REPEATED MEASURE DATA MODELING ASSUMING A FINITE MIXTURE OF NORMAL DISTRIBUTIONS FOR THE ERROR: A BAYESIAN ANALYSIS

Jorge Alberto ACHCAR[1]
Edson Zangiacomi MARTINEZ[1]
Margaret de CASTRO[2]

▪ *ABSTRACT: In this paper we introduce a Bayesian analysis of repeated measure data using MCMC (Markov Chain Monte Carlo) methods to obtain the posterior summaries of interest. We assume a finite mixture of normal distributions for the error, which gives better fit for the longitudinal data in the presence of covariates. We illustrate the proposed methodology considering a real data set introduced by Castro et al. (2003) related to a dose-response study with different dosages of dexamethasone (dex) to assess the cordicotropic resistance in Cushing disease (CD) using salivary Cortisol.*

▪ *KEYWORDS: Bayesian analysis; MCMC methods; repeated measures; cushing disease.*

1 Introduction

Longitudinal data are of great interest in the analysis of clinical trials. In this paper, we consider a Bayesian analysis for a repeated measure model applied to a real data set introduced by Castro et al. (2003). This data set is related to a dose-response study with different dosages of dexamethasone (dex) to assess the cordicotropic resistance in Cushing disease (CD) using salivary cortisol. The salivary cortisol experiment also was compared to plasma cortisol determination using dex suppression test (DST) and after oCRH (ovine corticotropin-releasing hormone) test in the differential diagnosis of Cushing syndrome (CS).

[1]Departamento de Medicina Social, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo – USP, CEP: 14049-900, Ribeirão Preto, SP, Brazil. E-mail: *achcar@fmrp.usp.br / edson@fmrp.usp.br*
[2]Departamento de Clínica Médica, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo – USP, CEP: 14049-900, Ribeirão Preto, SP, Brazil. E-mail: *castrom@fmrp.usp.br*

Cushing syndrome is a disease caused by an excess of cortisol production or by excessive use of cortisol or other similar steroid (glucocorticoid) hormones. Cortisol is a normal hormone produced in the outer portion, or cortex, of the adrenal glands, located above each kidney. The normal function of cortisol is to help the body respond to stress and change. It mobilizes nutrients, modifies the body's response to inflammation, stimulates the liver to raise the blood sugar, and it helps control the amount of water in the body. Cortisol production is regulated by adrenocorticotrophic hormone (ACTH), made in the pituitary gland, which is located just below the brain. When too much cortisol is produced in the adrenal glands, or an excess is taken in treating other diseases, significant changes occur in all of the tissues and organs of the body. All of these effects together are called Cushing Syndrome.

The main goal of this study was to verify the performance of salivary cortisol after different dosages of dex compared to plasma cortisol and plasma ACTH determination in patients with ACTH - secreting pituitary tumors.

In order to compare the dose-response of salivary cortisol, plasma cortisol and ACTH to different doses of dex, we have in Table 1 (data set introduced by Castro et al., 2003), the percentages of suppression of plasma ACTH, plasma cortisol and salivary cortisol for 18 patients with CD evaluated before and after oral DST: 2 mg/day (0.5 mg every 6 hours for 2 days); 8 mg/day (2 mg every 6 hours for 2 days) and 24 mg/day (6 mg every 6 hours for 1 day), administered sequentially and always in ascending order of doses.

In this paper, we consider the use of Bayesian methods to analyse repeated measure data based on MCMC (Markov Chain Monte Carlo) methods as the Gibbs sampling algorithm (see for example, Gelfand and Smith, 1990) to obtain the posterior summaries of interest. In Section 2, we describe the model assuming a finite mixture of normal distributions for the error. In Section 3, we introduce a Bayesian analysis considering a normal distribution for the error of the repeated measure data model and finally in Sections 4 and 5 we discuss the results for the Cushing disease data of Table 1.

2 The model

Let x_{lki} be the observed values of the response variables for group l ($l = 1, ..., L$), repeated measure or treatment k for the same individual ($k = 1, ..., T$) and individual i ($i = 1, ..., n$) assuming the model

$$x_{lki} = \beta_k + \gamma_l + \delta_{kl} + \omega_i + \epsilon_{lki}, \tag{1}$$

where β_k denotes the effect of doses of dex, γ_l denotes the effect of groups (% salivary cortisol, % plasma cortisol, % ACTH), δ_{kl} denotes the interaction group x treatment and ω_i is a random effect which captures the correlation among the repeated measures within the same individual. Here the random effect ω_i is assumed to be normal, mean 0 and variance σ^2_ω, with independence between subjects.

Table 1 - Cushing disease data (Castro et al., 2003)

	% salivary cortisol doses of dex			% plasma cortisol doses of dex			% ACTH doses of dex		
	2	8	24	2	8	24	2	8	24
1	0.13030	0.04727	0.04727	0.40541	0.06081	0.05743	0.41620	0.37151	0.20810
2	0.70270	0.15135	0.11351	2.10810	1.18920	0.36486	0.26004	0.28999	0.22052
3	0.24038	0.17308	0.09038	1.08570	0.68571	0.50000	0.53771	0.27095	0.29749
4	1.28000	0.06240	0.06240	0.91176	0.05882	0.04412	1.90950	0.97835	0.60236
5	1.13330	0.26667	0.10400	1.39130	0.33152	0.11413	1.99140	1.39220	0.58621
6	0.39070	0.39070	0.08372	0.73059	0.87671	0.19178	0.61860	0.45581	0.45581
7	0.23478	0.06783	0.06783	0.70455	0.10227	0.06818	0.64207	0.16728	0.30996
8	0.45946	0.04216	0.04216	1.20000	0.11000	0.06000	0.49724	0.23204	0.30939
9	0.74359	0.07436	0.10769	1.04000	0.36000	0.60000	1.92170	1.28310	1.74700
10	0.63636	0.45454	0.30909	0.49537	0.27778	0.20833	0.85333	0.05356	1.24440
11	1.16070	0.11786	0.14286	0.85366	0.41463	0.11585	0.40340	0.40097	0.33718
12	0.51136	0.05341	0.03864	0.33333	0.05208	0.02917	0.87649	0.34035	0.15096
13	0.39167	0.25000	0.20000	0.44000	0.24000	0.19000	0.13684	0.13684	0.14316
14	1.68420	0.37895	0.08947	1.93100	0.96552	0.19828	0.89231	0.40000	0.30769
15	0.29231	0.12000	0.12000	0.94737	0.06579	0.09649	0.55000	0.14857	0.12857
16	0.22727	0.03546	0.03546	0.56897	0.11207	0.09052	1.49000	1.16670	0.40000
17	0.33766	0.15584	0.10130	0.97500	0.27500	0.21250	1.83650	0.45673	1.07690
18	0.38462	0.30000	0.30000	0.46154	0.30769	0.30769	3.55240	1.13290	1.00000

