Brazilian Journal of Probability and Statistics (2001), 15, pp. 1–14.

A HIERARCHICAL BAYES ANALYSIS OF DATA ON DAILY ACTIVITY CYCLES

Carlos Alberto Ribeiro Diniz and **José Galvão Leite**

Departamento de Estatística, Universidade Federal de São Carlos, CP 676, São Carlos, SP, Brazil. E-mails: dcad@power.ufscar.br and leite@power.ufscar.br

Summary
The classical approach in analysis of data on daily activity cycles is difficult due to the randomness of the daily total group activity and dependency in the samples. The dependence occurs because the daily activities are recorded for the same animal group over the entire study period. A hierarchical Bayes solution to such a type of data is presented. The new approach is applied in a real data set collected to compare daily activity behavior in a group of Black Lion Tamarin, *L. chrysopygus,* in two different annual seasons in Brazil. Gibbs with Metropolis-Hastings algorithms are used in order to determine the posterior distributions of success probabilities of any specific activity at rainy and dry seasons. The marginal posterior distributions were compared using the Kullback-Leibler divergence measure.

Key Words: Activity behavior; Gibbs-with-Metropolis-Hastings algorithms; hierarchical Bayesian analysis; Kullback-Leibler divergence measure.

1 Introduction

Data on daily activity cycles are common in studies of groups of animals in their natural environments. The classical approach in such type of data is difficult due to the randomness of the daily total group activity and the dependence in the samples. This dependence occurs because the daily activities are recorded for the same group of animals over the entire study period. Sussman *et al.* (1979) present an alternative statistical methods for analyzing data on daily activity. They propose a technique based on fewer assumptions than those required by the chi-square test procedure.

In this paper a Bayesian solution to a set of data on daily activity cycles in primate is given. This primate study (Costa, 1997) was conducted in order to compare daily activity behavior in a group of Black Lion Tamarin, *L. chrysopygus,* in two different annual seasonal climatic conditions in Brazil, the rainy and dry seasons. The basic data set was obtained through a systematic observation of a single group of six individual: two adult male, three adult female and an infant female. The group was observed from November 1992 to October 1993, three consecutive days by month, totalling 36 observation days. The sampling technique known as instantaneous scan-sampling (Altmann, 1974) was used. Counts, for each animal, of the number of each of the following six activities were recorded: resting (A1); animal-feeding (A2); plant-feeding (A3); foraging (A4); moving (A5) and not in sight (A6). A fifteen-minute-interval was taken between the counts of an animal to the next one. At the end of each day the total group activity record for activity is collected. The period of observation per day was associated to the awakeness and the retiringness of the group, thus the total group activity record for the ith day in the rainy season, M_i, $i = 1, 2, \ldots, r$, and for the jth day in the dry season, N_j, $j = 1, 2, \ldots, s$, were not fix. The activities are not assumed to be independent and all the 6 animals of the group take part on all activities.

2 Bayesian approach

Initially, let us assume that the main interest is to compare the *resting behavior* of the two seasons. Suppose, in the rainy season, M is the total *group activity* record, p is the probability that an element of the group be in resting behavior and X is the total *group resting* record. In the dry season, N is the total *group activity* record, q is the probability that an element of the group be in resting behavior and Y is the total *group resting* record. Suppose that the random vector (M, N) assumes values in $\mathbb{N}^2$, where $\mathbb{N} = \{0,1,\ldots\}$ and the marginal probability densities of M and N, $f_M(\cdot|\lambda)$ and $f_N(\cdot|\vartheta)$, are indexed by the hyperparameters λ and ϑ, respectively. Given M, p, N, and q, it is assumed that X and Y have binomial distributions with parameters M, p and N, q respectively.

Even the random vectors (X, M) and (Y, N) are dependent, once the study was based on the same animal group, it is possible, assuming the previous assumptions, to write the likelihood functions of those vectors.

2.1 Likelihood functions

The likelihood function in the rainy season, for $m \in \mathbb{N}$ and $x \in \{0, 1, \ldots m\}$, is given by

$$\begin{aligned} L(\lambda, p|x, m) &= P(X = x,\ M = m|\lambda,\ p) \\ &= \binom{m}{x} p^x (1-p)^{m-x} f_M(m|\lambda) I_{(0,1)}(p), \end{aligned} \tag{2.1}$$

and in the dry season, for $n \in \mathbb{N}$ and $y \in \{0, 1, \ldots n\}$, is given by

$$L(\vartheta,\ q|y,n) = \binom{n}{y} q^y (1-q)^{n-y} f_N(n|\vartheta) I_{(0,1)}(q).$$

2.2 Posterior distributions

All statistical inferences in the Bayesian analysis of resting behaviors will be based on the posterior distributions of $p|x, m$ and $q|y, n$.

