

Multiple Change Point Analysis for the Regular Exponential Family using the Product Partition Model

R. H. Loschi¹, F. R. B. Cruz¹ and R. B. Arellano-Valle²

¹*Universidade Federal de Minas Gerais*
and ²*Pontificia Universidad Catolica de Chile*

Abstract: As an extension to previous research efforts, the PPM is applied to the identification of multiple change points in the parameter that indexes the regular exponential family. We define the PPM for Yao's prior cohesions and contiguous blocks. Because the exponential family provides a rich set of models, we also present the PPM for some particular members of this family in both continuous and discrete cases and the PPM is applied to identify multiple change points in real data. Firstly, multiple changes are identified in the rates of crimes in one of the biggest cities in Brazil. In order to illustrate the continuous case, multiple changes are identified in the volatility (variance) and in the expected return (mean) of some Latin America emerging markets return series.

Key words: Conjugacy, Gibbs sampling, relevance, Yao's cohesions.

1. Introduction

Most of the methodologies considered in change points analysis assume that the number of change points is known and fixed (see Chen and Lee, 1995; Geweke and Terui, 1993; Hawkins, 2001, and many others). Other authors have studied the one-change-point problem using a Bayesian approach (see Menzefricke, 1981; Hsu, 1984; Smith, 1975, for example). The product partition model (PPM) developed by Hartigan(1990) introduces more flexibility into the analysis of these problems since it considers the number of change points as a random variable. As shown by Barry and Hartigan (1992), by applying the PPM one can easily obtain product estimates for the parameters of interest at each point of the time, the posterior distribution of the random partition generated by the change points, and also the posterior distribution of the number of change points. Applications and extensions of the PPM can be found in Barry and Hartigan (1993), Crowley (1997), Quintana and Iglesias (2003), Loschi and Cruz (2005) and others (see more on Bayesian approach of multiple change points in Elliott and Shope, 2003).

Despite of all flexibility introduced by the PPM in the analysis of change point problems, in many situations, we have noticed that the posterior distribution for the random partition, as originally defined by Barry and Hartigan (1993), puts similar or equal mass for a large number of partitions. For such cases, it is a hard task to identify the change points. One way to measure the evidence of a change is to compute the posterior probability of each instant being a change point as established by Loschi and Cruz (2005). Another way to solve this problem is to consider the theoretical decision approach for choosing the best partition introduced by Quintana and Iglesias (2003).

In this paper we apply the PPM and its extensions proposed by Loschi and Cruz (2005) to identify multiple change points in the parameter that indexes the regular exponential family. We suppose that only contiguous blocks are possible and consider the prior cohesions proposed by Yao (1984). The change point model considered by Yao (1984) is a discrete time version of the model discussed by Barnard (1959). We also present the PPM for some particular members of the exponential family in both discrete and continuous cases. Consequently, we extend to a more general family the methodology earlier developed by Loschi *et al.* (2003) to identify change points in normal means and variances.

Two real data sets are analyzed. Firstly, to illustrate the discrete cases, we identify multiple change points in the rates of crimes in Belo Horizonte, one of the biggest cities in Brazil. Then, multiple change points are identified in the volatility (variance) and in the expected return (mean) of some Latin America emerging market return series, illustrating the use of the PPM in the continuous case.

This paper is organized as follows. In Section 2, the parametric PPM is defined. Some computational procedures proposed in the literature to manage the PPM are also described. In Section 3, we apply the PPM to the exponential family and to some particular distributions. In Section 4, the results are applied to the analysis of two real data sets. Finally, in Section 5 some conclusions close the paper.

2. The Product Partition Model

In this section we briefly describe the PPM as well as its extensions to compute the posterior probability of each instant being a change point. The computational method involved is also presented.

2.1 PPM using Yao's cohesions

Let X_1, \dots, X_n be a data sequence and consider the index set $I = \{1, \dots, n\}$. Consider a random partition $\rho = \{i_0, i_1, \dots, i_b\}$ of set I , $0 = i_0 < i_1 < \dots <$

$i_b = n$. Consider that each partition divides the data sequence X_1, \dots, X_n into $B = b$ contiguous subsequences (blocks), which are denoted here by $\mathbf{X}_{[i_{r-1}i_r]} = (X_{i_{r-1}+1}, \dots, X_{i_r})^t$, for $r = 1, \dots, b$. Let $c_{[ij]}$ be the prior cohesion associated with the block $[ij] = \{i + 1, \dots, j\}$, for $i, j \in I \cup \{0\}$, and $j > i$, that represents the degree of similarity among the observations in $\mathbf{X}_{[ij]}$ and that may be interpreted here as transition probabilities in the Markov chain defined by the change points.

Let $0 \leq p \leq 1$ be the probability that a change occurs at any instant in the sequence. The prior cohesion for block $[ij]$ proposed by Yao (1984) is given by:

$$c_{[ij]} = \begin{cases} p(1 - p)^{j-i-1}, & \text{if } j < n, \\ (1 - p)^{j-i-1}, & \text{if } j = n, \end{cases}$$

for all $i, j \in I, i < j$. These prior cohesions imply that the sequence of change points establishes a discrete renewal process, with occurrence times identically distributed with geometric distribution.

Let $\theta_1, \dots, \theta_n$ be a sequence of unknown parameters conditional on which the sequence of random variables X_1, \dots, X_n has conditional marginal densities $f_1(X_1|\theta_1), \dots, f_n(X_n|\theta_n)$, respectively. The prior distribution of $\theta_1, \dots, \theta_n$ is constructed as follows. Given a partition $\rho = \{i_0, \dots, i_b\}$, for $b \in I$, one has that $\theta_i = \theta^{[i_{r-1}i_r]}$, for every $i_{r-1} < i \leq i_r$ and $r = 1, \dots, b$, and that $\theta^{[i_0i_1]}, \dots, \theta^{[i_{b-1}i_b]}$ are independent and also independent from p , with $\theta^{[ij]}$ having prior (block) density $\pi_{[ij]}(\theta), \theta \in \Theta_{[ij]}$, where $\Theta_{[ij]}$ is the parameter space corresponding to the common parameter, say, $\theta^{[ij]} = \theta_{i+1} = \dots = \theta_j$, that indexes the conditional density of $\mathbf{X}_{[ij]}$. Hence, the PPM for Yao's prior cohesions is defined as follows.

The random quantity $(X_1, \dots, X_n; \rho)$ follows a PPM, denoted by $(X_1, \dots, X_n; \rho) \sim PPM$, if:

- i) given p , the prior distribution of ρ is the following product distribution:

$$P(\rho = \{i_0, \dots, i_b\} | p) = p^{b-1}(1 - p)^{n-b}, \quad b \in I,$$

for every partition $\{i_0, \dots, i_b\}$.

- ii) Given $\rho = \{i_0, \dots, i_b\}$ and p , the sequence X_1, \dots, X_n is independent from p and has the joint density given by:

$$f(X_1, \dots, X_n | \rho = \{i_0, \dots, i_b\}, p) = \prod_{j=1}^b f_{[i_{j-1}i_j]}(\mathbf{X}_{[i_{j-1}i_j]}),$$

in which $f_{[ij]}(\mathbf{X}_{[ij]}) = \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta$ is the joint density of the random vector, called data factor.

It is also shown that the posterior distribution of θ_s is given by:

$$\pi(\theta_s | X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=s}^n r_{[ij]}^* \pi_{[ij]}(\theta_s | \mathbf{X}_{[ij]}), \quad s = 1, \dots, n, \quad (2.1)$$

and that the posterior expectation (or product estimate) of θ_s is given by:

$$E(\theta_s | X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=s}^n r_{[ij]}^* E(\theta_s | \mathbf{X}_{[ij]}), \quad s = 1, \dots, n, \quad (2.2)$$

in which $r_{[ij]}^* = P([ij] \in \rho | X_1, \dots, X_n)$ denotes the posterior relevance for the block $[ij]$. More details can be found in Barry and Hartigan (1992).

Let us assume that p has prior distribution $\pi(p)$. As a consequence of these assumptions, it follows that the posterior distributions of ρ and B are given, respectively, by:

$$\begin{aligned} &P(\rho = \{i_0, \dots, i_b\} | X_1, \dots, X_n) \\ &\propto \prod_{j=1}^b f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]}) \int_0^1 p^{b-1} (1-p)^{n-b} \pi(p) dp, \end{aligned}$$

and

$$\begin{aligned} &P(B = b | X_1, \dots, X_n) \\ &\propto \sum_{\mathcal{C}_b} \prod_{j=1}^b f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]}) \int_0^1 p^{b-1} (1-p)^{n-b} \pi(p) dp, \end{aligned}$$

where \mathcal{C}_b denotes the set of all partitions of I into b contiguous blocks.

Because the product estimates are strongly influenced by the prior distribution of p , it is important to obtain its posterior distribution. Assuming that $p \sim \pi(p)$, such posterior distribution is given by:

$$\pi(p | X_1, \dots, X_n) \propto \sum \prod_{j=1}^b f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]}) p^{b-1} (1-p)^{n-b} \pi(p),$$

in which the summation is over all partitions of I .

