

AVALIAÇÃO DE UMA MEDIDA DE EVIDÊNCIA DE UM PONTO DE MUDANÇA E SUA UTILIZAÇÃO NA IDENTIFICAÇÃO DE MUDANÇAS NA TAXA DE CRIMINALIDADE EM BELO HORIZONTE

Rosangela H. Loschi *

Flávio B. Gonçalves

Frederico R. B. Cruz

Departamento de Estatística

Universidade Federal de Minas Gerais (UFMG)

Belo Horizonte – MG

loschi@est.ufmg.br

* *Corresponding author* / autor para quem as correspondências devem ser encaminhadas

Recebido em 09/2004; aceito em 06/2005 após 1 revisão

Received September 2004; accepted June 2005 after one revision

Resumo

A probabilidade *a posteriori* de um instante ser um ponto de mudança foi proposta por Loschi & Cruz (2005) como uma medida de evidência de que o comportamento de uma seqüência de dados mude em tal instante. A proposta deste trabalho é avaliar a eficiência desta medida na identificação de mudanças na taxa da distribuição Poisson, em dados seqüencialmente observados e compará-la com a medida proposta por Hartigan (1990), isto é, com a probabilidade *a posteriori* da partição aleatória formada pelos pontos de mudança. Cenários ou seqüências de dados com e sem pontos de mudanças são considerados. Em cenários sem pontos de mudanças, assumem-se taxas pequenas e grandes para avaliar a eficiência da medida proposta na presença de pouca e muita variabilidade. Em cenários com pontos de mudanças, consideram-se tanto mudanças estruturais quanto observações atípicas. Conclui-se que, em geral, a medida proposta teve melhor desempenho para identificar pontos de mudança. Uma análise para dados de criminalidade da cidade de Belo Horizonte também é feita utilizando-se o modelo proposto e observou-se que esta taxa muda freqüentemente ao longo do tempo.

Palavras-chave: amostrador de Gibbs; modelo partição produto; mudança estrutural; observação atípica.

Abstract

The posterior probability of each instant being a change point was proposed by Loschi & Cruz (2005) as a measure of evidence that a change occurred in that instant. In this paper, the efficiency of such a measure is evaluated and it is compared with the measure proposed by Hartigan (1990). Poisson data sequences, with and without changes, are considered. The method is evaluated in the presence of both structural changes and atypical observations. The proposed method over performs Hartigan's proposal. A series of violent crimes data from Belo Horizonte is analyzed by the proposed method. It was observed that the rate of violent crimes changes frequently throughout the time.

Keywords: Gibbs sampling; product partition model; structural change; atypical observation.

1. Introdução

O modelo partição produto (MPP) introduzido por Hartigan (1990) em sua forma mais geral é, entre os propostos na literatura (por exemplo, os modelos *threshold* e o método considerado por Hawkins, 2001), o mais flexível para o estudo de problemas de ponto de mudança (mudança estrutural ou observação atípica). Sua flexibilidade se deve ao fato de que considera como variáveis aleatórias tanto o número de pontos de mudança como também os instantes em que as mudanças ocorreram (instantes estes identificados por uma partição aleatória do conjunto de índices das observações). Por esta razão, este modelo vem sendo amplamente utilizado. Veja, por exemplo, Barry & Hartigan (1993), Crowley (1997), os quais aplicam o MPP para a identificação de pontos de mudança na média de dados normalmente distribuídos, Quintana & Iglesias (2003), para uma aplicação do MPP na identificação de *outliers* em modelos de regressão, Loschi *et al.* (2003), para a identificação de pontos de mudança na média e variância de dados seqüencialmente observados e normalmente distribuídos entre outros.

Ao estudar problemas de pontos de mudanças utilizando o MPP, constata-se, no entanto, que a probabilidade *a posteriori* da partição aleatória que indica os instantes em que as mudanças ocorreram – medida de evidência de mudança sugerida por Hartigan (1990) e Barry & Hartigan (1992) – pode não fornecer uma idéia precisa sobre os instantes em que, de fato, tais mudanças ocorreram. Este problema surge por que várias partições podem ocorrer com mesma probabilidade *a posteriori* dificultando a tomada de decisão sobre os instantes em que as mudanças ocorreram. Numa tentativa de solucionar este problema surgiram duas alternativas. Quintana & Iglesias (2003) propõem uma maneira de escolher a partição ótima usando técnicas de Teoria de Decisão e, posteriormente, Loschi & Cruz (2005) propõem avaliar a evidência de uma mudança ter ocorrido calculando a probabilidade *a posteriori* de cada instante ser um ponto de mudança.