The error terms $\epsilon_{lki,}$ are independently distributed with a finite mixture of normal distributions with density

$$\varphi\left(\epsilon_{lki}\right)=\sum_{j=1}^{J}p_j\phi_j\left(\epsilon_{lki}\left|\mu_j,\sigma_j^2\right.\right), \tag{2}$$

where $\sum_{j=1}^{J}p_j=1$, $\sum_{j=1}^{J}\mu_j p_j=0$ and $\phi_j\left(\epsilon_{lki}\left|\mu_j,\sigma_j^2\right.\right)$ denotes a normal density $N\left(\mu_j,\sigma_j^2\right)$ for the error ϵ_{lki}. To avoid the well recognized identifiability problem, we can impose an ordering constraint on the means $(\mu_1<\mu_2<...<\mu_J)$ (see, for example, Mergersen and Robert, 1996; Carroll, Roeder and Wasserman, 1999; Titterington, Smith and Makov, 1986, or Roeder and Wasserman, 1997).

The assumption of a mixture of normal distributions for the error term gives more flexibility of fit for the data in comparison with the standard assumption of normal errors (see for example, Roeder and Wasserman, 1997). This assumption also implies in more robust inferences in the presence of one or more discordant observations (outliers).

For a Bayesian analysis of the data of Table 1 (Cushing disease data), we assume a special case of (2) considering $J=2$ (a mixture of two normal distributions). In this case, the likelihood function for $\boldsymbol{\theta}'=(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{\omega},p_1,p_2,\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\sigma_\omega^2)$ where $\boldsymbol{\beta}'=(\beta_1,...,\beta_T)$, $\boldsymbol{\gamma}'=(\gamma_1,...,\gamma_L)$, $\boldsymbol{\delta}'=(\delta_{11},...,\delta_{TL})$ and $\boldsymbol{\omega}'=(\omega_1,...,\omega_n)$ is given by

$$L\left(\boldsymbol{\theta}\right)=\prod_{l=1}^{L}\prod_{k=1}^{T}\prod_{i=1}^{n}\left[\sum_{j=1}^{2}p_j\phi_j\left(\epsilon_{lki}\left|\mu_j,\sigma_j^2\right.\right)\right], \tag{3}$$

where $\epsilon_{lki}=x_{lki}-\beta_k-\gamma_l-\delta_{kl}-\omega_i$. Observe that $p_2=1-p_1$ and $\mu_2=-\mu_1p_1/p_2$.

Assuming prior independence among the parameters, consider the following prior densities for β_k, γ_l, δ_{kl}, p_1, μ_1, μ_2, σ_1^2 and σ_2^2:

$$\begin{array}{rcll}
\beta_k & \sim & N(0,a_k^2), & a_k \text{ known, } k=1,...,T,\\
\gamma_l & \sim & N(0,b_l^2), & b_l \text{ known, } l=1,...,L,\\
\delta_{kl} & \sim & N(0,c_{kl}^2), & c_{kl} \text{ known,} k=1,...,T,\ l=1,...,L,\\
p_1 & \sim & Beta(d,e), & d,e \text{ known,}\\
\mu_1 & \sim & N(f_1,g_1^2), & f_1,g_1 \text{ known,}\\
\mu_2 & \sim & N(f_2,g_2^2), & f_2,g_2 \text{ known,}\\
\sigma_1^2 & \sim & IG(f_3,g_3), & f_3,g_3 \text{ known,}\\
\sigma_2^2 & \sim & IG(f_4,g_4), & f_4,g_4 \text{ known,}\\
\sigma_\omega^2 & \sim & IG(f_\omega,g_\omega), & f_\omega,g_\omega \text{ known,}
\end{array} \tag{4}$$

where $IG(\cdot,\cdot)$ denotes an inverse gamma distribution (see Casella and Berger, 2002, p.51), $N(\cdot,\cdot)$ denotes a normal distribution and $Beta(\cdot,\cdot)$ denotes a beta

distribution. The joint posterior distribution for $\boldsymbol{\theta}^* = \left(\boldsymbol{\theta}, \sigma_\omega^2\right)$ is given by

$$\begin{aligned}\pi\left(\boldsymbol{\theta}^*|\mathbf{x}\right) \quad \propto \quad & \exp\left(-\frac{1}{2}\sum_{k=1}^{T}\frac{\beta_k^2}{a_k^2}\right)\exp\left(-\frac{1}{2}\sum_{l=1}^{L}\frac{\gamma_l^2}{b_l^2}\right)\exp\left(-\frac{1}{2}\sum_{k=1}^{T}\sum_{l=1}^{L}\frac{\delta_{kl}^2}{c_{kl}^2}\right)\times \\ & \times\left(\sigma_\omega^2\right)^{-\frac{n}{2}}\exp\left(-\frac{1}{2\sigma_\omega^2}\sum_{i=1}^{n}\omega_i^2\right)p_1^{d-1}\left(1-p_1\right)^{e-1}\times \\ & \times\exp\left[-\frac{1}{2g_1^2}\left(\mu_1-f_1\right)^2\right]\exp\left[-\frac{1}{2g_2^2}\left(\mu_2-f_2\right)^2\right]\left(\sigma_1^2\right)^{-(f_3+1)}\times \\ & \times\exp\left(-\frac{g_3}{\sigma_1^2}\right)\left(\sigma_2^2\right)^{-(f_4+1)}\exp\left(-\frac{g_4}{\sigma_2^2}\right)\times \\ & \times\left(\sigma_\omega^2\right)^{-(f_\omega+1)}\exp\left(-\frac{g_\omega}{\sigma_\omega^2}\right)L\left(\boldsymbol{\theta}'\right), \end{aligned} \tag{5}$$

where $L\left(\boldsymbol{\theta}\right)$ is given in (3) and $\mathbf{x}$ is the observed matrix of data values.