An evidence of a change point is the posterior probability of each instant being a change point. Let \mathcal{C}_l be the set that contains all partitions that include the instant l as a change point, that is, each partition in \mathcal{C}_l assume the form $\{i_0, \dots, i_{k-1}, i_k = l, i_{k+1}, \dots, i_b\}$ for any $k \in I$. The event A_l denotes that the instant l is a change point, for $l = 1, \dots, n-1$. Thus, the posterior probability of A_l is given by:

$$P(A_l | X_1, \dots, X_n) = \frac{\sum_{\mathcal{C}_l} \prod_{j=1}^b f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]}) \int_0^1 p^{b-1} (1-p)^{n-b} \pi(p) dp}{\sum \prod_{j=1}^b f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]}) \int_0^1 p^{b-1} (1-p)^{n-b} \pi(p) dp}$$

where the summation in the denominator is over all partition of I .

2.2 Computational methods

To compute the posterior distributions of ρ , B and p , the product estimates of θ_k , and the posterior probability of A_l , the following algorithm was proposed in the literature. Assume that, given ρ , $\theta_l = \theta_{i+1} = \dots = \theta_j$, for $l = 1, \dots, n$ and $i, j \in I$, $i < j$. Let $\mathbf{X}_{[0n]} = (X_1, \dots, X_n)$ and $\theta = (\theta_1, \dots, \theta_n)$ and denote by θ_{-l} the vector $(\theta_1, \dots, \theta_{l-1}, \theta_{l+1}, \dots, \theta_n)$. The full conditional distributions of p , ρ , and θ_l , for $l = 1, \dots, n$ are given, respectively, by:

$$\begin{aligned}\pi(p|\rho, \theta, \mathbf{X}_{[0n]}) &\propto p^{b-1}(1-p)^{n-b}\pi(p); \\ \pi(\rho|p, \theta, \mathbf{X}_{[0n]}) &\propto \left(\prod_{j=1}^b f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]})\right) p^{b-1}(1-p)^{n-b}; \\ \pi(\theta_l|\rho, p, \theta_{-l}, \mathbf{X}_{[0n]}) &\propto f_{[ij]}(\theta_l|X_{[ij]}).\end{aligned}$$

Notice that, because all partitions must be considered, it may be very difficult to sample directly from the full conditional distribution of ρ in case of long sequences are assumed. Let us define the auxiliary random quantity U_l , such that $U_l = 1$, if $\theta_l = \theta_{l+1}$, and $U_l = 0$, otherwise, for $l = 1, \dots, n-1$. Notice that the random partition ρ is immediately identified by considering vectors $\mathbf{U} = (U_1, \dots, U_{n-1})$ of these random quantities. Each partition $(U_1^s, \dots, U_{n-1}^s)$, $s \geq 1$, is generated by using Gibbs sampling by considering the following ratio:

$$R_r = \frac{f_{[xy]}(X_{[xy]}) \int_0^1 p^{b-2}(1-p)^{n-b+1} d\pi(p)}{f_{[xr]}(X_{[xr]}) f_{[ry]}(X_{[ry]}) \int_0^1 p^{b-1}(1-p)^{n-b} d\pi(p)}, \quad (2.3)$$

for $r = 1, \dots, n-1$, and where x is the last change point before r and y is the next change point following r (see, Loschi and Cruz (2005) for all the details).

3. The Product Partition Model for the Regular Exponential Family

In this section we extend the PPM to identify multiple change points in the parameter that indexes the exponential family. In order to permit a tractable implementation of the PPM, some results related to the conjugacy for the regular exponential family established in the literature are considered (see Bernardo and Smith, 1994, for example).

Let $\theta_l \in \Theta_l \subseteq \mathcal{R}^d$, $d \geq 1$, $l = 1, \dots, n$. Assume that, given $\theta_1, \dots, \theta_n$, X_1, \dots, X_n , $X_l \in \mathcal{X} \subseteq \mathcal{R}$, $l = 1, \dots, n$, are independent and $X_l|\theta_l$ has density

function in the regular d -parameter exponential family, that is,

$$p(X_l|\theta_l) = h(X_l)L(\theta_l) \exp \left\{ \sum_{k=1}^d c_k \phi_k(\theta_l) f_k(X_l) \right\}$$

where h (positive), $f = (f_1, \dots, f_d)$ and $\phi = (\phi_1, \dots, \phi_d)$ are functions and c_l , $l = 1, \dots, d$ are constants such that

$$[L(\theta_l)]^{-1} = \int_{\mathcal{X}} h(X_l) \exp \left\{ \sum_{k=1}^d c_k \phi_k(\theta_l) f_k(X_l) \right\} d(X_l) < \infty.$$

It follows that the common parameter $\theta_{[ij]}$ for the block $X_{[ij]}$ has the following conjugate prior distribution:

$$p(\theta^{[ij]}|\tau^{[ij]}) = [K(\tau^{[ij]})]^{-1} [L(\theta^{[ij]})]^{\tau_0^{[ij]}} \exp \left\{ \sum_{k=1}^d c_k \phi_k(\theta^{[ij]}) \tau_k^{[ij]} \right\}, \quad (3.1)$$

where $\tau^{[ij]} = (\tau_0^{[ij]}, \dots, \tau_d^{[ij]})$ such that

$$K(\tau^{[ij]}) = \int_{\Theta} [L(\theta^{[ij]})]^{\tau_0^{[ij]}} \exp \left\{ \sum_{k=1}^d c_k \phi_k(\theta^{[ij]}) \tau_k^{[ij]} \right\} d(\theta^{[ij]}) < \infty.$$

As a consequence of such assumptions it follows that the prior predictive distribution of $X_{[ij]}$ is

$$p(X_{[ij]}|\tau^{[ij]}) = \Pi_{l=i+1}^j h(X_l) \frac{K(\tau^{[ij]} + t_{j-i}(X_{[ij]}))}{K(\tau^{[ij]})} \quad (3.2)$$

where $t_{(j-i)}(X_{[ij]}) = (j-i, \sum_{l=i+1}^j f_1(X_l), \dots, \sum_{l=i+1}^j f_d(X_l))$ and the block posterior distribution of $\theta^{[ij]}$, given $X_{[ij]}$, is

$$\begin{aligned} p(\theta^{[ij]}|X_{[ij]}, \tau^{[ij]}) &= [K(\tau^{[ij]} + t_{(j-i)}(X_{[ij]}))]^{-1} [L(\theta^{[ij]})]^{\tau_0^{[ij]} + j - i} \\ &\times \exp \left\{ \sum_{k=1}^d c_k \phi_k(\theta^{[ij]}) \left(\tau_k^{[ij]} + \sum_{l=i+1}^j f_k(X_l) \right) \right\}. \end{aligned} \quad (3.3)$$

Consequently, from expressions (3.3) and (2.1) the posterior distribution of θ_s , $s = 1, \dots, n$ assumes the following expression:

$$\begin{aligned} p(\theta_s|X_1, \dots, X_n, \tau^{[ij]}) &= \sum_{i=1}^{s-1} \sum_{j=s}^n r_{[ij]}^* [K(\tau^{[ij]} + t_{(j-i)}(X_{[ij]}))]^{-1} [L(\theta_s)]^{\tau_0^{[ij]} + j - i} \\ &\exp \left\{ \sum_{k=1}^d c_k \phi_k(\theta_s) \left(\tau_k^{[ij]} + \sum_{l=i+1}^j f_k(X_l) \right) \right\}. \end{aligned}$$

However, we should notice that the general formula for the product estimates can be specified only for the natural parameter. In such a case, we should consider some results presented in Diaconis and Ylvisaker (1979). The product estimates for special members of the exponential family will be shown shortly in Sections 3.1 to 3.6.

From (2.3) and (3.2), the estimates of the posterior distributions of ρ , B , A_l and p as well as the estimates of the posterior relevancies can be generated considering the following expression:

$$R_r = \frac{K(\tau^{[xr]})K(\tau^{[ry]})K(\tau^{[xy]} + t_{y-x}(X_{[xy]})) \int_0^1 p^{b-2}(1-p)^{n-b+1}d\pi(p)}{K(\tau^{[xy]})K(\tau^{[xr]} + t_{r-x}(X_{[xr]}))K(\tau^{[ry]} + t_{y-r}(X_{[ry]})) \int_0^1 p^{b-1}(1-p)^{n-b}d\pi(p)}.$$

Notice that, for the exponential family, the sampling scheme is simplified since it is not necessary to know the complete expression of the predictive distribution.

Since the exponential family provides a rich set of models the PPM will be applied to some particular members of this family, as follows.