The joint posterior distribution of $(\lambda,\ p)$ is

$$\pi(\lambda, p|x, m) \propto L(\lambda, p|x, m)\pi(\lambda, p), \tag{2.2}$$

where $\pi(\lambda, p)$ is the joint prior density for (λ, p).

For the primate daily activity cycles p will be the probability of success of the specific activity and, in cases where $f_M(m|\lambda)$ corresponds to the Poisson density, λp will be the average of success of the specific activity. If λ and p are independent a priori, $\pi(\lambda, p) = \pi(\lambda)\pi(p)$, which is natural in this case, then the marginal posterior distribution of p is given by

$$\pi(p|x, m) \propto p^x (1-p)^{m-x} \pi(p) I_{(0,1)}(p). \tag{2.3}$$

Analogously, the marginal posterior distribution of q, assuming that ϑ and q are independent a priori, is

$$\pi(q|y, n) \propto q^y (1-q)^{n-y} \pi(q) I_{(0,1)}(q). \tag{2.4}$$

If p and q have natural conjugate Beta distributions with hyperparameters (a, b) and (c, d), respectively, then the expressions (2.3) and (2.4) become

$$\pi(p|x, m) = \frac{1}{B(a+x, m+b-x)} p^{a+x-1} (1-p)^{m+b-x-1} I_{(0,1)}(p) \tag{2.5}$$

and

$$\pi(q|y, n) = \frac{1}{B(c+y, n+d-y)} q^{c+y-1} (1-q)^{n+d-y-1} I_{(0,1)}(q) \tag{2.6}$$

respectively.

2.3 Selection of priors

The values of the hyperparameters a, b, c, and d are frequently based on the previous knowledge of the experimenter involved in the investigation.

If those values are not available they must be determined from any information about the marginal distribution. Some methods (see Berger, 1985, chapter 3) used to select the priors are, for instance, the type II maximum likelihood prior (ML-II) and the moment approach.

The determination of a ML-II prior in this work will be associated with the class Γ of priors

$$\Gamma = \left\{ \pi : \pi(\lambda, p) = \pi(\lambda) \frac{1}{B(a,b)} p^{a-1}(1-p)^{b-1} I_{(0,1)}(p), a > 0, b > 0 \right\}.$$

Thus, if $\widehat{\pi} \in \Gamma$ satisfies

$$P(X = x, M = m | \widehat{\pi}) = \sup_{(a,b):a>0,b>0} P(X = x, M = m | a, b) \tag{2.7}$$

then $\widehat{\pi}$ is the ML-II prior. Similar class of priors is used for $(\vartheta,\, q)$.

The predictive distribution of (X, M) given a and b is given by

$$\begin{aligned} P(X &= x,\ M = m|a,b) = \int_0^1 \int_0^\infty P(X = x,\ M = m|\lambda,\ p)\pi(\lambda, p) d\lambda dp \\ &\propto \int_0^1 \int_0^\infty p^x (1-p)^{m-x} f_M(m|\lambda) \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)} \pi(\lambda) d\lambda dp \\ &\propto \frac{B(a+x, m-x+b)}{B(a,b)}. \end{aligned} \tag{2.8}$$

and the predictive distribution of (Y, N) given c and d is given by

$$P(Y = y,\ N = n|c,d) \propto \frac{B(c+y, n-y+d)}{B(c,d)}. \tag{2.9}$$

In order to determine a and b in (2.5), considering the class Γ, it is necessary to find a and b that satisfy (2.7). It is easy to show that there are no a and b values such that the function (2.8) can be maximized. It indicates that the ML-II approach can not be used in this case. The moment approach could also not be used, without subjective experience, since the data set consist of just one value for X, the total *group resting* record in the rainy season. Similar argument can be used about c and d.