Observe that, independent of the adopted prior distribution, the term $\prod_{l=1}^{L}\prod_{k=1}^{T}\prod_{i=1}^{n}\left[\sum_{j=1}^{2}p_j\phi_j\left(\epsilon_{lki}\left|\mu_j,\sigma_j^2\right.\right)\right]$ prevent us from expressing the conditional densities as a product of independent components. To simplify the joint posterior and the full conditional distributions needed for the Gibbs sampling algorithm, we introduce artificial variables (see, for example, Tanner and Wong, 1987) $\boldsymbol{\upsilon}_{lki} = (\upsilon_{lki1}, \upsilon_{lki2})$ where $l = 1, ..., L$, $k = 1, ..., T$ and $i = 1, ..., n$ for each observation. In this way, we assume that the conditional distribution for υ_{lki1} given $\boldsymbol{\theta}$ and x has a Bernoulli distribution with success probability

$$h_{lki1} = \frac{p_1\phi_1\left(\epsilon_{lki}\left|\mu_1,\sigma_1^2\right.\right)}{\sum_{j=1}^{2}p_j\phi_j\left(\epsilon_{lki}\left|\mu_j,\sigma_j^2\right.\right)}. \tag{6}$$

That is,

$$\pi\left(\upsilon_{lki1}\right) \propto h_{lki1}^{\upsilon_{lki1}}\left(1-h_{lki1}\right)^{1-\upsilon_{lki1}}. \tag{7}$$

Noting that $\upsilon_{lki1} + \upsilon_{lki2} = 1$, we have $\upsilon_{lki1} = 1$ with probability h_{lki1} and $\upsilon_{lki1} = 0$ with probability $1 - h_{lki1}$, or $\upsilon_{lki2} = 1$ with probability $1 - h_{lki1}$ and $\upsilon_{lki2} = 0$ with probability h_{lki1}. Thus,

$$\pi\left(\upsilon_{1111}, ..., \upsilon_{LTn1}\right) \propto \prod_{l=1}^{L}\prod_{k=1}^{T}\prod_{i=1}^{n}\frac{\prod_{j=1}^{2}p_j\phi_j\left(\epsilon_{lki}\left|\mu_j,\sigma_j^2\right.\right)^{\upsilon_{lki1}}}{\sum_{j=1}^{2}p_j\phi_j\left(\epsilon_{lki}\left|\mu_j,\sigma_j^2\right.\right)}. \tag{8}$$

Combining equation (5) with (8), we have the joint posterior distribution for $\boldsymbol{\theta}^*$ and $\boldsymbol{\upsilon}$ given the observed $\mathbf{x}$ matrix of data values as a product of independent components. That is,

$$\pi\left(\upsilon_{1111}, ..., \upsilon_{LTn1}, \boldsymbol{\theta}^*|\mathbf{x}\right) \propto \pi\left(\boldsymbol{\theta}^*\right)\prod_{l=1}^{L}\prod_{k=1}^{T}\prod_{i=1}^{n}\prod_{j=1}^{2}\left[p_j\phi_j\left(\epsilon_{lki}\left|\mu_j,\sigma_j^2\right.\right)\right]^{\upsilon_{lki1}}. \tag{9}$$

To generate samples of the joint posterior distribution (9) for the case of $J = 2$ components, we use the Gibbs sampling algorithm. Starting with initial values $\boldsymbol{\theta}^{*(0)} = \left(\boldsymbol{\theta}_1^{*(0)}, ..., \boldsymbol{\theta}_p^{*(0)}\right)$, we follow the steps:

(i) Generate a sample $\boldsymbol{\upsilon}^{(1)} = \left(\upsilon_{111}^{(1)}, ..., \upsilon_{LTn}^{(1)}\right)$ from equation (7).

(ii) Generate a sample of $\boldsymbol{\theta}^*$ from the full conditional distributions $\pi(\boldsymbol{\theta}_1^* | \boldsymbol{\theta}_2^{*(0)}, ..., \boldsymbol{\theta}_p^{*(0)}, \boldsymbol{\upsilon}^{(1)}, \mathbf{x})$, ..., $\pi(\boldsymbol{\theta}_p^* | \boldsymbol{\theta}_1^{*(1)}, ..., \boldsymbol{\theta}_{p-1}^{*(1)}, \boldsymbol{\upsilon}^{(1)}, \mathbf{x})$ (see Cancho, Achcar and Ortega, 2006).

Thus, continue iterations by repeating steps (i) and (ii).

To simulate samples of the joint posterior distribution (5) we could have a great computational simplification using the software WinBugs (Spiegelhalter et al., 1999) which only requires the specification of the prior distributions and the likelihood function for the parameters.

Similarly, we could assume $J > 2$ components in the mixture density (2), considering a Dirichlet prior distribution for p_1, p_2, ..., p_J. In practical work, we could consider different values for J and to choose the best model using standard existing Bayesian discrimination methods (see for example, Smith and Bernardo, 2007).

3 A normal distribution for the error

A standard situation to analyse medical longitudinal data is to assume a normal distribution for the error term ϵ_{lki} in equation (1). Assuming that the error terms ϵ_{lki} in the model (1) are independently distributed as $N(0, \sigma^2)$, the same prior distributions for β_k, γ_l, δ_{kl}, and ω_i given in (4) and the prior distributions $\sigma^2 \sim IG(f, g)$ with f, g known and $\sigma_\omega^2 \sim IG(f_\omega, g_\omega)$ with f_ω and g_ω known, the joint posterior distribution for $\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{\omega}, \sigma^2$ and σ_ω^2 is given by

$$\begin{aligned}
\pi\left(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{\omega}, \sigma^2, \sigma_\omega^2 | \mathbf{x}\right) \quad \propto \quad & \exp\left(-\frac{1}{2}\sum_{k=1}^{T}\frac{\beta_k^2}{a_k^2}\right)\exp\left(-\frac{1}{2}\sum_{l=1}^{L}\frac{\gamma_l^2}{b_l^2}\right)\times \\
& \times \exp\left(-\frac{1}{2}\sum_{k=1}^{T}\sum_{l=1}^{L}\frac{\delta_{kl}^2}{c_{kl}^2}\right)\left(\sigma_\omega^2\right)^{-\frac{n}{2}}\exp\left(-\frac{1}{2\sigma_\omega^2}\sum_{i=1}^{n}\omega_i^2\right)\times \\
& \times \left(\sigma^2\right)^{-(f+1)}\exp\left(-\frac{g}{\sigma^2}\right)\left(\sigma_\omega^2\right)^{-(f_\omega+1)}\exp\left(-\frac{g_\omega}{\sigma_\omega^2}\right)\times \\
& \times \prod_{l=1}^{L}\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^{T}\epsilon_{lki}^2\right), \qquad (10)
\end{aligned}$$

where $\epsilon_{lki} = x_{lki} - \beta_k - \gamma_l - \delta_{kl} - \omega_i$.