3.1 Bernoulli data sequence

Let $\theta_l \in [0, 1] \times \mathcal{R}$, $l = 1, \dots, n$. Assume that, conditional on $\theta_1, \dots, \theta_n$, the sequence X_1, \dots, X_n , $X_l \in \{0, 1\}$, $l = 1, \dots, n$, are independent and $X_l|\theta_l$ has Bernoulli distribution with parameter θ_l denoted by $X_l|\theta_l \sim \mathcal{BE}(\theta_l)$. Consequently, for the block $X_{[ij]}$ we have that

$$f(X_{[ij]}|\theta^{[ij]}) = \left(\frac{\theta^{[ij]}}{1-\theta^{[ij]}}\right)^{\sum_{l=i+1}^j X_l} (1-\theta^{[ij]})^{j-i}. \tag{3.4}$$

Let $\tau^{[ij]} = (\tau_0^{[ij]}, \tau_1^{[ij]})$. From (3.1) it follows that the conjugate prior distribution for the common parameter $\theta^{[ij]}$ is the beta distribution with parameters $\tau_1^{[ij]} + 1$ and $\tau_0^{[ij]} - \tau_1^{[ij]} + 1$, denoted by $\theta^{[ij]} \sim \mathcal{B}(\tau_1^{[ij]} + 1, \tau_0^{[ij]} - \tau_1^{[ij]} + 1)$, whose density function is:

$$f(\theta^{[ij]}|\tau^{[ij]}) = [K(\tau^{[ij]})]^{-1}(\theta^{[ij]})^{\tau_1^{[ij]}}(1-\theta^{[ij]})^{\tau_0^{[ij]}-\tau_1^{[ij]}}$$

where $\tau_1^{[ij]} > -1$, $\tau_0^{[ij]} > \tau_1^{[ij]} - 1$ and

$$K(\tau^{[ij]}) = \frac{\Gamma(\tau_1^{[ij]} + 1)\Gamma(\tau_0^{[ij]} - \tau_1^{[ij]} + 1)}{\Gamma(\tau_0^{[ij]} + 2)}. \tag{3.5}$$

Since from (3.4) we have that $t_{(j-i)} = (j-i, \sum_{l=i+1}^j X_l)$, from (3.3) it follows that $\theta^{[ij]}|X_{[ij]} \sim \mathcal{B}(\tau_1^{[ij]*}, \tau_0^{[ij]*})$ in which $\tau_1^{[ij]*} = \tau_1^{[ij]} + \sum_{l=i+1}^j X_l + 1$ and $\tau_0^{[ij]*} =$

$\tau_0^{[ij]} + j - i - \tau_1^{[ij]} - \sum_{k=i+1}^j X_k + 1$. Thus, from (2.2) the product estimates of θ_s , $s = 1, \dots, n$, are given by:

$$E(\theta_s | X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=s}^n r_{[ij]}^* \frac{\tau_1^{[ij]} + \sum_{l=i+1}^j X_l + 1}{\tau_0^{[ij]} + j - i + 2}.$$

The posterior relevancies and the posterior distributions of ρ , p , B and A_l can be generated by substituting (3.5) into (2.3).

3.2 Poisson data sequence

Let $\theta_l > 0$, $l = 1, \dots, n$. Assume that, conditional on $\theta_1, \dots, \theta_n$, the sequence X_1, \dots, X_n , $X_l \in \{0, 1, \dots\}$, $l = 1, \dots, n$, are independent and $X_l | \theta_l$ has Poisson distribution with parameter θ_l denoted by $X_l | \theta_l \sim \mathcal{P}(\theta_l)$. Consequently, for the block $X_{[ij]}$ we have that

$$f(X_{[ij]} | \theta^{[ij]}) = \prod_{l=i+1}^j [X_l!]^{-1} \exp\{-(j-i)\theta^{[ij]}\} (\theta^{[ij]})^{\sum_{l=i+1}^j X_l}. \tag{3.6}$$

Let $\tau^{[ij]} = (\tau_0^{[ij]}, \tau_1^{[ij]})$. From (3.1) it follows that the conjugate prior distribution for the common parameter $\theta^{[ij]}$, related to the block $[ij]$ is the gamma distribution with parameters $\tau_1^{[ij]} + 1$ and $\tau_0^{[ij]}$, denoted by $\theta_{[ij]} \sim \mathcal{G}(\tau_1^{[ij]} + 1, \tau_0^{[ij]})$, with density function given by:

$$f(\theta^{[ij]} | \tau_0^{[ij]}, \tau_1^{[ij]}) = [K(\tau^{[ij]})]^{-1} (\theta^{[ij]})^{\tau_1^{[ij]}} \exp(-\tau_0^{[ij]} \theta^{[ij]}).$$

in which $\tau_0^{[ij]} > 0$, $\tau_1^{[ij]} > -1$ and

$$K(\tau^{[ij]}) = \Gamma(\tau_1^{[ij]} + 1) (\tau_0^{[ij]})^{-(\tau_1^{[ij]} + 1)}. \tag{3.7}$$

Since from (3.6) we have that $t_{(j-i)} = (j-i, \sum_{l=i+1}^j X_l)$, from (3.3) it follows that $\theta^{[ij]} | X_{[ij]} \sim \mathcal{G}(\tau_1^{[ij]*}, \tau_0^{[ij]*})$ in which $\tau_1^{[ij]*} = \tau_1^{[ij]} + \sum_{l=i+1}^j X_l + 1$ and $\tau_0^{[ij]*} = \tau_0^{[ij]} + j - i$. Thus, from (2.2) the product estimates of θ_s , $s = 1, \dots, n$, are given by:

$$E(\theta_s | X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=s}^n r_{[ij]}^* \frac{\tau_1^{[ij]} + \sum_{l=i+1}^j X_l + 1}{\tau_0^{[ij]} + j - i}.$$

The posterior relevancies and the posterior distributions of ρ , p , B and A_k can be generated by substituting (3.7) into (2.3).

3.3 Exponential data sequence

Let $\theta_l > 0, l = 1, \dots, n$. Assume that, conditional on $\theta_1, \dots, \theta_n$, the sequence $X_1, \dots, X_n, X_l \in \mathcal{R}, l = 1, \dots, n$, are independent and $X_l|\theta_l$ has exponential distribution with parameter θ_l denoted by $X_l|\theta_l \sim \mathcal{E}(\theta_l)$. Consequently, for the block $X_{[ij]}$ we have that

$$f(X_{[ij]}|\theta^{[ij]}) = \prod_{l=i+1}^j \mathbf{1}\{X_l > 0\} \exp\left\{-\theta^{[ij]} \sum_{l=i+1}^j X_l\right\} (\theta^{[ij]})^{j-i},$$

in which $\mathbf{1}\{A\}$ denotes the indicator function of event A .

Let $\tau^{[ij]} = (\tau_0^{[ij]}, \tau_1^{[ij]})$. From (3.1) it follows that the conjugate prior distribution for the common parameter $\theta^{[ij]}$, related to the block $[ij]$ is the gamma distribution with parameters $\tau_0^{[ij]} + 1$ and $\tau_1^{[ij]}$, denoted by $\theta_{[ij]} \sim \mathcal{G}(\tau_0^{[ij]} + 1, \tau_1^{[ij]})$, with density function given by:

$$f(\theta^{[ij]}|\tau_0^{[ij]}, \tau_1^{[ij]}) = [K(\tau^{[ij]})]^{-1} (\theta^{[ij]})^{\tau_0^{[ij]}} \exp(-\tau_1^{[ij]}\theta^{[ij]}).$$

in which $\tau_1^{[ij]} > 0, \tau_0^{[ij]} > -1$ and

$$K(\tau^{[ij]}) = \Gamma(\tau_0^{[ij]} + 1) (\tau_1^{[ij]})^{-(\tau_0^{[ij]} + 1)}. \tag{3.8}$$

Because we have that $t_{(j-i)} = (j - i, \sum_{l=i+1}^j X_l)$, from (3.3) it follows that $\theta^{[ij]}|X_{[ij]} \sim \mathcal{G}(\tau_0^{[ij]*}, \tau_1^{[ij]*})$ in which $\tau_1^{[ij]*} = \tau_1^{[ij]} + \sum_{l=i+1}^j X_l$ and $\tau_0^{[ij]*} = \tau_0^{[ij]} + j - i + 1$. Thus, from (2.2) the product estimates of $\theta_s, s = 1, \dots, n$, are given by:

$$E(\theta_s|X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=s}^n r_{[ij]}^* \frac{\tau_0^{[ij]} + j - i + 1}{\tau_1^{[ij]} + \sum_{l=i+1}^j X_l}.$$

The posterior relevancies and the posterior distributions of ρ, p, B and A_k can be generated by substituting (3.8) into (2.3).

