(a+\sum_{i=1}^{n} y_{ij}, b+\sum_{i=1}^{n} e^{w_i}e^{\beta_j x_{ij}}\right);$$

where $\boldsymbol{\alpha}_{(j)} = \left(\alpha_1,\ldots,\alpha_{j-1},\alpha_{j+1},\ldots,\alpha_k\right)$; $j = 1,\ldots,k$.

(*ii*)**:**

$$\pi\left(\beta_j|\boldsymbol{\alpha},\boldsymbol{\beta}_{(j)},\mathbf{w},\tau^2,\mathbf{y},\mathbf{x}\right) \propto N\left(c,d^2\right)\psi_1\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\tau^2,\mathbf{y},\mathbf{x}\right),$$

where

$$\psi_1\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\tau^2,\mathbf{y},\mathbf{x}\right) = \exp\left(-\alpha_j\sum_{i=1}^{n} e^{w_i}e^{\beta_j x_{ij}} + \beta_j\sum_{i=1}^{n} y_{ij}x_{ij}\right); \quad j = 1,\ldots,k.$$

(*iii*)**:**

$$\pi\left(w_i|\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_{(i)},\tau^2,\mathbf{y},\mathbf{x}\right) \propto N\left(0,\tau^2\right)\psi_2\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\tau^2,\mathbf{y},\mathbf{x}\right),$$

where

$$\psi_2\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\tau^2,\mathbf{y},\mathbf{x}\right)=\exp\left[-e^{w_i}\sum_{j=1}^{k}\alpha_j e^{\beta_j x_{ij}}+\sum_{j=1}^{k}w_i y_{ij}\right];\quad i=1,\ldots,n.$$

(*iv*)**:**

$$\tau^2|\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\mathbf{y},\mathbf{x}\sim IG\left(f+\frac{n}{2},g+\frac{1}{2}\sum_{i=1}^{n}w_i^2\right).$$

Observe that we need to use the Metropolis-Hastings algorithm to simulate samples for β_j and w_i.

Model 2.

(*i*)**:**

$$\pi\left(\alpha_j|\boldsymbol{\alpha}_{(j)},\boldsymbol{\beta},\mathbf{w},\tau^2,\mathbf{y},\mathbf{x}\right)\propto\alpha_j^{a-1}e^{-b\beta_j}\psi_1\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\phi,\mathbf{y},\mathbf{x}\right),$$

where

$$\psi_1\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\phi,\mathbf{y},\mathbf{x}\right)=\exp\left(-\alpha_j\sum_{i=1}^{n}w_i e^{\beta_j x_{ij}}+\sum_{i=1}^{n}y_{ij}\ln\left(\alpha_j\right)\right);\quad j=1,\ldots,k.$$

(*ii*)**:**

$$\pi\left(\beta_j|\boldsymbol{\alpha},\boldsymbol{\beta}_{(j)},\mathbf{w},\phi,\mathbf{y},\mathbf{x}\right)\propto N\left(c,d^2\right)\psi_2\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\phi,\mathbf{y},\mathbf{x}\right),$$

where

$$\psi_2\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\phi,\mathbf{y},\mathbf{x}\right)=\exp\left(-\alpha_j\sum_{i=1}^{n}w_i e^{\beta_j x_{ij}}+\beta_j\sum_{i=1}^{n}y_{ij}x_{ij}\right);\quad j=1,\ldots,k.$$

(*iii*)**:**

$$\pi\left(w_i|\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_{(i)},\phi,\mathbf{y},\mathbf{x}\right)\propto Gamma\left(\phi^{-1},\phi^{-1}\right)\psi_3\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\phi,\mathbf{y},\mathbf{x}\right),$$

where

$$\psi_3\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\phi,\mathbf{y},\mathbf{x}\right)=\exp\left[-w_i\sum_{j=1}^{k}\alpha_j e^{\beta_j x_{ij}}+\sum_{j=1}^{k}y_{ij}\ln\left(w_i\right)\right];\quad i=1,\ldots,n.$$

(*iv*)**:**

$$\pi\left(\phi|\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\mathbf{y},\mathbf{x}\right)\propto Gamma\left(f,g\right)\psi_4\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\phi,\mathbf{y},\mathbf{x}\right),$$

where

$$\begin{aligned}\psi_4\left(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w},\phi,\mathbf{y},\mathbf{x}\right) &= \exp\left[-n\phi^{-1}\ln\left(\phi\right)-n\ln\Gamma\left(\phi^{-1}\right)\right]\times\\ &\quad\times\exp\left[\phi^{-1}\sum_{i=1}^{n}\ln\left(w_i\right)-\phi^{-1}\sum_{i=1}^{n}w_i\right];\quad i=1,\ldots,n.\end{aligned}$$

Observe that we need to use the Metropolis-Hastings algorithm to simulate samples for α_j, β_j, w_i and ϕ.