2.4 Hierarchical models

An alternative approach is to model such a problem hierarchically. That is, given a and b, $a > 0$ and $b > 0$, if λ and p are independent a priori,

λ and the vector (a, b) are independent and p can be modelled by a beta distribution with hyperparameters a and b, then

$$\begin{aligned} \pi(\lambda, p|a, b) &= \pi(\lambda)\pi(p|a, b) \\ &= \pi(\lambda)\frac{1}{B(a, b)}p^{a-1}(1-p)^{b-1}I_{(0,1)}(p). \end{aligned}$$

Suppose that $\pi(a, b) \propto \exp\{-(a+b)/k\}$. This prior put uniform probability on $a/(a+b)$, the mean of the beta distribution with hyperparameters a and b (George, 1992). Different values of k were used in this study and the posterior results showed insensitive to all of them. For this reason $k = 10$ was used in the analysis.

The posterior density for all the parameters (λ, p, a, b) is

$$\begin{aligned} \pi(\lambda, p, a, b|x, m) &\propto L(\lambda, p|x, m)\pi(\lambda, p|a, b)\pi(a, b) \\ &\propto \frac{1}{B(a, b)}p^{a+x-1}(1-p)^{b+m-x-1}f_M(m|\lambda)\pi(\lambda)\pi(a, b)I_{(0,1)}(p). \end{aligned}$$

The posterior density for the parameters (p, a, b) is given by

$$\pi(p, a, b|x, m) \propto \frac{1}{B(a, b)}p^{a+x-1}(1-p)^{b+m-x-1}\exp\{-(a+b)/k\}I_{(0,1)}(p) \tag{2.10}$$

and for the parameters (a, b) is given by

$$\pi(a, b|x, m) \propto \frac{B(a+x, b+m-x)}{B(a, b)}\exp\{-(a+b)/k\}. \tag{2.11}$$

The marginal posterior densities of the parameters p, a, and b in equation (2.10) are not easily obtained. An alternative to handle situations where the integration is difficult or impossible to be calculated exactly is the use of the Gibbs with Metropolis-Hastings algorithm, which allows one to simulate observations from a complex joint distribution.

In this study the posterior simulation is performed in part using the Gibbs sampler algorithm (Casella and George, 1992), for the cases where the conditional posterior distributions of the parameters are available for sampling, and in part using the Metropolis-Hastings algorithm (Chib and Greenberg, 1995), for the cases where the conditional distribution are not of standard form.

All results in this section and in the next subsection can automatically be used to simulate from the full conditionals of q, c, and d.

2.4.1 The conditional posterior distributions of the parameters

The conditional posterior density for p given a, b, x and m is

$$\pi(p|a,b,x,m) \propto p^{a+x-1}(1-p)^{b+m-x-1} I_{(0,1)}(p). \qquad (2.12)$$

The expression in (2.12) is readily recognized to be the kernel of a beta distribution with parameters $a+x$ and $b+m-x$.

The conditional posterior density for a given b, p, x and m is

$$\pi(a|b,p,x,m) \propto \frac{\Gamma(a+b)}{\Gamma(a)} p^{a} \exp\{-a/k\} I_{(0,\infty)}(a) \qquad (2.13)$$

and the conditional posterior density for b given a, p, x and m is

$$\pi(b|a,p,x,m) \propto \frac{\Gamma(a+b)}{\Gamma(b)} p^{b} \exp\{-b/k\} I_{(0,\infty)}(b). \qquad (2.14)$$

Observe that it is almost impossible to simulate from the densities (2.13) and (2.14). Thus in order to generate the samples from a and b we use the Metropolis-Hastings algorithm.

3 Data analysis

The Bayesian methods outlined in the previous sections are now applied to a set of data on daily activity cycles of a group of Black Lion Tamarin, in two different annual seasonal climatic conditions in Brazil. Counts of six activities, for each animal, were recorded. At the end of each day the total group activity record for activity was collected (see the original data set in Table 4).

Characteristics of the posterior distributions of the parameters p a, and b (q, c, and d) can be calculated from the samples generated by the Gibbs with Metropolis-Hastings technique. A Pentium II 333 MHZ and the statistical software SAS were used to generate two chains of 10,000 iterations for the three parameters. The first 3000 iterations were ignored and then every $5th$ sample run was taken forward for analysis.

The posterior distributions of p and q are summarized in Table 1, with characteristics such as the posterior means and the posterior medians (the averages and the medians of the simulated values), and the 95% credible intervals. The activity *not in sight* was not relevant for the research.

Table 1
Mean, median, sd and HPD interval for p and q.