Este artigo estende para o caso Poisson os resultados obtidos em Loschi & Cruz (2005) e apresenta um estudo sobre a eficiência desta medida na identificação de pontos de mudança. Serão consideradas seqüências de dados gerados da distribuição de Poisson, independentemente. Seqüências ou cenários contendo ou não pontos de mudanças são considerados. Em cenários sem pontos de mudanças, consideram-se seqüências com diferentes variabilidades e em cenários com uma ou mais mudanças estudam-se tanto mudanças estruturais quanto observações atípicas. Para ilustrar a metodologia, utiliza-se o modelo proposto na análise de dados de criminalidade registrados na 17^a Cia da Polícia Militar de Belo Horizonte (região que inclui o campus da UFMG). O interesse é estudar o comportamento da taxa de criminalidade nesta região ao longo do tempo e avaliar se o programa instituído pelo Comando da Polícia Militar do Estado de Minas Gerais no final de 1999 foi eficiente para reduzir a criminalidade na área.

Este trabalho está assim organizado. Na seção 2, apresenta-se o MPP, em sua versão paramétrica introduzida por Barry & Hartigan (1992), com as coesões de Yao (1984) e os métodos computacionais utilizados para a obtenção das distribuições *a posteriori* de interesse. Na seção 3, avalia-se o comportamento e a eficiência da medida proposta para detecção de pontos de mudança. Para este estudo, consideram-se dados gerados, independentemente, de distribuições de Poisson. Na seção 4, aplica-se o modelo proposto para analisar o comportamento da taxa de criminalidade da cidade de Belo Horizonte de janeiro de 1998 a setembro de 2001. Para finalizar, na seção 5, apresentam-se as principais conclusões.

2. Modelo Estatístico

O MPP foi introduzido em sua forma mais geral por Hartigan (1990). Na sua forma geral, este modelo considera, em sua estrutura, todas as partições possíveis do conjunto de dados levando em conta tanto blocos contíguos quanto não-contíguos. Posteriormente, Barry & Hartigan (1992) particularizam o MPP para a situação em que os dados são seqüencialmente observados e apenas blocos contíguos são permitidos. Neste contexto eles introduzem uma versão paramétrica para o MPP em que os blocos presentes na partição são induzidos pela igualdade dos parâmetros que indexam as distribuições das observações. Para ilustrar o uso do MPP, Barry & Hartigan (1993) aplicam o MPP em sua versão paramétrica para identificar mudanças na média de dados seqüencialmente observados e normalmente distribuídos com variância comum.

Nesta seção serão brevemente apresentados o MPP e sua aplicação à seqüência de dados distribuídos segundo uma distribuição de Poisson e também o método computacional utilizado para obter-se as distribuições *a posteriori* de interesse.

2.1 Modelo partição produto para coesões de Yao

Seja X_1, \dots, X_n uma seqüência de dados observados e considere o conjunto de índices $I = \{1, \dots, n\}$. Denote por $\rho = \{i_0, \dots, i_b\}$ a partição aleatória do conjunto de índice I tal que $0 = i_0 < i_1 < \dots < i_b = n$ e por B a variável aleatória que denota o número de blocos em ρ . Assuma que cada partição divide a seqüência X_1, \dots, X_n em $B = b$ subseqüências contíguas, as quais serão denotadas aqui por $X_{[ij]} = (X_{i+1}, \dots, X_j)'$. Seja C_{ij} a coesão *a priori* associada ao bloco $[ij] = \{i+1, \dots, j\}$, para $i, j \in I \cup \{0\}$, e $j > i$. Por coesão entende-se o grau de similaridade que se julga existir entre as observações em um mesmo bloco, que pode ser interpretada, aqui, como as probabilidades de transição na cadeia de Markov definida por pontos de mudança (Barry & Hartigan, 1992).