Considering the special case of the Cushing disease data of Table 1, with $L = 3$ groups (%salivary, %plasma, %ACTH), $T = 3$ treatments (doses of dex: 2, 8 and

24), the conditional posterior distributions for the parameters needed in the Gibbs sampling algorithm are given by

$$\beta_k \left| \boldsymbol{\theta}^*_{(\beta_k)}, \mathbf{x} \sim N \left(\frac{a_k^2 \sum_{l=1}^{3} \sum_{i=1}^{n} \mu_{lki}^{(1)}}{\sigma^2 + 3na_k^2}, \frac{\sigma^2 a_k^2}{\sigma^2 + 3na_k^2} \right) \right.,$$

$$\gamma_l \left| \boldsymbol{\theta}^*_{(\gamma_l)}, \mathbf{x} \sim N \left(\frac{b_l^2 \sum_{k=1}^{3} \sum_{i=1}^{n} \mu_{lki}^{(2)}}{\sigma^2 + 3nb_l^2}, \frac{\sigma^2 b_l^2}{\sigma^2 + 3nb_l^2} \right) \right.,$$

$$\delta_{kl} \left| \boldsymbol{\theta}^*_{(\delta_{kl})}, \mathbf{x} \sim N \left(\frac{c_{kl}^2 \sum_{i=1}^{n} \mu_{lki}^{(3)}}{\sigma^2 + nc_{kl}^2}, \frac{\sigma^2 c_{kl}^2}{\sigma^2 + nc_{kl}^2} \right) \right.,$$

$$\omega_i \left| \boldsymbol{\theta}^*_{(\omega_i)}, \mathbf{x} \sim N \left(\frac{\sigma_\omega^2 \sum_{l=1}^{3} \sum_{k=1}^{3} \mu_{lki}^{(4)}}{\sigma^2 + 9\sigma_\omega^2}, \frac{\sigma^2 \sigma_\omega^2}{\sigma^2 + 9\sigma_\omega^2} \right) \right.,$$

$$\begin{aligned} \sigma^2 \left| \boldsymbol{\theta}^*_{(\sigma^2)}, \mathbf{x} \right. &\sim IG \left(f + \frac{9n}{2}, g + \frac{1}{2} \sum_{l=1}^{3} \sum_{k=1}^{3} \sum_{i=1}^{n} \epsilon_{lki}^2 \right), \\ \sigma_\omega^2 \left| \boldsymbol{\theta}^*_{(\sigma_\omega^2)}, \mathbf{x} \right. &\sim IG \left(f_\omega + \frac{n}{2}, g_\omega + \frac{1}{2} \sum_{i=1}^{n} \omega_i^2 \right), \end{aligned} \tag{11}$$

where $\mu_{lki}^{(1)} = x_{lki} - \gamma_l - \delta_{kl} - \omega_i$, $\mu_{lki}^{(2)} = x_{lki} - \beta_k - \delta_{kl} - \omega_i$, $\mu_{lki}^{(3)} = x_{lki} - \beta_k - \gamma_l - \omega_i$, $\mu_{lki}^{(4)} = x_{lki} - \beta_k - \gamma_l - \delta_{kl}$, $l = 1, 2, 3$, $k = 1, 2, 3$ and $i = 1, ..., n$; $\boldsymbol{\theta}^*_{(\delta)}$ is the vector of all parameters except δ.

4 An application with the Cushing disease data

For the Cushing disease data of Table 1, let us first assume the model (1) with a normal distribution for the error. Assuming the prior distributions (4) with $a_k^2 = b_l^2 = c_{kl}^2 = 10$, $l = 1, 2, 3$, $k = 1, 2, 3$, $\sigma_\omega^2 \sim IG(0.1, 0.1)$ and $\sigma^2 \sim IG(0.1, 0.1)$, we generated from the conditional posterior distributions (11) 101000 Gibbs samples for the joint posterior distribution (10) The samples were generated using the WinBugs software. From these 101000 samples, we discarded the first 1000 ("burn-in samples") and since the successive realizations of the chain were correlated, we considered a spacing of size 10, obtaining a final sample of size 10000 for each parameter. To monitor the convergence of the chain, we have used

standard graphical methods and the potential scale redution index introduced by Gelman and Rubin (1992). In our case, these values were less than 1.01 indicating the convergence of the chain.

In Table 2, we present Monte Carlo estimates for the posterior means, standard deviations and 95% credible intervals for all parameters based on the $S = 10000$ Gibbs samples. We also have in Table 2 Monte Carlo estimates for the posterior means of $\theta_{lk} = E(X_{lki}) = \beta_k + \gamma_l + \delta_{kl}$, $l = 1, 2, 3$, $k = 1, 2, 3$ It is interesting to observe that from a preliminary analysis of the data of Table 1, we have the sample means $\overline{x}_{11} = 0.608$, $\overline{x}_{12} = 0.1745$, $\overline{x}_{13} = 0.1142$, $\overline{x}_{21} = 0.9210$, $\overline{x}_{22} = 0.3603$, $\overline{x}_{23} = 0.1916$, $\overline{x}_{31} = 1.0770$, $\overline{x}_{32} = 0.5380$ and $\overline{x}_{33} = 0.5290$ ($\overline{x}_{lk}$, l indexes group and k indexes doses of treatment). We observe good fit of the data for the proposed model given in (1).