3.4 Normal data sequence with unknown means

Let $\theta_l \in \mathcal{R}$, $l = 1, \dots, n$. Assume that, conditional on $\theta_1, \dots, \theta_n$, the sequence X_1, \dots, X_n , $X_l \in \mathcal{R}$, $l = 1, \dots, n$, are independent and $X_l|\theta_l$ has normal distribution with unknown mean θ_l and the same known variance σ^2 , denoted by $X_l|\theta_l \sim \mathcal{N}(\theta_l, \sigma^2)$. Consequently, for the block $X_{[ij]}$ we have that

$$f(X_{[ij]}|\theta^{[ij]}) = \left(\frac{1}{2\pi\sigma^2}\right)^{(j-i)/2} \exp\left\{-\frac{\sum_{l=i+1}^j X_l^2}{2\sigma^2}\right\} \exp\left\{-(j-i)\frac{(\theta^{[ij]})^2}{2\sigma^2}\right\} \exp\left\{\frac{\theta^{[ij]} \sum_{l=i+1}^j X_l}{\sigma^2}\right\}. \quad (3.9)$$

Let $\tau^{[ij]} = (\tau_0^{[ij]}, \tau_1^{[ij]})$. From (3.1) it follows that the conjugate prior distribution for the common parameter $\theta^{[ij]}$ is the normal distribution with mean $\frac{\tau_1^{[ij]}}{\tau_0^{[ij]}}$ and variance $\frac{\sigma^2}{\tau_0^{[ij]}}$, denoted by $\theta^{[ij]} \sim \mathcal{N}\left(\frac{\tau_1^{[ij]}}{\tau_0^{[ij]}}, \frac{\sigma^2}{\tau_0^{[ij]}}\right)$, whose density function is:

$$f(\theta^{[ij]}|\tau^{[ij]}) = [K(\tau^{[ij]})]^{-1} \exp\left\{-\frac{\tau_0^{[ij]}(\theta^{[ij]})^2}{2\sigma^2}\right\} \exp\left\{\frac{\theta^{[ij]} \tau_1^{[ij]}}{\sigma^2}\right\}, \quad (3.10)$$

where $\tau_0^{[ij]} > 0$, $\tau_1^{[ij]} \in \mathcal{R}$ and

$$k(\tau^{[ij]}) = \left(\frac{2\pi\sigma^2}{\tau_0^{[ij]}}\right)^{1/2} \exp\left\{\frac{(\tau_1^{[ij]})^2}{2\sigma^2\tau_0^{[ij]}}\right\}. \quad (3.11)$$

Since from (3.9) we have that $t_{(j-i)} = (j-i, \sum_{l=i+1}^j X_l)$, from (3.3) it follows that $\theta^{[ij]}|X_{[ij]} \sim \mathcal{N}(\tau_1^{[ij]*}, \tau_0^{[ij]*})$ in which $\tau_1^{[ij]*} = \frac{\tau_1^{[ij]} + \sum_{l=i+1}^j X_l}{\tau_0^{[ij]} + j - i}$ and $\tau_0^{[ij]*} = \frac{\sigma^2}{\tau_0^{[ij]} + j - i}$. Thus, from (2.2) the product estimates of θ_s , $s = 1, \dots, n$, are given by:

$$E(\theta_s|X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=s}^n r_{[ij]}^* \frac{\tau_1^{[ij]} + \sum_{k=i+1}^j X_k}{\tau_0^{[ij]} + j - i}.$$

The posterior relevancies and the posterior distributions of ρ , p , B and A_k can be generated by substituting (3.11) into (2.3). For different approaches for this PPM model see Barry and Hartigan (1993) and Crowley (1997).

3.5 Normal data sequence with unknown variances

Let $\theta_l > 0, l = 1, \dots, n$. Assume that, conditional on $\theta_1, \dots, \theta_n$, the sequence $X_1, \dots, X_n, X_l \in \mathcal{R}, l = 1, \dots, n$, are independent and $X_l|\theta_l$ has normal distribution with unknown variance θ_l and mean equal to zero, denoted by $X_l|\theta_l \sim \mathcal{N}(0, \theta_l)$. Consequently, for the block $X_{[ij]}$ we have that

$$f(X_{[ij]}|\theta^{[ij]}) = \left(\frac{1}{2\pi}\right)^{(j-i)/2} \left(\frac{1}{\theta^{[ij]}}\right)^{(j-i)/2} \exp\left\{-\frac{\sum_{l=i+1}^j X_l^2}{2\theta^{[ij]}}\right\}. \tag{3.12}$$

Let $\tau^{[ij]} = (\tau_0^{[ij]}, \tau_1^{[ij]})$. From (3.1) it follows that the conjugate prior distribution for the common parameter $\theta^{[ij]}$ is the inverted-gamma distribution with parameters $\frac{\tau_1^{[ij]}}{2}$ and $\frac{\tau_0^{[ij]}-2}{2}$, denoted by $\theta^{[ij]} \sim \mathcal{IG}\left(\frac{\tau_1^{[ij]}}{2}, \frac{\tau_0^{[ij]}-2}{2}\right)$, whose density function is:

$$f(\theta^{[ij]}|\tau^{[ij]}) = [K(\tau^{[ij]})]^{-1} \left(\frac{1}{\theta^{[ij]}}\right)^{\frac{\tau_0^{[ij]}}{2}} \exp\left\{-\frac{\tau_1^{[ij]}}{2\theta^{[ij]}}\right\},$$

where $\tau_1^{[ij]} > 0, \tau_0^{[ij]} > 2$ and

$$K(\tau^{[ij]}) = \Gamma\left(\frac{\tau_0^{[ij]}-2}{2}\right) \left(\frac{\tau_1^{[ij]}}{2}\right)^{-\frac{\tau_0^{[ij]}-2}{2}}. \tag{3.13}$$

Since from (3.12) we have that $t_{(j-i)} = (j-i, \sum_{l=i+1}^j X_l^2)$, from (3.3) it follows that $\theta^{[ij]}|X_{[ij]} \sim \mathcal{IG}\left(\frac{\tau_1^{[ij]*}}{2}, \frac{\tau_0^{[ij]*}}{2}\right)$ in which $\tau_1^{[ij]*} = \tau_1^{[ij]} + \sum_{l=i+1}^j X_l^2$ and $\tau_0^{[ij]*} = \tau_0^{[ij]} + j - i - 2$. Thus, from (2.2) the product estimates of $\theta_s, s = 1, \dots, n$, are given by:

$$E(\theta_s|X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=s}^n r_{[ij]}^* \frac{\tau_1^{[ij]} + \sum_{l=i+1}^j X_l^2}{\tau_0^{[ij]} + j - i - 4}.$$

The posterior relevancies and the posterior distributions of ρ, p, B and A_k can be generated by substituting (3.13) into (2.3).

3.6 Normal data sequence with unknown means and variances

Let $\theta_l = (\mu_l, \sigma_l^2) \in \mathcal{R} \times \mathcal{R}^+$. Assume that, conditional on $\theta_1 = (\mu_1, \sigma_1^2), \dots, \theta_n = (\mu_n, \sigma_n^2)$ the sequence X_1, \dots, X_n are independent and $X_l | \mu_l, \sigma_l^2$ has normal distribution with parameters μ_l and σ_l^2 , denoted by $X_l | \mu_l, \sigma_l^2 \sim \mathcal{N}(\mu_l, \sigma_l^2)$. Consequently, for the block $X_{[ij]}$ we have that

$$f(X_{[ij]} | \mu^{[ij]}, (\sigma^{[ij]})^2) = \left(\frac{1}{2\pi} \right)^{(j-i)/2} \left[\frac{1}{\sigma^{[ij]}} \exp \left\{ -\frac{(\mu^{[ij]})^2}{2(\sigma^{[ij]})^2} \right\} \right]^{j-i} \exp \left\{ \frac{\mu^{[ij]} + \sum_{l=i+1}^j X_l}{(\sigma^{[ij]})^2} - \frac{\sum_{l=i+1}^j X_l^2}{2(\sigma^{[ij]})^2} \right\}. \quad (3.14)$$

Let $\tau^{[ij]} = (\tau_0^{[ij]}, \tau_1^{[ij]}, \tau_2^{[ij]})$. From (3.1) it follows that the conjugate prior distribution for the common parameter $(\mu^{[ij]}, (\sigma^{[ij]})^2)$ is the normal-inverted-gamma distribution with parameters $\frac{\tau_1^{[ij]}}{\tau_0^{[ij]}}$, $\frac{1}{\tau_0^{[ij]}}$, $\frac{1}{2} \left[\tau_2^{[ij]} - \frac{(\tau_1^{[ij]})^2}{\tau_0^{[ij]}} \right]$ and $\frac{\tau_0^{[ij]} - 3}{2}$ denoted by $\theta^{[ij]} \sim \mathcal{NIG} \left(\frac{\tau_1^{[ij]}}{\tau_0^{[ij]}}, \frac{1}{\tau_0^{[ij]}}; \frac{1}{2} \left[\tau_2^{[ij]} - \frac{(\tau_1^{[ij]})^2}{\tau_0^{[ij]}} \right], \frac{\tau_0^{[ij]} - 3}{2} \right)$, whose density function is:

$$f(\theta^{[ij]} | \tau^{[ij]}) = [K(\tau^{[ij]})]^{-1} \left[\frac{1}{\sigma^{[ij]}} \exp \left\{ -\frac{(\mu^{[ij]})^2}{2(\sigma^{[ij]})^2} \right\} \right]^{\tau_0^{[ij]}} \times \exp \left\{ \frac{\mu^{[ij]} \tau_1^{[ij]}}{(\sigma^{[ij]})^2} - \frac{\tau_2^{[ij]}}{2(\sigma^{[ij]})^2} \right\}.$$

where $\tau_0^{[ij]} > 0$, $\tau_1^{[ij]} \in \mathcal{R}$, $\tau_2^{[ij]} > \frac{(\tau_1^{[ij]})^2}{\tau_0^{[ij]}}$ and

$$K(\tau^{[ij]}) = \left(\frac{2\pi}{\tau_0^{[ij]}} \right)^{1/2} \Gamma \left(\frac{\tau_0^{[ij]} - 3}{2} \right) \left(\tau_2^{[ij]} - \frac{(\tau_1^{[ij]})^2}{\tau_0^{[ij]}} \right)^{-\frac{\tau_0^{[ij]} - 3}{2}}. \quad (3.15)$$