Model 3.

(*i*)**:**

$$\pi(\alpha_l|\boldsymbol{\alpha}_{(l)},\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\mathbf{w}_1,\mathbf{w}_2,\phi_1,\phi_2,\mathbf{y},\mathbf{x}) \propto \alpha_l^{a_l-1}(1-\alpha_l)^{b_l-1}\psi_1(\boldsymbol{\theta}),$$

where $\boldsymbol{\theta}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\mathbf{w}_1,\mathbf{w}_2,\phi_1,\phi_2,\mathbf{y},\mathbf{x})$ and

$$\psi_1(\boldsymbol{\theta}) = \exp\left[-\sum_{i=1}^{n}\sum_{j=1}^{k}\lambda_{ij}\right]\times\prod_{i=1}^{n}\prod_{j=1}^{k}\lambda_{ij}^{y_{ij}};\quad l=1,2$$

where λ_{ij} is given in (2.12).

(*ii*)**:**

$$\pi\left(\beta_{lj}|\boldsymbol{\beta}_{(lj)},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathbf{w}_1,\mathbf{w}_2,\phi_1,\phi_2,\mathbf{y},\mathbf{x}\right) \propto \exp\left[-\frac{1}{2d_l^2}\left(\beta_{lj}-\beta_{lj}^{*}\right)^2\right]\psi_1(\boldsymbol{\theta}),$$

for $l=1,2$; $j=1,\ldots,k$.

(*iii*)**:**

$$\pi(w_{li}|\mathbf{w}_{(li)},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\phi_1,\phi_2,\mathbf{y},\mathbf{x}) \propto w_{li}^{\phi_l^{-1}-1}\exp\left[-\phi_l^{-1}w_{li}\right]\psi_1(\boldsymbol{\theta}),$$

for $i=1,\ldots,n$; $l=1,2$.

(*iv*)**:**

$$\pi\left(\phi_l|\phi_{(l)},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\mathbf{w}_1,\mathbf{w}_2,\mathbf{y},\mathbf{x}\right) \propto \phi_l^{f_l-1}e^{-g_l\phi_l}\psi_2(\boldsymbol{\phi}_l),$$

where

$$\psi_2(\boldsymbol{\phi}_l) = \exp\left[-\frac{n}{\phi_l}\ln\phi_l - n\ln T\left(\phi_l^{-1}\right) - \frac{1}{\phi_l}\sum_{i=1}^{n}\ln w_{li} - \frac{1}{\phi_l}\sum_{i=1}^{n}w_{li}\right],$$

for $l=1,2$; observe that we need to use the Metropolis-Hastings algorithm to simulate samples for all parameters.

Appendix B: Winbugs Codes .

Model 1.

LISTING 1. Main code of the Winbugs program (Model 1).

```
model
  {
    for(i in 1:rats)
      {
        for(j in 1:times)
          {
            y[i,j] ~ dpois(lambda[i,j])
            lambda[i,j] <- alpha[j]*exp(beta[j]*x[i,j]+w[i])
          }
        w[i] ~ dnorm(a1,sigma)
      }
    for(j in 1:times)
      {
```

```
        alpha[j] ~ dgamma(a2,b2)
        beta[j]  ~ dnorm(a3,b3)
      }
    sigma ~ dgamma(a4,b4)
    tau2 <- 1/sigma
  }
```

(ak, bk), $k = 1, \ldots, 4$, represent known hyperparameters.

LISTING 2. Main code of the Winbugs program (Model 2).

Model 2.

```
model
  {
    for(i in 1:rats)
      {
        for(j in 1:times)
          {
            y[i,j] ~ dpois(lambda[i,j])
            lambda[i,j] <- w[i]*alpha[j]*exp(beta[j]*x[i,j])
          }
        w[i] ~ dgamma(pri,pri)
      }
    for(j in 1:times)
      {
        alpha[j] ~ dgamma(a1,b1)
        beta[j]  ~ dnorm(a2,b2)
      }
    pri <-  1/phi
    phi ~ dgamma(a3,b3)
  }
```

(ak, bk), $k = 1, 2, 3$, represent known hyperparameters.