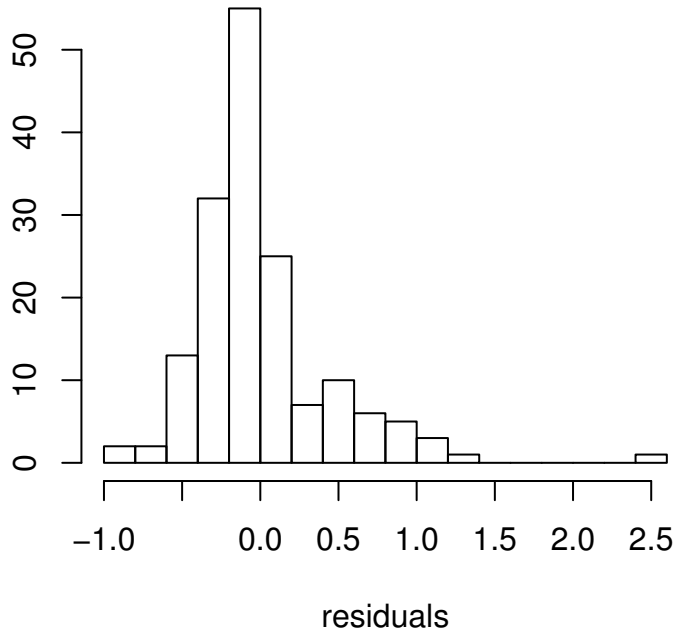

Figure 1 - Histogram of the residuals $\widehat{\epsilon}_{lki} = x_{lki} - \theta_{lk}$.

To compare the profiles of the three groups we have in Table 2, the posterior summaries for the parameters $\lambda_1 = \theta_{11} - \theta_{21}$, $\lambda_2 = \theta_{12} - \theta_{22}$, $\lambda_3 = \theta_{13} - \theta_{23}$ (comparison of salivary with %plasma), $\lambda_4 = \theta_{11} - \theta_{31}$, $\lambda_5 = \theta_{12} - \theta_{32}$, $\lambda_6 = \theta_{13} - \theta_{33}$ (comparison of %salivary with %ACTH), $\lambda_7 = \theta_{21} - \theta_{31}$, $\lambda_8 = \theta_{22} - \theta_{32}$, $\lambda_9 = \theta_{23} - \theta_{33}$ (comparison of %plasma with %ACTH). To compare the effects of the doses of dex, we have the posterior summaries for $\lambda_{10} = \theta_{11} - \theta_{12}$, $\lambda_{11} = \theta_{11} - \theta_{13}$ (%salivary), $\lambda_{12} = \theta_{21} - \theta_{22}$, $\lambda_{13} = \theta_{21} - \theta_{23}$ (%plasma), $\lambda_{14} = \theta_{31} - \theta_{32}$ and $\lambda_{15} = \theta_{31} - \theta_{33}$ (%ACTH). From the results of Table 2, we observe that $\lambda_1 = \theta_{11} - \theta_{21}$, $\lambda_4 = \theta_{11} - \theta_{31}$, $\lambda_5 = \theta_{12} - \theta_{32}$, $\lambda_6 = \theta_{13} - \theta_{33}$, $\lambda_9 = \theta_{23} - \theta_{33}$ are different of zero,

since zero is not included in the 95% credible intervals for λ_1, λ_4, λ_5, λ_6 and λ_9. Also for λ_{10}, λ_{11}, λ_{12}, λ_{13}, λ_{14} and λ_{15}.

In Figure 1, we have the histogram of the residuals $\widehat{\epsilon}_{lki} = x_{lki} - \theta_{lk}$. We observe bimodality for the residuals. Therefore, we consider a reanalysis of the Cushing disease data of Table 1 assuming a mixture of two normal distributions for the error in model (1) and the same prior distributions (4). Assuming the same values for the hyperparameters considered for the case of normal errors, we have in Table 3 the posterior summaries for the quantities of interest based on $S = 10000$ simulated Gibbs samples also with a "burn-in" of size 1000 and taking every 10^{th} observation to eliminate the correlation among the samples (Winbugs code in the Appendix).

We observe from the posterior summaries presented in Tables 2 and 3, similar inference results considering the model (1) with a normal distribution or a mixture of two normal distributions for the error.

To check the overall performance for a model we calculate the sum of squares of the residuals for both models. Assuming the model (1) with a normal distribution for the error, we have $\sum_{l=1}^{3}\sum_{k=1}^{3}\sum_{i=1}^{18}|\widehat{\epsilon}_{lki}| = 47.6267$. Assuming the model (1) with a mixture of two normal distributions for the error , we have $\sum_{l=1}^{3}\sum_{k=1}^{3}\sum_{i=1}^{18}|\widehat{\epsilon}_{lki}| = 47.5675$. That is, the model (1) with a mixture of two normal distributions gives a little improvement in terms of fitting for the Cushing disease data of Table 1. For model selection, we also could use the deviance information criterion (DIC) (Spiegelhalter, 2002). Smaller values of DIC indicates better models. For model 1 (normal error), the Monte Carlo estimate for DIC obtained for the 10000 simulated Gibbs samples is given by 190.247 and for model 2 (mixture of two normal components for the error) is given by 188.146. From these results, we also conclude that model 2 gives a slight improvement for the fit considering the Cushing disease data of Table 1.

Alternatively, we can introduce in the model a correlation structure among the repeated measurements replacing the random effect ω_i in model (1) by ω_{ik}, $i = 1, ..., n$, $k = 1, 2, 3$, where

$$\begin{bmatrix} \omega_{i1} \\ \omega_{i2} \\ \omega_{i3} \end{bmatrix} \sim MN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \sigma^2_{\omega}\begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}\right), \tag{12}$$

with independence between subjects. Here, MN denotes a multivariate normal distribution. We are assuming a first-order autoregressive structure to allow for higher correlation of measurements on the same individuals when measured in adjacent doses of dex. In the Bayesian analysis of the model, we impose $0 < \rho < 1$, assuming a positive correlation among the measurements. Assuming the prior distributions (4) with $a_k^2 = b_l^2 = c_{kl}^2 = 10$, $l = 1, 2, 3$, $k = 1, 2, 3$, $\sigma^2_{\omega} \sim IG(0.1, 0.1)$, $\sigma^2 \sim IG(0.1, 0.1)$ and $\rho \sim Uniform(0, 1)$, we obtained a DIC value of 142.924 for the model with a normal distribution for the error and a DIC value of 166.416 for the model that considers a mixture of normal distributions for the error. The estimates of the parameters of the model considering the first-order autoregressive structure