Since from (3.14) we have that

$$(\mu^{[ij]}, (\sigma^{[ij]})^2) | X_{[ij]} \sim \mathcal{NIG} \left(\frac{a}{b}, \frac{1}{b}; \frac{1}{2} \left[\tau_2^{[ij]} + \sum_{l=i+1}^j X_l^2 \right] - \frac{a^2}{2b}, \frac{b-3}{a} \right),$$

where

$$a = \tau_1^{[ij]} + \sum_{l=i+1}^j X_l, \quad b = \tau_0^{[ij]} + j - i.$$

Thus, from (2.2) the product estimates of $(\mu_s, (\sigma_s)^2)$, $s = 1, \dots, n$, are given by:

$$E(\mu_s | X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=s}^n r_{[ij]}^* \frac{a}{b}$$

$$E(\sigma_s^2 | X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=s}^n r_{[ij]}^* \frac{(\tau_2^{[ij]} + \sum_{l=i+1}^j X_l^2)b - a^2}{b(b-5)},$$

where, again,

$$a = \tau_1^{[ij]} + \sum_{l=i+1}^j X_l, \quad b = \tau_0^{[ij]} + j - i.$$

The posterior relevancies and the posterior distributions of ρ , p , B and A_k can be generated by substituting (3.15) into (2.3). See also Loschi and Cruz (2005) for a sensitivity analysis to this model.

4. Applications

We shall now illustrate some of the results given in the previous sections by applying them to two real data sets. As an example for the discrete case, we consider the series of counts of violent crimes in a central neighborhood of Belo Horizonte, the capital city of the State of Minas Gerais, Brazil. Given the rate, we assume that the number of crimes are distributed according to a Poisson distribution. For illustrating the continuous case, we consider the return series of four Latin American emerging market indices and assume that, conditional on the expected return and in the volatility (measured here as variance), the returns are normally distributed with unknown mean and variance.

In order to estimate the posterior distributions 10,000 samples of 0–1 values were generated with the dimension of the time series, starting from a sequence of zeros. Since the convergence was reached before the 1,000th step (not shown), the initial 1,000 iterations were discharged as *burn-in*. In order to avoid correlation among vectors a lag of 10 was selected. Further discussions about the number of iterations to be discharged, as well as the lag to be taken, can easily be found in the literature (Gamerman, 1997).

4.1 Change point analysis for criminality data

In this section we consider the series of counts of violent crimes in a central neighborhood (5th CIA) of Belo Horizonte, recorded monthly from January, 1998, to September, 2001. The time series is plotted in Figure 1 jointly with the product

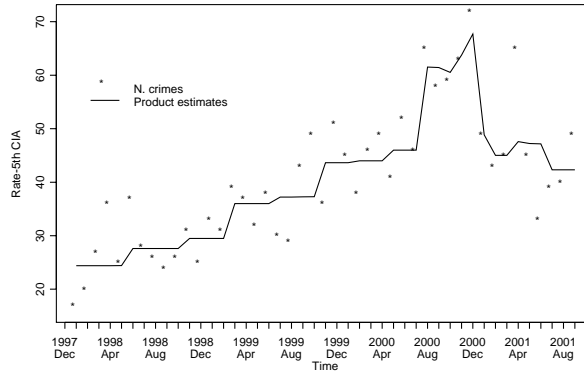
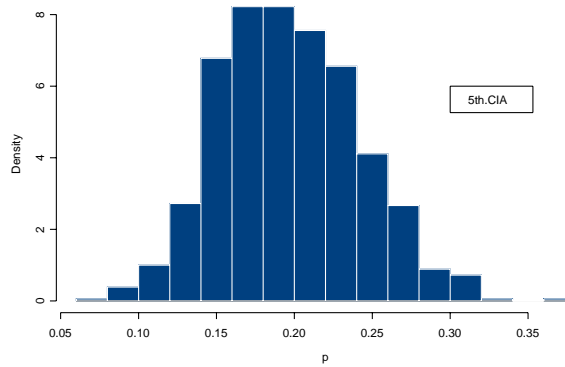


Figure 1: Product estimates for the rate of crimes.

Figure 2: Posterior distribution of p .

estimates for the rate of crimes. Our interest is to verify if the “Policing with Results” a program introduced by the State of Minas Gerais Military Police Command in the late 1990s, made the crime rate to decrease.

We assume that, given the rate, the number of violent crimes within the same block are independent and distributed according to the Poisson distribution $\mathcal{P}(\theta_{[ij]})$. We also adopt the natural conjugate prior distribution for the rate of crimes $\theta_{[ij]}$ which, in this case, is a gamma distribution. As we do not have a precise information about the rate of violent crimes, we assume that $\theta_{[ij]} \sim \mathcal{G}(2.01, 0.001)$ which is a low informative prior distribution and has mean and standard deviation equal to 2.01×10^3 and 1.43×10^3 violent crimes per month, respectively.

Since a small number of changes is expected in the series, we also assume that the probability p of having a change in any instant has a beta prior distribution

Table 1: Descriptive statistics for the posterior distributions of B and p .

parameter	Mean	Median	Standard Deviation	Q1	Q3
B	14.03	14	0.161	14	14
p	0.1971	0.1937	0.0448	0.1645	0.2278

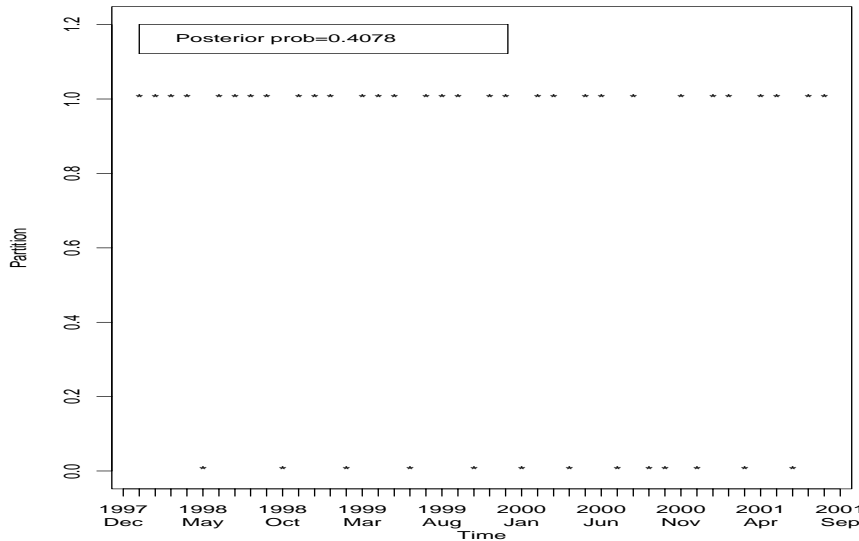


Figure 3: The posterior most probable partition

with parameters $\alpha = 1.5$ and $\beta = 28.5$ (see detail in Section 4.2). Consequently, we are considering in the prior evaluation that the expected number of blocks in the series is 2.2 and the standard deviation is 2.2 blocks.

From Figure 1 we can notice the increase of the rate of crimes from January, 1998, to December, 2000. The rate reaches its maximum in December, 2000 (approximately 68 crimes per month), and decreases in January and June, 2001. These reduction could be a positive effect of the “Policing with Results” program.

Figure 2 presents the posterior distribution for the probability of occurring a crime in any instant p . This distribution is slightly asymmetric and has only one mode. The posterior distribution of p concentrate most of its mass in values close to 0.20.

Notice from Table 1 that, if a square loss function is considered, the posterior estimate for the probability of a crime being committed in any instant is 19.71% which is higher than the prior estimate, 5.0%. The posterior estimate for the number of change points ($B - 1$) is also higher than in the prior evaluation (mean

= 13.03 crimes per month). The posterior variance of B is small which means that there is less uncertainty about B .

From Figures 3 and 4 we can see the posterior most probable partition and the posterior probability of each instant being a change point, respectively. The posterior most probable partition occurs with probability 40.78% and indicates that the following months are change points: June and October, 1998; March, July, and November, 1999; February, May, August, October, and November, 2000; and January, April, and July, 2001. Beside this, we can also observe that the following months has probability higher than 80% of being a change point: July and December, 1998; April, August, and December, 1999; March, June, September, and November, 2000; and March and June, 2001.