Seasonal	Activity	mean	median	sd	0.95 HPD interval
rainy	resting - p	0.3860	0.3860	0.0056	(0.3750; 0.3970)
dry	resting - q	0.3670	0.3670	0.0106	(0.3470; 0.3880)
rainy	animal-feeding -p	0.0111	0.0111	0.0012	(0.0088; 0.0136)
dry	animal-feeding -q	0.0078	0.0076	0.0020	(0.0044; 0.0121)
rainy	plant-feeding - p	0.1120	0.1120	0.0037	(0.1050; 0.1190)
dry	plant-feeding - q	0.1200	0.1200	0.0071	(0.1060; 0.1340)
rainy	foraging - p	0.0783	0.0782	0.0031	(0.0723; 0.0845)
dry	foraging - q	0.0955	0.0952	0.0065	(0.0831; 0.1090)
rainy	moving - p	0.1280	0.1280	0.0039	(0.1210; 0.1360)
dry	moving - q	0.1410	0.1410	0.0076	(0.1270; 0.1560)

4 Convergence diagnosis

The CODA software (Best, *et al.*, 1995) was used to perform convergence disgnostics of the chains. At least four criteria available in that package were used and the results showed that the chains have converged. One of them, the Gelman and Rubin criterion (Gilks *et al.*, 1995), is presented in Table 2. Histograms of the combined chains for each activity are showed in Figure 1.

Table 2
Gelman and Rubin 50% and 97.5% shrink factors

Variable	Point est.	97.5% quantile
resting - p	1.00	1.00
resting - q	1.02	1.08
animal-feeding -p	1.00	1.00
animal-feeding -q	1.00	1.00
plant-feeding - p	1.00	1.00
plant-feeding - q	1.00	1.00
foraging - p	1.00	1.00
foraging - q	1.00	1.00
moving - p	1.00	1.00
moving - q	1.00	1.00