Table 2 - Posterior summaries – normal distribution for the error

Parameter	posterior mean	s.d.	95% credible interval
β_1	0.0599	1.889	(−3.685 ; 3.718)
β_2	0.2174	1.873	(−3.434 ; 3.892)
β_3	0.3790	1.886	(−3.331 ; 4.085)
γ_1	0.4737	1.861	(−3.176 ; 4.177)
γ_2	0.1087	1.905	(−3.564 ; 3.817)
γ_3	0.0536	1.881	(−3.702 ; 3.779)
δ_{11}	0.0734	2.069	(−3.963 ; 4.152)
δ_{12}	0.0058	2.074	(−4.093 ; 3.984)
δ_{13}	−0.0099	2.058	(−3.993 ; 4.050)
δ_{21}	0.2290	2.061	(−3.855 ; 4.235)
δ_{22}	0.0340	2.069	(−3.981 ; 4.153)
δ_{23}	−0.0781	2.064	(−4.226 ; 3.923)
δ_{31}	0.2237	2.070	(−3.843 ; 4.253)
δ_{32}	0.0493	2.047	(−3.969 ; 4.087)
δ_{33}	0.0941	2.058	(−3.945 ; 4.156)
θ_{11}	0.6071	0.112	(0.387 ; 0.827)
θ_{12}	0.1744	0.113	(−0.050 ; 0.396)
θ_{13}	0.1134	0.113	(−0.107 ; 0.340)
θ_{21}	0.9201	0.110	(0.703 ; 1.139)
θ_{22}	0.3601	0.112	(0.141 ; 0.582)
θ_{23}	0.1929	0.112	(−0.024 ; 0.415)
θ_{31}	1.0760	0.113	(0.853 ; 1.301)
θ_{32}	0.5370	0.112	(0.312 ; 0.754)
θ_{33}	0.5268	0.113	(0.306 ; 0.749)
λ_1	−0.3130	0.134	(−0.575 ; −0.050)
λ_2	−0.1856	0.137	(−0.449 ; 0.081)
λ_3	−0.0794	0.136	(−0.342 ; 0.191)
λ_4	−0.4694	0.135	(−0.735 ; −0.205)
λ_5	−0.3625	0.137	(−0.629 ; −0.090)
λ_6	−0.4133	0.136	(−0.682 ; −0.150)
λ_7	−0.1564	0.135	(−0.423 ; 0.106)
λ_8	−0.1769	0.136	(−0.446 ; 0.088)
λ_9	−0.3339	0.134	(−0.595 ; −0.071)
λ_{10}	0.4327	0.136	(0.170 ; 0.701)
λ_{11}	0.4936	0.136	(0.2266 ; 0.761)
λ_{12}	0.5600	0.134	(0.2942 ; 0.826)
λ_{13}	0.7272	0.135	(0.4611 ; 0.996)
λ_{14}	0.5395	0.137	(0.2726 ; 0.815)
λ_{15}	0.5497	0.137	(0.2863 ; 0.816)

Table 3 - Posterior summaries – mixture of two normal distributions

Parameter	posterior mean	standard.deviation	95% credible interval
β_1	0.0280	1.889	$(-3.674; 3.711)$
β_2	0.2142	1.877	$(-3.459; 3.916)$
β_3	0.3559	1.873	$(-3.243; 4.057)$
γ_1	0.5038	1.881	$(-3.139; 4.199)$
γ_2	0.1129	1.867	$(-3.596; 3.788)$
γ_3	0.0735	1.890	$(-3.565; 3.768)$
δ_{11}	0.0849	2.072	$(-3.931; 4.145)$
δ_{12}	0.0293	2.095	$(-4.051; 4.120)$
δ_{13}	0.0097	2.065	$(-3.979; 4.124)$
δ_{21}	0.2205	2.087	$(-3.835; 4.332)$
δ_{22}	0.0445	2.071	$(-3.993; 4.085)$
δ_{23}	-0.0931	2.083	$(-4.152; 3.984)$
δ_{31}	0.1730	2.085	$(-3.938; 4.250)$
δ_{32}	0.0685	2.053	$(-3.928; 4.121)$
δ_{33}	0.1094	2.043	$(-3.914; 4.180)$
θ_{11}	0.6166	0.110	$(0.400; 0.830)$
θ_{12}	0.1702	0.109	$(-0.045; 0.386)$
θ_{13}	0.1112	0.110	$(-0.105; 0.329)$
θ_{21}	0.9384	0.111	$(0.728; 1.158)$
θ_{22}	0.3716	0.110	$(0.156; 0.587)$
θ_{23}	0.1946	0.112	$(-0.0242; 0.415)$
θ_{31}	1.0330	0.114	$(0.807; 1.254)$
θ_{32}	0.5373	0.112	$(0.319; 0.753)$
θ_{33}	0.5389	0.110	$(0.323; 0.752)$
λ_1	-0.3218	0.133	$(-0.582; -0.061)$
λ_2	-0.2014	0.133	$(-0.462; 0.055)$
λ_3	-0.0834	0.134	$(-0.346; 0.182)$
λ_4	-0.4160	0.136	$(-0.683; -0.148)$
λ_5	-0.3672	0.134	$(-0.632; -0.106)$
λ_6	-0.4277	0.133	$(-0.689; -0.166)$
λ_7	-0.0942	0.137	$(-0.361; 0.178)$
λ_8	-0.1657	0.134	$(-0.429; 0.096)$
λ_9	-0.3443	0.135	$(-0.606; -0.078)$
λ_{10}	0.4465	0.133	$(0.184; 0.711)$
λ_{11}	0.5055	0.133	$(0.241; 0.762)$
λ_{12}	0.5668	0.133	$(0.306; 0.826)$
λ_{13}	0.7438	0.134	$(0.483; 1.005)$
λ_{14}	0.4953	0.138	$(0.224; 0.764)$
λ_{15}	0.4938	0.138	$(0.224; 0.763)$
p_1	0.5987	0.106	$(0.384; 0.796)$
σ_1^2	0.1380	0.022	$(0.101; 0.188)$
σ_2^2	0.2215	0.044	$(0.150; 0.322)$

(12) are similar to those obtained considering the model (1) with the random effect $\omega_i \sim N(0, \sigma^2_\omega)$. Although it is not the objective of this study to explore and identify correlation structures to enhance the fit of the models to the available data, the DIC values suggest that better results are obtained through elicitations of the association among the repeated measurements.