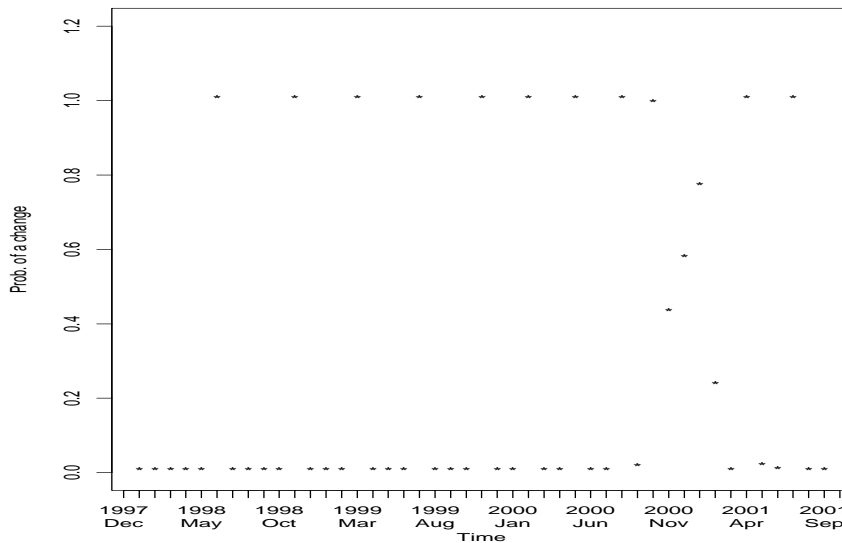


Figure 4: Posterior probability of a change.

4.2 Change point analysis for Latin American emerging indices

In this section we compare the behavior of four important Latin American stock market indices by means of their return series: the Merval of Argentina (*Índice de Mercado de Valores de Buenos Aires*), the IBOVESPA of Brazil (*Índice da Bolsa de Valores do Estado de São Paulo*), the IPSA of Chile (*Índice de Precios Selectivos de Acciones*), and the IPyC of Mexico (*Índice de Precios y Cotizaciones*), from October 31, 1995, to October 31, 2000, recorded fortnightly. A return series is defined by using the transformation $R_t = (P_t - P_{t-1})/P_{t-1}$,

where P_t is the last price observed in the t th fortnight. In this period, three important financial crises involving emerging markets occurred: Mexico's crisis, in January, 1995, Asia's crisis, in August, 1997, and Russia's crisis, in July, 1998. As it is well-known, these events could eventually produce changes in Latin American stock market behavior.

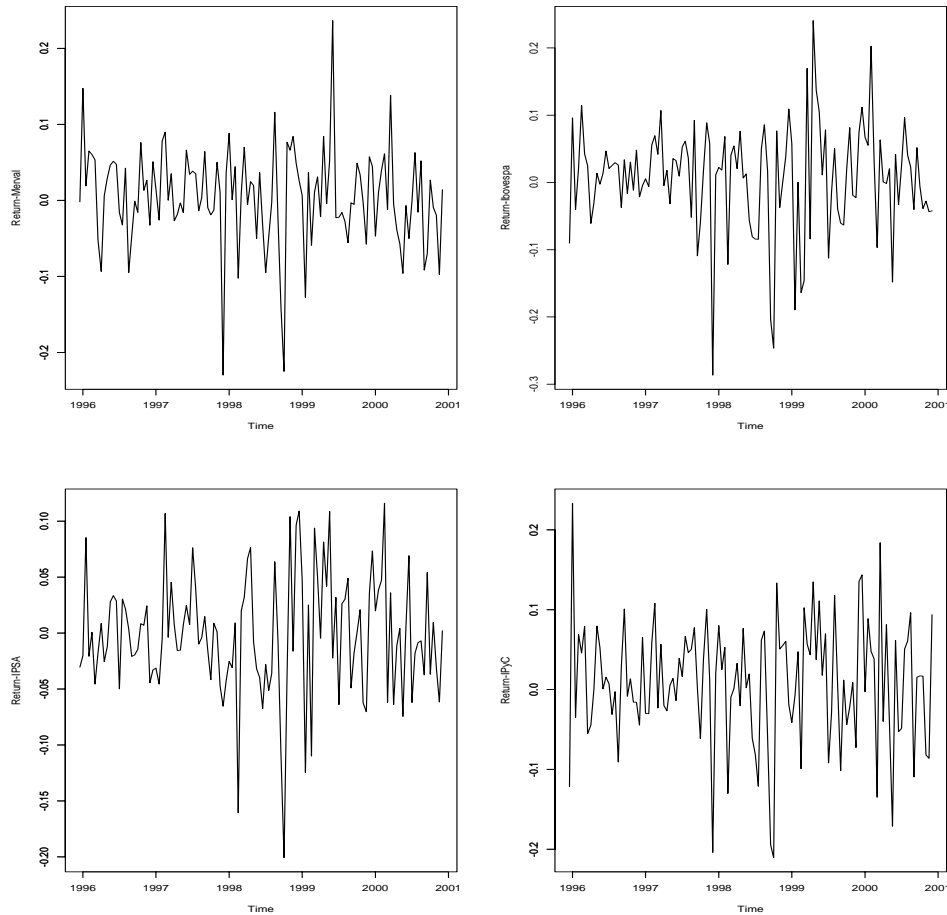


Figure 5: Return series.

It is noticeable from Figure 5 that the behavior of the returns of all these indices suggests the existence of some changes in the variance and in the expected return. In fact, Lochi *et al.* (2005) proposed a PPM to analyze the behavior of the volatility of these indices concluding that they possess volatility clusters. Our purpose here is to extend Loschi *et al.*'s (2005) analysis to show that within the period from 31 October, 1995, to 31 October, 2000, Merval, IBOVESPA, IPSA,

and IPyC series possess both expected return and volatility clusters and also to provide the probability of each instant being a change point for each return series.

We suppose that returns within the same block are conditionally independent and distributed according to the normal distribution $\mathcal{N}(\mu_{[ij]}, \sigma_{[ij]}^2)$. We also adopt the natural conjugate prior distribution for the parameter $(\mu_{[ij]}, \sigma_{[ij]}^2)$ which, in this case, is the normal-inverted-gamma distribution such that

$$\begin{aligned}\mu_{[ij]}|\sigma_{[ij]}^2 &\sim \mathcal{N}(0, \sigma_{[ij]}^2); \\ \sigma_{[ij]}^2 &\sim \mathcal{IG}(a_{[ij]}/2, b_{[ij]}/2),\end{aligned}$$

where $a_{[ij]}$ and $b_{[ij]}$ assume the different values pointed out in Table 2. Consequently, we are assuming that the returns are distributed according to a Student- t PPM, which has heavier tails than the normal distribution and discloses a structure of correlation amongst the returns within the same block. Table 2 presents the descriptive statistics for prior distributions of $\sigma_{[ij]}^2$ for each index.

Table 2: Parameters and descriptive statistics for the prior distribution of the volatility.

Index	$a_{[ij]}$	$d_{[ij]}$	Mean	Variance	Mode
MERVAL, IBOVESPA, and IPyC	0.001	6	2.50×10^{-4}	6.25×10^{-5}	1.25×10^{-4}
IPSA	0.001	8	1.67×10^{-4}	1.39×10^{-5}	1.00×10^{-4}

Notice from Table 2 that we are assuming that in the prior evaluation the IPSA has the smallest volatility (the prior mean and mode are 1.67×10^{-4} and 1.00×10^{-4} , respectively). The volatility for MERVAL, IBOVESPA and IPyC are considered the same (the prior mean and mode are 2.50×10^{-4} and 1.25×10^{-4} , respectively). Observe also that there is less uncertainty about the volatility of IPSA since the variance for the volatility is the smallest.

Since a small number of changes is expected to MERVAL, IBOVESPA, IPSA, and IPyC, we also assume that the probability of having a change in any instant p has a beta prior distribution with parameters $\alpha = 5$ and $\beta = 50$. This prior distribution has modal value equal to 0.076, mean equal to 0.091, standard deviation equal to 3.84×10^{-2} , and concentrate most of its mass in small values of p . We can observe that this prior specification implies that the expected number of blocks in the prior evaluation for all indices is 11.82 and the standard deviation is 5.53 blocks. The modal value of the number of blocks is 10.0 which means that the most probable number of change points in the four indices is 9.0. That is, for us the four indices are equally susceptible to shocks and to the global political atmosphere.

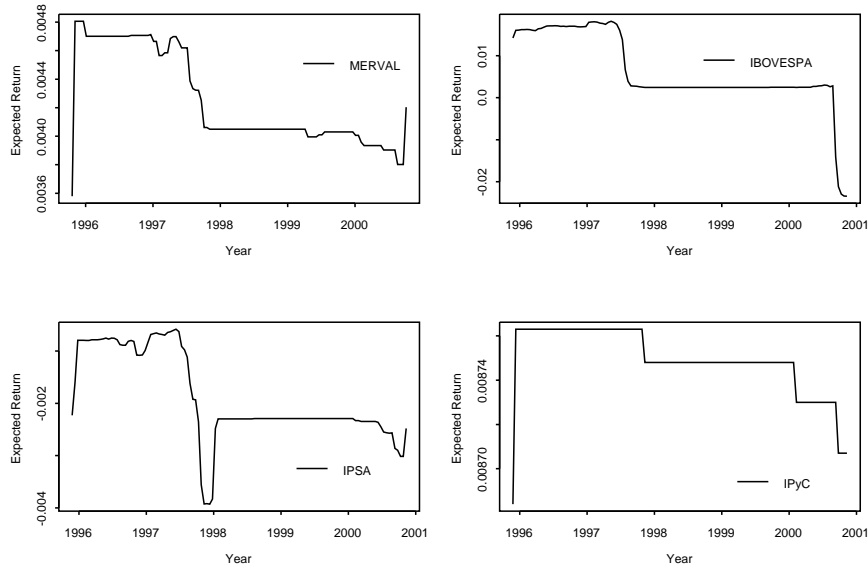


Figure 6: Product estimates for the expected returns.