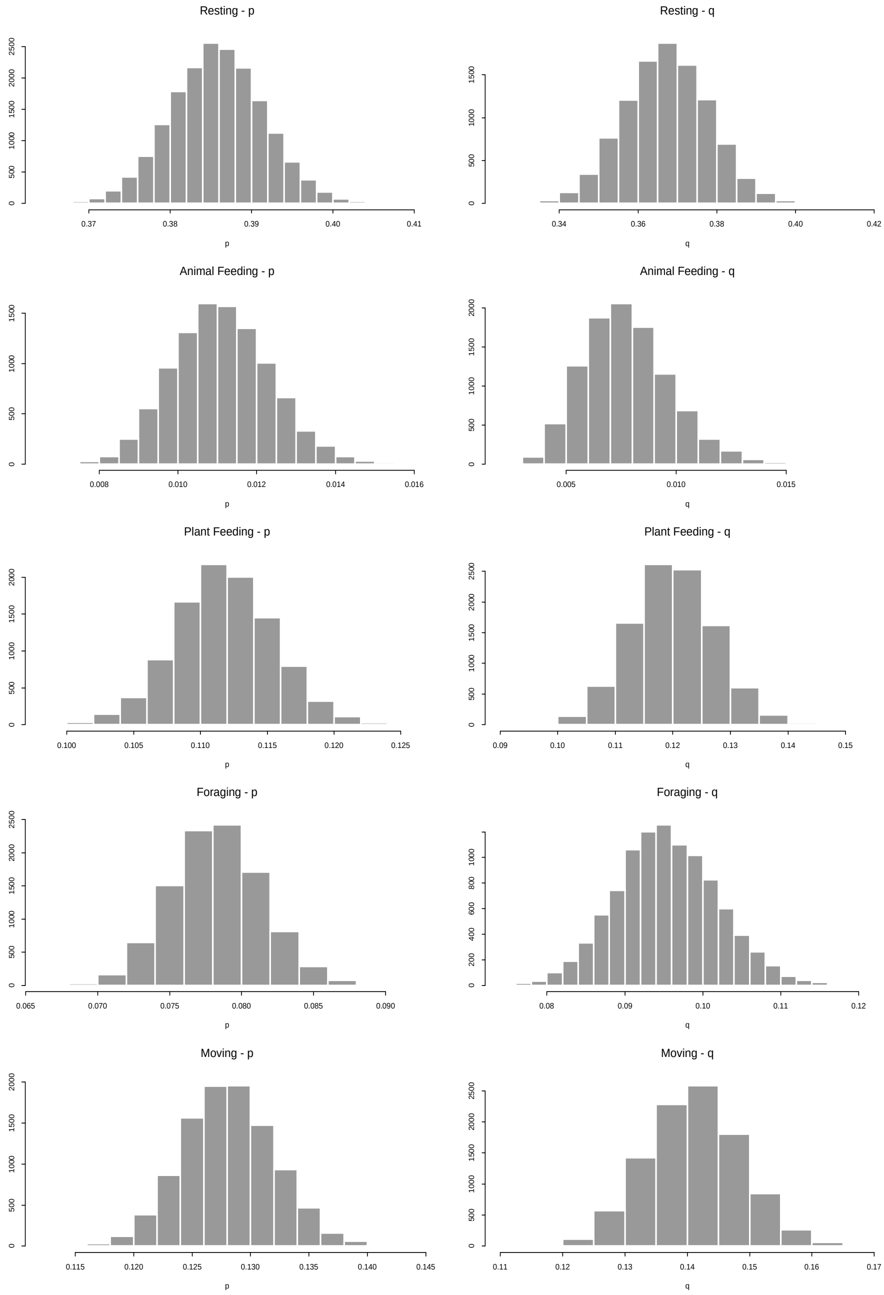

Figure 1
Histogram of the chains

5 Comparing the posterior distributions

The Kullback-Leibler (KL) divergence measure and Monte Carlo methods are used here in order to compare the posterior distributions of p and q. Suppose that the random variables W_1 and W_2 have pdf $f_{W_1}(z) = \pi_1(z|x,m)$ and pdf $f_{W_2}(z) = \pi_2(z|y,n)$, respectively, and, for some $j = 1, 2, \ldots J$,

$$\begin{aligned} f_{W_1}(z_j) &= \iint \pi_1(z_j|a,b,x,m)\pi(a,b|x,m)dadb \\ &= E_{a,b}(\pi_1(z_j|a,b,x,m)). \end{aligned} \tag{5.1}$$

Thus, for values $(a_1, b_1), (a_2, b_2), \ldots, (a_I, b_I)$ generated from $\pi(a,b|x,m)$ the expectation (5.1) can be estimated by (Mackay, 1996; Gelfand *et al.*, 1992)

$$f_{W_1}(z_j) \cong \frac{1}{I}\sum_{i=1}^{I} \pi_1(z_j|a_i, b_i, x, m), \qquad \text{for all } I \text{ sufficiently large.} \tag{5.2}$$

Similarly, for values $(c_1, d_1), (c_2, d_2), \ldots, (c_I, d_I)$ generated from $\pi(c, d|y, n)$, $f_{W_2}(z_j)$ can be estimated by

$$f_{W_2}(z_j) \cong \frac{1}{I}\sum_{i=1}^{I} \pi_2(z_j|c_i, d_i, y, n), \qquad \text{for all } I \text{ sufficiently large.} \tag{5.3}$$

The KL divergence between f_{W_1} and f_{W_2} is defined as (Csiszár, 1967)

$$\begin{aligned} K(f_{W_1},\ f_{W_2}) &= \int \log\left(\frac{f_{W_2}(z)}{f_{W_1}(z)}\right) f_{W_1}(z)dz \\ &= K(\pi_1(\cdot|x,m),\ \pi_2(\cdot|y,n)). \end{aligned}$$

Using (5.2) and (5.3) the discrepancy between π_1 and π_2 can be approximately measured by

$$K(\pi_1(\cdot|x,m),\ \pi_2(\cdot|y,n)) \cong \frac{1}{J}\sum_{J=1}^{J} \log\left(\frac{\frac{1}{I_2}\sum_{i=1}^{I_2} \pi_2(z_j|c_i, d_i, y, n)}{\frac{1}{I_1}\sum_{i=1}^{I_1} \pi_1(z_j|a_i, b_i, x, m)}\right), \tag{5.4}$$

for all J, I_1 and I_2 sufficient large.

The use of $K(\pi_1,\ \pi_2)$ as a measure of the discrepancy between π_1 and π_2 depends on the calibration of $K(\cdot,\cdot)$. Thus, given $K(\pi_1,\ \pi_2) = K$, we have to create a scale to decide if K indicates a weak or a strong divergence between the posteriors.