5 Presence of outliers

The use of a mixture of two normal distributions for the error in model (1) gave a little improvement in terms of better fit for the Cushing disease data of Table 1. In this case, we could consider a normal distribution for the error to get accurate Bayesian inference results for the data.

In the presence of outliers, we have more robust inference results using a mixture of normal distributions for the error in model (1). To see this, let us assume the same Cushing disease data of Table 1, but replacing observation $x_{218} = 1.200$ (8^{th} observation in the %plasma cortisol group with a dose 2 of dex) by $x_{218} = 26.000$. That is, we have the presence of an outlier observation. Considering the same values for the hyperparameters of the prior distribution (4) as considered in section 4 and a normal distribution for the error in model (1), we have in Table 4, the posterior summaries for $\theta_{lk} = E(X_{lki}) = \beta_k + \gamma_l + \delta_{kl}$, $l = 1, 2, 3$, $k = 1, 2, 3$. We also have in Table 4, the obtained Monte Carlo estimates for the posterior mean for θ_{lk} assuming a mixture of two normal distribution for the error (1), also considering the same values for the hyperparameters of the prior distributions as considered in section 4.

From the results of Table 4, we observe much better inference results considering the mixture of two normal distributions for the error in model (1). We observe that the posterior means for θ_{lk} are very close to the sample means $\overline{x}_{lk}$, $l = 1, 2, 3$, $k = 1, 2, 3$ (see section 4). Considering the model (1) with a normal distribution for the error, we have large posterior standard deviations for the parameters θ_{lk}. Considering a mixture of two normal distributions for the error, we have accurate results (small standard deviations) for θ_{lk}. We also observe that the calculated sum of squares of the residuals are given by $\sum_{l=1}^{3}\sum_{k=1}^{3}\sum_{i=1}^{18}|\widehat{\epsilon}_{lki}| = 65.6828$ assuming a normal distribution for the error and by $\sum_{l=1}^{3}\sum_{k=1}^{3}\sum_{i=1}^{18}|\widehat{\epsilon}_{lki}| = 49.0939$ assuming a mixture of two normal distributions for the error. That is, considering mixture of normal distributions for the error in model (1), we get more robust Bayesian inference results, especially in the presence of one or more outliers, an usual situation for medical repeated measure data. For the model assuming a normal distribution for the error, the Monte Carlo estimate for DIC is given by 700.615 and for model considering the mixture of two normal components for the error, the estimate for DIC is given by 594.586.

Table 4 - Posterior summaries for θ_{lk}, l, $k = 1, 2, 3$ (presence of a outlier)

	Normal error		Mixture of normals error	
Parameter	posterior mean	s.d.	posterior mean	s.d.
θ_{11}	0.6155	0.4823	0.5907	0.1431
θ_{12}	0.1674	0.4858	0.2073	0.1464
θ_{13}	0.1090	0.4746	0.1435	0.1433
θ_{21}	2.2820	0.4921	0.8347	0.1412
θ_{22}	0.3636	0.4781	0.3469	0.1437
θ_{23}	0.2050	0.4859	0.2045	0.1447
θ_{31}	1.0750	0.4781	1.1560	0.1431
θ_{32}	0.5412	0.4786	0.5228	0.1462
θ_{33}	0.5326	0.4775	0.6169	0.1406

6 Concluding remarks

To analyse repeated measures data, statisticians usually consider standard existing methods based on the Hottelling's T^2 statistics (see, for example, Wichern and Johnson, 2007) assuming independent multivariate normal populations with same variance-covariance matrix. This data structure, usually is not appropriate for real data sets.

The use of Bayesian methods (Smith and Bernardo, 2007) is an useful alternative to analyse repeated measure data in medical studies. The introduction of random effects which captures the possible correlation among the repeated measure data gives a great flexibility of fit. Also, the introduction of mixture of normal distributions for the error gives a robust inference approach to repeated measure data, in the presence of outliers. It is important to point out that the use of MCMC methods gives accurate Bayesian results.

Some other important advanteges of the proposed Bayesian methodology:

(i) It does not depend on asymptotical inference results.

(ii) It is possible to incorporate expert medical opinion in the prior distributions.

(iii) The presence of missing data is easily handed in the Bayesian approach.

(iv) It is not required high computational effort, especially using the software WinBugs.

ACHCAR, J. A.; MARTINEZ, E. Z.; CASTRO, M. Modelagem Bayesiana de medidas repetidas assumindo uma mistura finita de distribuições normais para o erro. *Rev. Mat. Estat.,* São Paulo, v.25, n.3, p.127-143, 2007.

▪ RESUMO: Neste artigo, introduzimos uma análise Bayesiana de medidas repetidas usando métodos MCMC (Monte Carlo em Cadeias de Markov) para a obtenção de medidas resumo de interesse. Assumimos uma mistura finita de distribuições normais para o erro, para um melhor ajuste de dados longitudinais na presença de covariáveis. Para exemplificar a aplicabilidade do modelo proposto, consideramos um banco de dados reais introduzidos por Castro et al (2003) em um estudo dose-resposta com diferentes dosagens de dexamethasone (dex).

▪ PALAVRAS-CHAVE: Análise Bayesiana; métodos MCMC; medidas repetidas; doença de Cushing.

References

CANCHO, V. G.; ACHCAR, J. A.; ORTEGA, E. M. M. Non-linear regression models assuming that the error term has a finite mixture of normal distributions. *Rev. Mat. Estat.*, São Paulo, v.24, n.3, p.35-60, 2006.

CARROLL, J. R.; ROEDER, K., WASSERMAN, L. Flexible parametric measurement error models. *Biometrics*, v.55, n.1, p. 44-54, 1999.

CASELLA, G.; BERGER, R. L. *Statistical inference*. 2nd ed. Pacific Grove: Duxbury, 2002. 660p.

CASTRO, M.; ELIAS, L. L. K.; ELIAS, P. C. L.; MOREIRA, A. C. A dose response study of salivary cortisol after dexamethasone suppression test in Cushing disease and its potential use in the differential diagnosis of Cushing syndrome. *Clin. Endocrinol.*, Oxford, v.59, n.6, p.800-805, 2003.