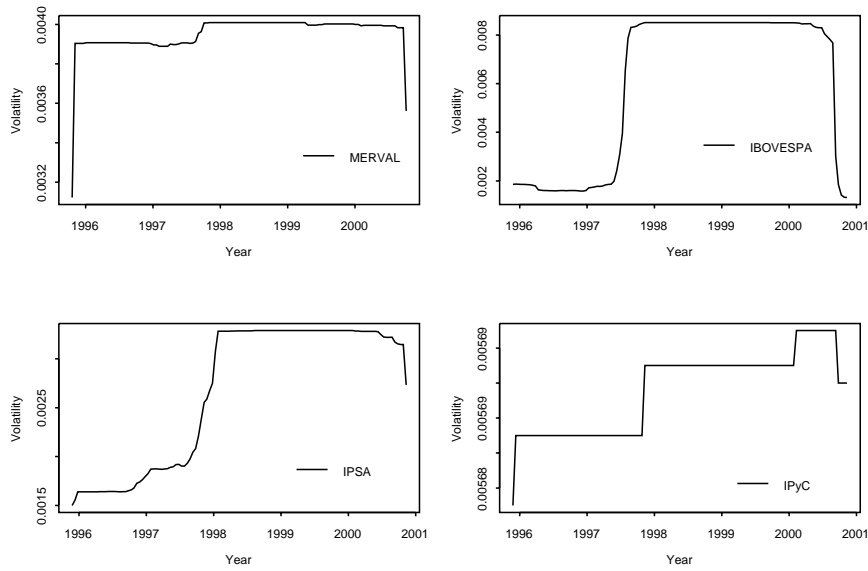


Figure 7: Product estimates for the volatility.

Figures 6 and 7 present the product estimates of the fortnightly expected returns and volatilities for Merval, Ibovespa, Ipsa, and Ipyc return series,

respectively. We can notice that all indices present changes in both expected return and volatility. It is also noticeable that the product estimates for the volatility tends to increase through time and the product estimates for expected return tends to decrease for all indices. Important changes can easily be identified.

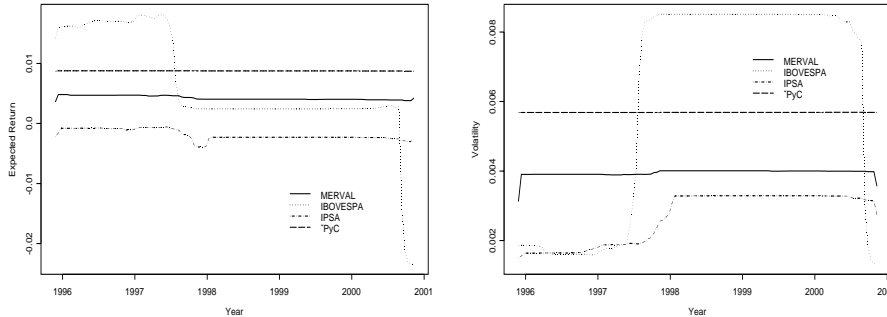


Figure 8: Product estimates for the expected return and volatility plotted altogether.

Table 3: Descriptive statistics for the posterior distribution of the number of blocks.

Index	Mean	Variance	Mode
Merval	1.39	0.3953	1
IBOVESPA	3.44	1.2089	3
IPSA	2.14	1.2309	2
IPyC	1.01	0.0044	1

In general, comparing the four indices we notice from Figure 8 that IPSA presents the smallest volatility and expected return. From the 1st fortnight, July, 1997, IBOVESPA presents the highest volatility and the expected return for IBOVESPA is smaller than that we obtain for Merval and IPyC. Before the 1st fortnight, July, 1997, the product estimates for the volatility of IBOVESPA and IPSA are very close. We also can notice that important changes in all indices occur almost simultaneously.

It is noticeable from Figure 9 and Table 3 that for all indices the posterior distribution for the number of blocks are asymmetric, there are unique modes and they typically concentrate most of their mass in small values. The posterior estimates for the number of blocks in each index are much smaller than we considered in the prior evaluation (mean, 11.82, and mode, 10). That is, these posterior distributions disclose that all four indices are more stable than we considered in the prior evaluation. Moreover, we can perceive that Merval and

IPyC are more stable than IPSA and IBOVESPA, and that, with higher posterior probability, Merval and IPyC do not present change points.

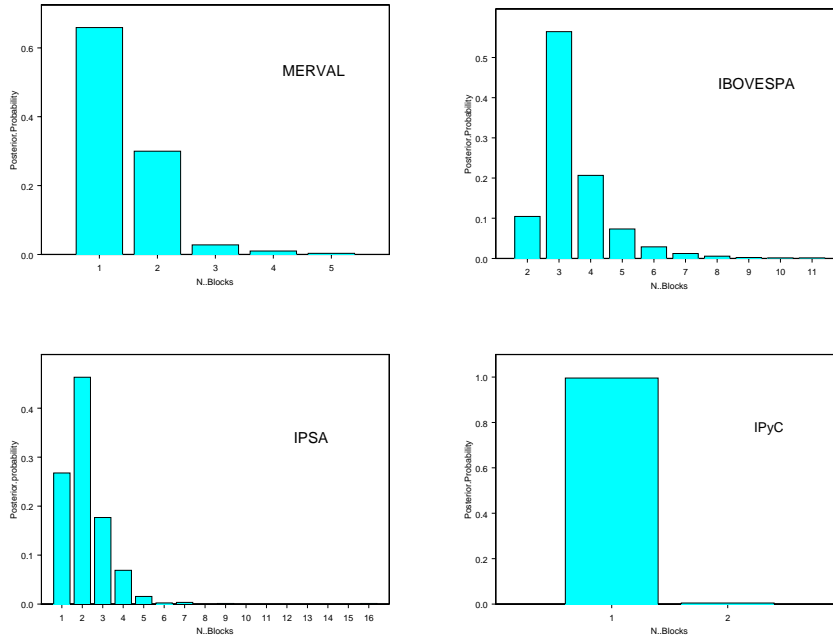


Figure 9: Posterior distribution for the number of blocks.

Table 4 presents the prior and posterior probabilities for the posterior mode of the number of blocks for the four indices. Notice, for example, that in spite of having a low prior probability for Merval and IPyC series to not experience a change point ($2.6 \times 10^{-1}\%$), the posterior probability for these events increase strongly reaching 65.90% and 99.59%, respectively. Assuming the 0-1 loss function, we conclude that IBOVESPA experiences two changes in its behavior with a probability of 56.45% and that only one change point occurs in IPSA series with posterior probability of 46.34%.

Table 4: Prior and posterior probabilities for the posterior mode of B .

Index	Posterior mode	Prior prob.	Posterior prob.
Merval	1	2.60×10^{-3}	0.6590
IBOVESPA	3	1.95×10^{-2}	0.5645
IPSA	2	9.19×10^{-3}	0.4634
IPyC	1	2.60×10^{-3}	0.9956

Table 5: The posterior most probable partition and their prior and posterior probabilities.

Index	Partition	Prior prob.	Posterior prob.
MERVAL	{0,120}	2.60×10^{-3}	0.6589
IBOVESPA	{0,40,115,120}	2.78×10^{-6}	0.1311
IPSA	{0,120}	2.60×10^{-3}	0.2678
IPyC	{0,120}	2.60×10^{-3}	0.9956

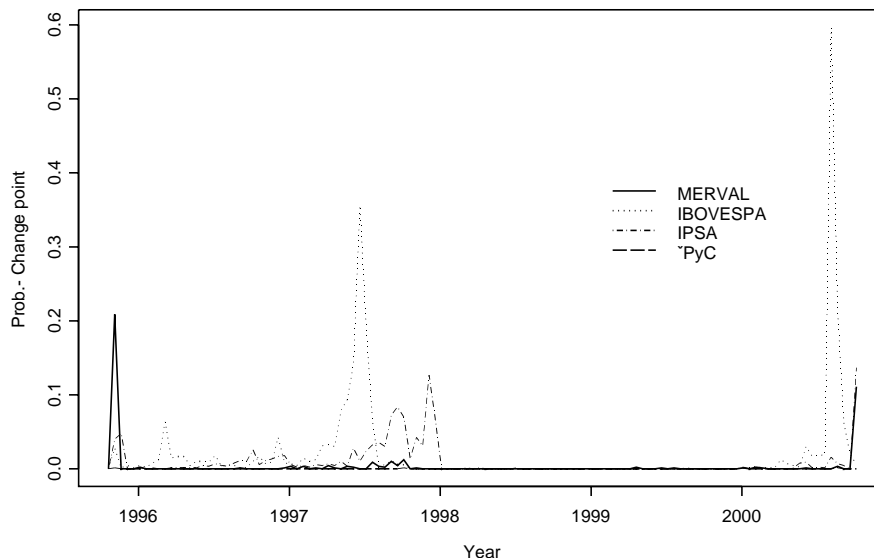


Figure 10: Posterior probability of a change point.