The calibration of $K(\cdot,\cdot)$ follows from McCulloch (1989) and Peng and Dey (1995). Let $\{g(\cdot|p) : p \in (0,1)\}$ be the family of probability functions corresponding to the Bernoulli distribution, that is, $g(x|p) = p^x(1-p)^{1-x}I_{\{0,1\}}(x)$, $p \in (0,1)$. Fixing, arbitrarily, the numbers p_0 and p_1, such that $p_0 \in (0,\ 0.5)$ and $p_1 \in (0.5,1)$, with p_0 and p_1 symmetrically about $p = 0.5$, as, for example, $p_0 = 0.02$ and $p_1 = 0.98$, it is reasonable to suppose that, for all $p < p_0$ or $p > p_1$ $g(\cdot|0.5)$ and $g(\cdot|p)$ are completely different, and for all $p \in [p_0,\ p_1]$ it is reasonable to suppose that $g(\cdot|0.5)$ and $g(\cdot|p)$ are similar. It implies that, for all $p \in (0,\ p_0) \cup (p_1,\ 1)$, the value of $K(g(\cdot|0.5),\ g(\cdot|p))$ could be considered large and, for all $p \in [p_0,\ p_1]$, the value of $K(g(\cdot|0.5),\ g(\cdot|p))$ could be considered small.

The function $K(g(\cdot|0.5),\ g(\cdot|p))$, which is equal to $-0.5\log[4p(1-p)]$, for $p \in (0,1)$, is symmetric about $p = 0.5$, monotone decreasing for $0 < p < 0.5$, monotone increasing for $0.5 < p < 1$, and zero for $p = 0.5$, (see Figure 2). For this reason the calibration is defined as follows. If $K(\pi_1,\ \pi_2) \geq K_0$, where $K_0 = K(g(\cdot|0.5),\ g(\cdot|p_0))$,then π_1 and π_2 are completely different, otherwise, π_1 and π_2 are similar.

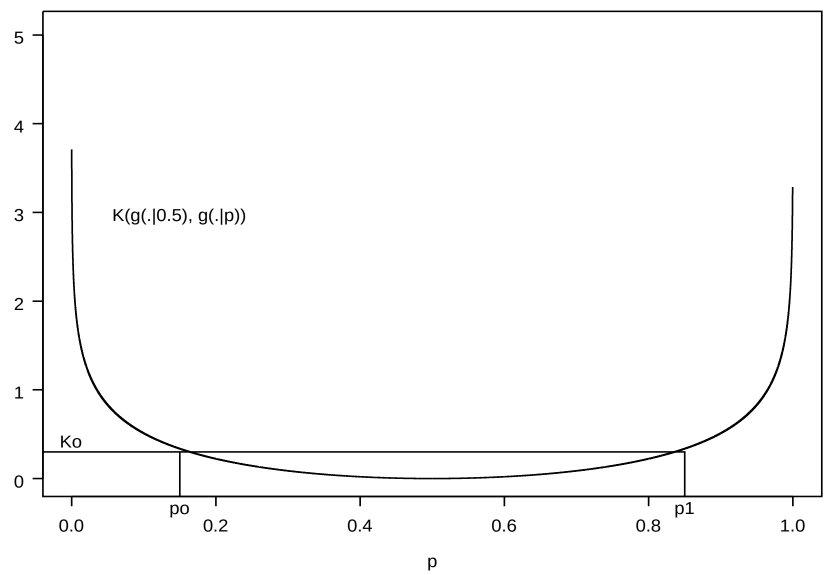

Figure 2
The calibration of $K(\cdot|\cdot)$

The values of the approximate KL divergence measure between the marginal posteriors (rainy and dry seasons) for each activity, measured by expression (5.4), are given in Table 3. The results suggest the plausibility of different behavior between the rainy and dry seasons.

Table 3
Divergence measure between the marginal posteriors

Activity	Approximate KL divergence measure
Resting	37.3586
Animal-Feeding	8.4387
Plant-Feeding	8.9072
Foraging	22.3466
Moving	9.8767

6 Discussion

The methodology proposed in this paper is an attempt to use an hierarchical Bayes approach for data on daily activity cycles. This approach provides a convenient way to compare activity behavior in two dependent samples. In the situation presented here, samples from the marginal posterior distributions of success probabilities at rainy and dry seasons, considering informative prior distributions, were generated by the Gibbs with Metropolis-Hastings technique. Characteristics of those posterior were calculated.

The posterior distributions for each season were compared using the Kullback-Leibler divergence measure. In all situations the group of animals seems to have different behavior with respect to the five activities studied here.