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

GELMAN, A.; RUBIN, B.D. Inference from iterative simulation using multiple sequences. *Stat. Sci.*, v.4, p.457-511, 1992.

MERGERSEN, K.; ROBERT, C. P. Testing for mixtures: a Bayesian entropic approach. In: BERNARDO, J. M., SMITH, A. F. M., DAWID, A. P., BERGER, J. O. (Ed.). *Bayesian Statistics 4*. New York: Oxford University Press, 1996.

ROEDER, K.; WASSERMAN, L. Practical bayesian density estimation using mixture of normals. *J. Am. Stat. Assoc.*, v.92, n.493, p.894-902, 1997.

SMITH, A. F.M.; BERNARDO, J.M. *Bayesian theory*. 2.ed. Chichester: Wiley, 2007. 640p.

SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; VAN DER LINDE, A. Bayesian measures of model complexity and fit (with discussion). *J. R. Stat. Soc. B*, London, v.64, n.4, p.583-639, 2002.

SPIEGELHALTER, D. J.; THOMAS, A.; BEST, N. G.; GILKS, W. R. *WinBugs version 1.3: Bayesian inference using Gibbs sampling*. Cambridge: MRC Biostatistics Unit., 1999.

TANNER, M. A.; WONG, W. H. The calculation of posterior distributions by data augmentation. *J. Am. Stat. Assoc.*, v.82, n.398, p.528-550, 1987.

TITTERINGTON, D. M.; SMITH, A. F. M.; MAKOV, U. E. *Statistical analysis of finite mixture distributions*. New York: Wiley, 1986. 254p.

WICHERN, D. W.; JOHNSON, R. A. *Applied multivariate statistical analysis*. 6.ed. New Jersey: Prentice Hall, 2007. 800p.

Received in 01.08.2007.

Approved after revised in 22.12.2007.

Appendix: Program in WinBugs

```
# Model with a mixture of two normal distributions for the error

model  {
 for( i in 1 : n ) {  vec[i] ~dcat(p[])
 for( k in 1 : T ) {
 for( l in 1 : L ) {
     x[i,k,l]~dnorm(m[i, k,l,vec[i]], tau[vec[i]])
     m[i,k,l,1] <- beta[k]+gamma[l]+delta[k,l]+omega[i]
     m[i,k,l,2] <- m[i,k,l,1]
     }}}

     p[1:2] ~ddirch(phi[])
     # priors and theta definition:
 for( k in 1 : T ) { beta[k] ~dnorm(0,0.1) }
 for( l in 1 : L ) {  gamma[l] ~dnorm(0,0.1)
 for( k in 1 : T ) {
      delta[k,l] ~dnorm(0,0.1)
      theta[k,l] <- beta[k]+gamma[l]+delta[k,l]
      }}
 for( i in 1 : n ) { omega[i] ~dnorm(0,tau2) }
     tau2 ~dgamma(0.1,0.1)
     tau[1] ~dgamma(0.1,0.1)
     tau[2] ~dgamma(0.1,0.1)
     sigomega <- 1/tau2
     sigma[1] <- 1/tau[1]
     sigma[2] <- 1/tau[2]

     # lambda definition:
     lambda[1] <- theta[1,1]-theta[2,1]
     lambda[2] <- theta[1,2]-theta[2,2]
     lambda[3] <- theta[1,3]-theta[2,3]
     lambda[4] <- theta[1,1]-theta[3,1]
     lambda[5] <- theta[1,2]-theta[3,2]
     lambda[6] <- theta[1,3]-theta[3,3]
     lambda[7] <- theta[2,1]-theta[3,1]
     lambda[8] <- theta[2,2]-theta[3,2]
     lambda[9] <- theta[2,3]-theta[3,3]
     lambda[10] <- theta[1,1]-theta[1,2]
     lambda[11] <- theta[1,1]-theta[1,3]
     lambda[12] <- theta[2,1]-theta[2,2]
     lambda[13] <- theta[2,1]-theta[2,3]
     lambda[14] <- theta[3,1]-theta[3,2]
```

```
    lambda[15] <- theta[3,1]-theta[3,3]
    }
list(n=18,T=3,L=3,phi=c(1,1),vec=c(1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2))
x[,1,1] x[,1,2] x[,1,3] x[,2,1] x[,2,2] x[,2,3] x[,3,1] x[,3,2] x[,3,3]
0.13030 0.04727 0.04727 0.40541 0.06081 0.05743 0.41620 0.37151 0.20810
0.70270 0.15135 0.11351 2.10810 1.18920 0.36486 0.26004 0.28999 0.22052
0.24038 0.17308 0.09038 1.08570 0.68571 0.50000 0.53771 0.27095 0.29749
1.28000 0.06240 0.06240 0.91176 0.05882 0.04412 1.90950 0.97835 0.60236
1.13330 0.26667 0.10400 1.39130 0.33152 0.11413 1.99140 1.39220 0.58621
0.39070 0.39070 0.08372 0.73059 0.87671 0.19178 0.61860 0.45581 0.45581
0.23478 0.06783 0.06783 0.70455 0.10227 0.06818 0.64207 0.16728 0.30996
0.45946 0.04216 0.04216 1.20000 0.11000 0.06000 0.49724 0.23204 0.30939
0.74359 0.07436 0.10769 1.04000 0.36000 0.60000 1.92170 1.28310 1.74700
0.63636 0.45454 0.30909 0.49537 0.27778 0.20833 0.85333 0.05356 1.24440
1.16070 0.11786 0.14286 0.85366 0.41463 0.11585 0.40340 0.40097 0.33718
0.51136 0.05341 0.03864 0.33333 0.05208 0.02917 0.87649 0.34035 0.15096
0.39167 0.25000 0.20000 0.44000 0.24000 0.19000 0.13684 0.13684 0.14316
1.68420 0.37895 0.08947 1.93100 0.96552 0.19828 0.89231 0.40000 0.30769
0.29231 0.12000 0.12000 0.94737 0.06579 0.09649 0.55000 0.14857 0.12857
0.22727 0.03546 0.03546 0.56897 0.11207 0.09052 1.49000 1.16670 0.40000
0.33766 0.15584 0.10130 0.97500 0.27500 0.21250 1.83650 0.45673 1.07690
0.38462 0.30000 0.30000 0.46154 0.30769 0.30769 3.55240 1.13290 1.00000
END
```