Table 5 presents the posterior most probable partition for Merval, Ibovespa, Ipsa, and Ipyc series. We see that the posterior most probable partitions for Merval, Ipsa, and Ipyc indicate that there are no changes in these indices with probability 65.89%, 26.78% and 99.56%, respectively. The posterior most probable partition for Ibovespa {0, 40, 115, 120} indicates that this index experiences changes in its behavior in the observations 40 (2nd fortnight, June, 1997) and 115 (1st fortnight, June, 2000). Notice that this partition occurs with probability 13.11% only.

Figure 10 presents the probability of each instant being a change point. We notice that for Merval, Ipsa, and Ipyc the probability of each fortnight being a change point is less than 20.89%. For Ibovespa we observe that the 1st

fortnight, July, 1997, is a change point with probability 35.56%, the 2nd fortnight, August, 2000, has probability 59.67% of being a change point and the other instants have probability not superior to 21.34% of being a change point.

Figure 11 and Table 6 present the posterior distributions for p for each index and their descriptive statistics, respectively. Notice that all these distributions are asymmetric, have unique modes, and typically concentrate most of their mass in small values. We can notice for all indices that the posterior estimates for the probability of a change in any instant are smaller than that we have assumed in the prior evaluation (prior mean = 0.0909, for all indices) and it is different for each index. We also notice that the probability p of a change in any instant is highest for IBOVESPA (posterior mean = 0.0439) and is smallest for IPyC (posterior mean = 0.0287).

Notice that the change points identified by the PPM in Merval, IPSA, and IPyC indices are close to an important international event, Asia's crisis in August, 1997.

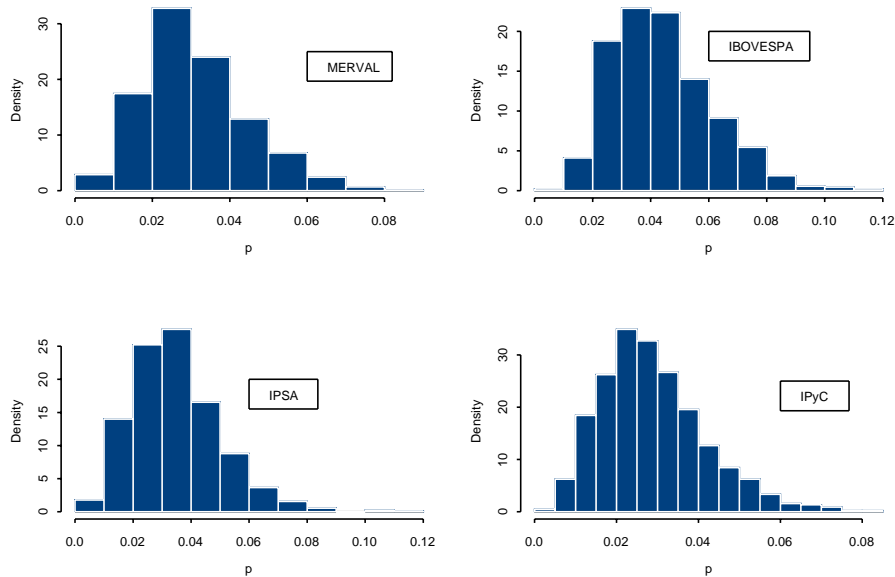


Figure 11: Posterior distribution of p .

Table 5: Descriptive statistics for the posterior distribution of p .

Index	Mean	Median	Standard Deviation
MERVAL	0.0309	0.0289	0.0133
IBOVESPA	0.0439	0.0416	0.0173
IPSA	0.0349	0.0330	0.0152
IPyC	0.0287	0.0266	0.0127

5. Summary and Conclusions

The product partition model (PPM) was defined for Yao's cohesions and applied to the identification of multiple change points in the parameter which index the exponential family. Because the exponential family provides a rich set of models the PPM was also defined the PPM for some particular members of this family.

In order to illustrate the use of PPM, a series of counts of violent crimes in Belo Horizonte, Brazil, and some Latin American emerging market indices (MERVAL-Argentina, IBOVESPA-Brazil, IPSA-Chile, and IPyC-Mexico) were analyzed. We concluded that the rate of violent crimes in a particular neighborhood of Belo Horizonte is high and presents a high number of change points. We also concluded that all indices possess clusters in volatility and expected return and the changes are almost simultaneous. We noticed that the changes experienced by MERVAL and IPyC have less magnitude and, as expected, IPSA felt later the effect of the crisis.

The results indicate that the PPM effective as it may provide useful inferences. Mainly, for the particular data sets analyzed here, it could be observed that the posterior probability of each instant being a change point provides an efficient tool for decision making.

Despite of the good performance of the PPM in the data analyzed here, some other Bayes estimates could be considered in case of the interest is not a retrospective analysis as considered here, see for instance, Quintana and Iglesias (2003).

Acknowledgments

Rosângela H. Loschi acknowledges CNPq (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*) of the Ministry for Science and Technology of Brazil (grants 300325/2003-7 and 472066/2004-8) and PRPq-UFMG (grant4801-UFMG/ RTR/ FUNDO/ PRPq/ RECEMDOUTORES/ 00) for a partial allowance to her research. The research of Frederico R. B. Cruz has been par-

tially funded by the CNPq, grants 301809/96-8, 201046/94-6, 307702/2004-9 and 472066/2004-8, the FAPEMIG (*Fundação de Amparo à Pesquisa do Estado de Minas Gerais*), grants CEX-289/98 and CEX-855/98, and PRPq-UFMG, grant 4081-UFMG/ RTR/ FUNDO/ PRPq/ 99. Reinaldo B. Arellano-Valle acknowledges FONDECYT, grant 1040865.

References

- Barnard, G. A. (1959). Control charts and stochastic processes (with discussion). *Journal of the Royal Statistical Society* **21**, 239-271 .
- Barry, D. and Hartigan, J. A. (1992). Product partition models for change point problems. *The Annals of Statistics* **20**, **1**, 260-279.
- Barry, D. and Hartigan, J. A. (1993). A Bayesian analysis for change point problem. *Journal of the American Statistical Association* **88**, No. 421, 309-319.
- Bernardo, J. M. and Smith, A. F. M. (1994). *Bayesian Theory*. John Wiley & Sons
- Chen, C. W. S. and Lee, J. C. (1995). Bayesian inference of threshold autoregressive models. *Journal of Time Series Analysis* **16**, 483-492.
- Crowley, E. M. (1997). Product partition models for normal means. *Journal of the American Statistical Association* **92**, 192-198.
- Diaconis, P. and Ylvisaker, D., (1979). Conjugate priors for exponential families. *The Annals of Statistics* **7**, 269-281.
- Elliott, M. R. and Shope, J. T., (2003). Use of a Bayesian changepoint model to estimate effects of a graduated driver's licensing program. *Journal of Data Science* **1**, 43-63.
- Gamerman, D., 1997. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall.
- Geweke, J. and Terui, N. (1993). Bayesian threshold autoregressive models for nonlinear time series. *Journal of Time Series Analysis* **14**, 441-454.
- Hartigan, J. A. (1990). Partition models. *Communication in Statistics - Theory & Method* **19**, 2745-2756.
- Hawkins, D. M. (2001). Fitting multiple change-point models to data. *Computational Statistics & Data Analysis* **37**, 323-341.
- Hsu, D. A. (1984) A Bayesian robust detection of shift in the risk structure of stock market returns. *Journal of the American Statistical Association* **77**, 407-416.
- Loschi, R. H., Cruz, F. R. B., Iglesias, P. L., Arellano-Valle, R. B. (2003). A Gibbs sampling scheme to the product partition model: An application to change point problems. *Computers & Operations Research* **30**, 463-482.
- Loschi, R. H., Cruz, F. R. B. (2005). Extension to the product partition model: Computing the probability of a change. *Computational Statistics & Data Analysis* **48**, 255-268.

- Loschi, R. H., Iglesias, P. L. and Moura, C. R. (2005) Bayesian analysis for change points in the volatility of Latin American emerging markets. *Journal of Data Science* **3**, 101-122.
- Menzeffricke, U. (1981). A Bayesian analysis of a change in the precision of a sequence of independent normal random variables at an unknown time point. *Applied Statistics* **30**, 141-146.
- Quintana, F. A. and P. L. Iglesias (2003). Nonparametric Bayesian clustering and product partition models. *Journal of the Royal Statistical Society B* **65**, 557-574.
- Smith, A. F. M. (1975). A Bayesian approach to inference about a change-point in a sequence of random variables. *Biometrika* **62**, 407-416.
- Yao, Y. (1984). Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches. *The Annals of Statistics* **12**, 1434-1447.

Received March 8, 2004; accepted June 17, 2004.

R. H. Loschi
Departamento de Estatística
Universidade Federal de Minas Gerais
31270-901 - Belo Horizonte - MG, Brazil
loschi@est.ufmg.br

F. R. B. Cruz
Departamento de Estatística
Instituto de Ciências Exatas
Universidade Federal de Minas Gerais
31270-901 - Belo Horizonte - MG, Brazil
fcruz@est.ufmg.br

R. B. Arellano-Valle
DEpartamento de Estadística
Facultad de Matemática
Pontificia Universidad Católica de Chile
Casilla 306, Santiago 22, Chile
reivalle@mat.puc.cl