Considere p , para $0 \leq p \leq 1$, a probabilidade de que uma mudança ocorra em um instante qualquer. A coesão *a priori* para o bloco $[ij]$ proposta por Yao (1984) e dada por:

$$C_{ij} = \begin{cases} p(1-p)^{j-i-1}, & j < n, \\ (1-p)^{j-i-1}, & j = n, \end{cases} \quad (1)$$

para todo $i, j \in I, i < j$. Essas coesões *a priori* implicam que a seqüência de pontos de mudança estabelece um processo de renovação, com tempos de ocorrência idêntica e geometricamente distribuídos. Conseqüentemente, tais coesões podem ser utilizadas quando acredita-se que pontos de mudança já ocorridos não são informativos sobre pontos de mudanças futuros.

Seja $\theta_1, \dots, \theta_n$, uma seqüência de parâmetros desconhecidos, condicionalmente nos quais, as variáveis aleatórias, X_1, \dots, X_n , têm densidades condicionais marginais dadas por $f_1(X_1 | \theta_1), \dots, f_n(X_n | \theta_n)$, respectivamente. A distribuição *a priori* de $\theta_1, \dots, \theta_n$ é construída como segue. Dado uma partição $\rho = \{i_0, \dots, i_b\}$, para $b \in I$, tem-se que $\theta_i = \theta_{[i_{r-1}i_r]}$, para todo $i_{r-1} < i \leq i_r, r = 1, \dots, b$, e que $\theta_{[i_0i_1]}, \dots, \theta_{[i_{b-1}i_b]}$ são independentes entre si e de p . Assuma que

$\theta_{[ij]}$ tem distribuição *a priori* $\pi_{[ij]}(\theta)$, $\theta \in \Theta_{[ij]}$, em que $\Theta_{[ij]}$ é o espaço paramétrico correspondente ao parâmetro comum $\theta_{[ij]}$ que indexa a densidade condicional de $X_{[ij]}$, ou seja, $\theta_{[ij]} = \theta_{i+1} = \dots = \theta_j$. Então, seguindo Barry & Hartigan (1992) define-se o modelo partição produto para coesões *a priori* de Yao como segue:

1. Dado p , a distribuição *a priori* de ρ é a seguinte distribuição produto:

$$P(\rho = \{i_0, i_1, \dots, i_b\} | p) = p^{b-1}(1-p)^{n-b}, \tag{2}$$

para toda partição $\{i_0, \dots, i_b\}$, satisfazendo $0 = i_0 < i_1 < \dots < i_b = n$.

2. condicionalmente em $\rho = \{i_0, \dots, i_b\}$ e p , a seqüência X_1, \dots, X_n é independente de p e tem densidade conjunta dada por:

$$f(X_1, \dots, X_n | \rho = \{i_0, \dots, i_b\}, p) = \prod_{j=1}^b f_{i_{(j-1)}i_j}(X_{i_{(j-1)}i_j}), \tag{3}$$

em que $f_{[ij]}(X_{[ij]}) = \int_{\Theta_{[ij]}} f_{[ij]}(X_{[ij]} | \theta) \pi_{[ij]}(\theta) d\theta$.

Barry & Hartigan (1992) também mostram que a esperança *a posteriori* (ou estimativa produto) para θ_k , $k=1, \dots, n$, é dada por:

$$E(\theta_k | X_1, \dots, X_n) = \sum_{i=0}^{s-1} \sum_{j=0}^n r^*_{[ij]} E(\theta_k | X_{[ij]}), \quad s=1, \dots, n, \tag{4}$$

em que $r^*_{[ij]} = P([ij] \in \rho | X_0, \dots, X_n)$ denota a relevância *a posteriori* para o bloco $[ij]$.

Assuma que p tem distribuição *a posteriori* Beta com parâmetros α e β . Como consequência dessa consideração, segue que as distribuições *a posteriori* de ρ e B são, respectivamente, dadas por:

$$P(\rho = \{i_0, \dots, i_b\} | X_1, \dots, X_n) \propto \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + b - 1)\Gamma(n + \beta - b)}{\Gamma(n + \beta + \alpha - 1)} \prod_{j=1}^b f_{i_{(j-1)}i_j}(X_{i_{(j-1)}i_j}), \tag{5}$$

$$P(B = b | X_1, \dots, X_n) \propto \binom{n-1}{b-1} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + b - 1)\Gamma(n + \beta - b)}{\Gamma(n + \beta + \alpha - 1)} \prod_{j=1}^b f_{i_{(j-1)}i_j}(X_{i_{(j-1)}i_j}).$$