LISTING 3. Main code of the Winbugs program (Model 3).

Model 3.

```
model
  {
    for( i in 1 : rats)
      {
        for( j in 1 : times)
          {
            y[i,j] ~ dpois(lambda[i,j])
            lambda[i,j] <- (alpha1*w1[i]+alpha2*w2[i])*
                           exp(beta1[j]+beta2[j]*x[i,j])
          }
        w1[i] ~ dgamma(pri1,pri1)
```

```
12          w2[i] ~ dgamma(pri2,pri2)
13        }
14      alpha1 ~ dbeta(a1,b1)
15      alpha2 ~ dbeta(a2,b2)
16
17      pri1 <- 1/phi1
18      pri2 <- 1/phi2
19
20      phi1 ~ dgamma(a3,b3)
21      phi2 ~ dgamma(a4,b4)
22
23      beta1[1]    ~  dnorm(a5,b5)
24      beta1[2]    ~  dnorm(a6,b6)
25      beta1[3]    ~  dnorm(a7,b7)
26      beta1[4]    ~  dnorm(a8,b8)
27      beta1[5]    ~  dnorm(a9,b9)
28      beta1[6]    ~  dnorm(a10,b10)
29      beta1[7]    ~  dnorm(a11,b11)
30      beta1[8]    ~  dnorm(a12,b12)
31      beta1[9]    ~  dnorm(a13,b13)
32      beta1[10]   ~  dnorm(a14,b14)
33      beta1[11]   ~  dnorm(a15,b15)
34      beta1[12]   ~  dnorm(a16,b16)
35      beta1[13]   ~  dnorm(a17,b17)
36      beta1[14]   ~  dnorm(a18,b18)
37      beta1[15]   ~  dnorm(a19,b19)
38      beta1[16]   ~  dnorm(a20,b20)
39      beta1[17]   ~  dnorm(a21,b21)
40      beta1[18]   ~  dnorm(a22,b22)
41      beta1[19]   ~  dnorm(a23,b23)
42      beta1[20]   ~  dnorm(a24,b24)
43      beta1[21]   ~  dnorm(a25,b25)
44      beta1[22]   ~  dnorm(a26,b26)
45      beta1[23]   ~  dnorm(a27,b27)
46      beta1[24]   ~  dnorm(a28,b28)
47
48      beta2[1]    ~  dnorm(a29,b29)
49      beta2[2]    ~  dnorm(a30,b30)
50      beta2[3]    ~  dnorm(a31,b31)
51      beta2[4]    ~  dnorm(a32,b32)
52      beta2[5]    ~  dnorm(a33,b33)
53      beta2[6]    ~  dnorm(a34,b34)
54      beta2[7]    ~  dnorm(a35,b35)
55      beta2[8]    ~  dnorm(a36,b36)
```

```
    beta2[9]    ~ dnorm(a37, b37)
    beta2[10]   ~ dnorm(a38, b38)
    beta2[11]   ~ dnorm(a39, b39)
    beta2[12]   ~ dnorm(a40, b40)
    beta2[13]   ~ dnorm(a41, b41)
    beta2[14]   ~ dnorm(a42, b42)
    beta2[15]   ~ dnorm(a43, b43)
    beta2[16]   ~ dnorm(a44, b44)
    beta2[17]   ~ dnorm(a45, b45)
    beta2[18]   ~ dnorm(a46, b46)
    beta2[19]   ~ dnorm(a47, b47)
    beta2[20]   ~ dnorm(a48, b48)
    beta2[21]   ~ dnorm(a49, b49)
    beta2[22]   ~ dnorm(a50, b50)
    beta2[23]   ~ dnorm(a51, b51)
    beta2[24]   ~ dnorm(a52, b52)
}
```

(ak, bk), $k = 1, \ldots, 52$, represent known hyperparameters.

Acknowledgements

The authors would like to thank the referees for their helpful comments and the Neurophysiology and Neuroethology Experimental Laboratory from Medical School of Ribeirão Preto for the data set. Emílio Augusto Coelho Barros and Jorge Alberto Achcar gratefully acknowledge partial financial support from FAPESP/Brazil (grant number 2007/01379-2) and CNPq/Brazil (grant number 300235/2005-5).