Table 4
Raw data set, by season, by month, by day, and by activity: resting (A1);animal-feeding(A2); plant-feeding(A3); foraging(A4); moving (A5) and not in sight(A6).

Season	Month	Day	A1	A2	A3	A4	A5	A6	Total
rainy	Nov	1	120	3	34	7	41	85	290
rainy	Nov	2	106	5	57	14	41	73	296
rainy	Nov	3	132	4	35	28	48	77	324
rainy	Dec	1	152	13	19	19	60	45	308
rainy	Dec	2	112	1	14	16	39	108	290
rainy	Dec	3	158	9	11	14	41	62	295
rainy	Jan	1	48	0	24	5	21	43	141
rainy	Jan	2	91	1	33	11	29	115	280
rainy	Jan	3	65	3	37	20	39	129	293
rainy	Feb	1	84	3	36	15	28	92	258
rainy	Feb	2	64	2	13	30	22	119	250
rainy	Feb	3	106	1	26	39	38	60	270
rainy	Mar	1	113	1	11	15	35	85	260
rainy	Mar	2	108	0	32	16	34	60	250
rainy	Mar	3	97	11	3	23	35	102	271
rainy	Apr	1	101	0	23	23	25	89	261
rainy	Apr	2	97	6	23	29	23	112	290
rainy	Apr	3	109	0	54	33	35	23	254
rainy	May	1	102	7	34	26	35	46	250
rainy	May	2	80	0	45	13	41	37	216
rainy	May	3	126	1	20	17	35	61	260
dry	Jun	1	73	1	30	25	28	74	231
dry	Jun	2	112	3	9	17	33	46	220
dry	Jun	3	58	1	52	19	28	42	200
dry	Jul	1	57	6	22	15	23	87	210
dry	Jul	2	82	1	39	22	32	54	230
dry	Jul	3	97	1	20	30	42	36	226
dry	Aug	1	98	1	25	20	30	95	269
dry	Aug	2	98	0	32	28	45	47	250
dry	Aug	3	90	2	20	22	32	75	241
rainy	Sep	1	140	0	30	29	23	58	280
rainy	Sep	2	71	6	35	29	38	91	270
rainy	Sep	3	107	2	54	36	44	67	310
rainy	Oct	1	102	1	33	22	31	101	290
rainy	Oct	2	74	1	41	33	26	85	260
rainy	Oct	3	45	0	13	3	14	205	280

Acknowledgements

The authors wish to thank a referee for helpful comments and suggestions. The research of José Galvão Leite is partially supported by CNPq/Brazil, grant 301339/91-0.

(Received July, 1999. Revised March, 2001.)

References

Altmann, J. (1974). Observational study of behavior: sampling methods. Behaviour, **49**, 227-267.

Berger, J. O. (1985). Statistical Decision Theory and Bayesian Analysis. New York: Springer-Verlag.

Best, N. G., Cowles, M. K. and Vines, S. K. (1995). *CODA manual version 0.30*. MRC Biostatistics Unit, Cambridge, UK.

Casella G. and George, E. (1992). Explaining the Gibbs Sampler. *The American Statistician*, **46**, 167-174.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, **49**, 4, 327-355.

Costa, A.C. M. (1997). Ecologia de um grupo de mico-Leão preto (*L. chrysopygus)* em um fragmento de mata. Unpublished M.Sc. dissertation, UNESP-Rio Claro, S.P.-Brazil.

Gelfand, A. E., Smith, A. F. M. and Lee, T. (1992). Bayesian analysis of constrained parameter and truncated data problem using Gibbs sampling. *Journal of the American Statistical Association*, **87,** 523-532.

George, E. I. (1992). Capture-recapture estimation via Gibbs sampling. *Biometrika*, **79**, 677-83.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. (1995). Markov Chain Monte Carlo in Practice. New York: Chapman-Hall.

Mackay, D. J. C. (1996). Introduction to Monte Carlo methods. Proceedings of Erice Summer School, ed. M.Jordan.

McCulloch, R. E. (1989). Local prior influence. *Journal of the American Statistical Association*, **84**, 473-478.

Peng, F. and Dey, D. K. (1995). Bayesian analysis of outlier problems using divergence measures. *The Canadian Journal of Statistics,* **23**, 199-213.

Sussman, R. W., O'fallon, W. M., Sussman, L. K. and Buettner-Janusch, J. (1979). Statistical methods for analyzing data on daily activity cycles of primates. *American Journal of Physical Anthropology*, **51**, 1-14.