References

Albert, J. H. & Chib, S. (1993), 'Bayesian analysis of binary and polychotonious response data', *Journal of the American Statistical Association* **88**, 669–679.

Bernardo, J. M. & Smith, A. F. M. (1995), *Bayesian Theory,* John Wiley & Sons, Chichester.

Chib, S. & Greenberg, E. (1995), 'Understanding the metropolis-hastings algorithm', *The American Statistician* **49**, 327–335.

Chib, S., Greenberg, E. & Winkelmann, R. (1998), 'Posterior simulation an Bayes factors in panel count data models', *Journal of Econometrics* **86**, 33–54.

Clayton, D. G. (1991), 'A Monte Carlo Method for Bayesian Inference in Frailty Models', *Biometrics* **47**, 467–485.

Crouchley, R. & Davies, R. B. (1999), 'A comparasion of population average and random effects models for the analysis of longitudinal count data with baseline information', *Journal of the Royal Statistical Society, A* **162**, 331–347.

Dunson, D. B. (2000), 'Bayesian latent variable models for clustered mixed outcomes', *Journal of the Royal Statistical Society B* **62**, 335–366.

Dunson, D. B. (2003), 'Dynamic latent trait models for multidimensional longitudinal data', *Journal of the American Statistical Association* **98**, 555–563.

Dunson, D. B. & Baird, D. D. (2002), 'A proportional hazard model for incidence and induced remission of disease', *Biometrics* **58**, 71–78.

Dunson, D. B. & Herring, A. H. (2005), 'Bayesian latent variable models for mixed discrete outcomes', *Biostatistics* **6**, 11–25.

Gelfand, A. E. & Ghosh, S. K. (1998), 'Model choice: a minimum posterior predictive loss approach', *Biometrika* **85**, 1–11.

Gelfand, A. E. & Smith, A. F. M. (1990), 'Sampling-based approaches to calculating marginal densities', *Journal of the American Statistical Association* **85**, 398–409.

Gelman, A. & Rubin, B. D. (1992), 'Inference from iterative simulation using multiple sequences', *Statistical Sciences* **4**, 457–511.

Henderson, R. & Shimakura, S. (2003), 'A serially correlated gamma frailty model for longitudinal count data', *Biometrika* **90**, 355–366.

Jorgensen, B., Lundbye-Christiansen, S., Song, P. X. K., & Sun, L. (1999), 'A state space model for multivariate longitudinal count data', *Biometrika* **86**, 169–191.

Korsgaard, I. R. & Andersen, A. H. (1998), 'The additive genetic gamma frailty model', *Scandinavian Journal of Statistics* **25**, 255–269.

Legler, J. M. & Ryan, L. M. (1997). 'Latent variable models for teratogenesis using multiple binary outcomes', *Journal of the American Statistical Association* **92**, 13–20.

Li, H. Z. (2002), 'An additive genetic gamma frailty model for linkage analysis of diseases with variable age of owsed using nuclear families', *Lifetime Data Analysis* **8**, 315–334.

Moustaki, I. (1996), 'A latent trait and a latent class model for mixed observed variables', *British Journal of Mathematical and Statistical Psychology* **49**, 313–334.

Moustaki, I. & Knott, M. (2000), 'Generalized latent trait models', *Psychometrika* **65**, 391–411.

Petersen, J. H. (1998), 'An additive frailty model for correlated life times', *Biometrics* **54**, 646–661.

Sammel, M. D., Ryan, L. M. & Legler, J. M. (1997), 'Latent variable models and mixed discrete and continuous outcomes', *Journal of the Royal Statistical Society B* **59**, 667–678.

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

Spiegelhalter, D. J., Best, N. G. & Vander Linde, A. (2000), 'A bayesian measure of model complexity and fit (with discussion)', *Journal of the Royal Statistical Society B* **64**, 583–639.

Spiegelhalter, D. J., Thomas, A., Best, N. G.& R., G. W. (1995), *BUGS: Bayesian Inference Using Gibbs Sampling, Version 0.50,* MRC Biostatistics Unit, Cambridge.

Yakovlev, A. Y. & Tsodikov, A. D. (1996), *Stochastic models of tumor latency and their biostatistical applications,* World Scientific, Singapore.

Departamento de Medicina Social, Faculdade de Medicina de Ribeirão Preto, 14048-900, Ribeirão Preto, SP, Brasil

E-mail address, Emílio A. Coelho-Barros: eacbarros@gmail.com