A distribuição *a posteriori* para p é dada por:

$$\pi(p | X_1, \dots, X_n) \propto \sum \prod_{j=1}^b f(X_{i_{j-1}i_j}) p^{b+\alpha-2} (1-p)^{n-b+\beta-1}, \tag{6}$$

em que a soma é sobre todas as partições de $\{1, \dots, n\}$ em b blocos com pontos finais em i_0, i_1, \dots, i_b , satisfazendo à condição $0 = i_0 < i_1 < \dots < i_b = n$. A probabilidade *a posteriori* de cada instante i ser um ponto de mudança é:

$$P(A_i | X_1, \dots, X_n) = \sum_{C_i} \prod_{j=1}^n f(X_{i_{j-1}i_j}) \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + b - 1)\Gamma(n + \beta - b)}{\Gamma(n + \beta + \alpha - 1)}, \tag{7}$$

em que C_i é o conjunto de todas as partições que incluem o i -ésimo instante como ponto de mudança. Para maiores detalhes ver Loschi & Cruz (2005).

2.2 Estimativas produto para a taxa da distribuição Poisson

Assuma que, dadas as taxas em cada instante k , $\theta_1, \dots, \theta_n$, as observações, X_1, \dots, X_n , são independentes e tais que $X_k | \theta_k \sim \text{Poisson}(\theta_k)$, para $k=1, \dots, n$. Assuma também que, *a priori*, o parâmetro comum $\theta_{[ij]}$, que indexa a distribuição das observações do bloco $[ij]$, tem distribuição gama com parâmetros $\tau > 0$ e $\varphi > 0$, a qual será denotada por $\theta_{[ij]} \sim \text{Gama}(\tau, \varphi)$ e cuja função densidade é dada por:

$$f(\theta_{[ij]} | \tau, \varphi) = \frac{\varphi^\tau}{\Gamma(\tau)} \cdot (\theta_{[ij]})^{\tau-1} \exp(-\varphi\theta_{[ij]}).$$

Como consequência tem-se que a distribuição preditiva *a priori* de $X_{[ij]}$, necessária para o cálculo das distribuições *a posteriori* exibidas na seção anterior, é a seguinte:

$$f(X_{[ij]}) = \prod_{k=i+1}^j \frac{1}{X_k!} \frac{\Gamma(\tau^*)}{\Gamma(\tau)} \left(\frac{\varphi}{\varphi^*}\right)^\tau \left(\frac{1}{\varphi^*}\right)^{\sum_{k=i+1}^j X_k},$$

em que

$$\begin{cases} \varphi^* = \varphi + j - 1, \\ \tau^* = \tau + \sum_{k=i+1}^j X_k, \end{cases}$$

para todo $i = 0, \dots, n-1$ e $j = i+1, \dots, n$. Também segue como consequência das suposições feitas que, dadas as observações do $X_{[ij]}$, a distribuição *a posteriori* por bloco de $\theta_{[ij]}$ é a distribuição gama com parâmetro τ^* e φ^* , isto é, $\theta_{[ij]} | X_{[ij]} \sim \text{Gama}(\tau^*, \varphi^*)$. Conseqüentemente, as estimativas *a posteriori* ou *estimativas produto* para a taxa θ no instante k , θ_k , para todo $k=1, \dots, n$ são dadas por :

$$\hat{\theta}_k = E(\theta_k | X_1, \dots, X_n) = \sum_{i=0}^{k-1} \sum_{j=k}^n \frac{\tau^*}{\varphi^*} r_{[ij]}^* \quad (8)$$

2.3 Métodos computacionais

Considere a quantidade aleatória auxiliar U_l , de forma que $U_l = 1$, se $\theta_l = \theta_{l+1}$ e $U_l = 0$, se $\theta_l \neq \theta_{l+1}$ para $l = 1, \dots, n-1$. Note que a partição aleatória ρ é imediatamente identificada ao considerar vetores $U = (U_1, \dots, U_{n-1})$ dessas quantidades aleatórias. Cada partição $(U_1^s, \dots, U_{n-1}^s)$, $s \geq 1$, é gerada usando o amostrador de Gibbs como segue. Partindo do valor inicial $(U_1^0, \dots, U_{n-1}^0)$, no passo s , o r -ésimo elemento, U_r^s , é gerado considerando a seguinte razão:

$$R_r = \frac{f_{[xy]}(X_{[xy]}) \int_0^1 p^{b-2} (1-p)^{n-b+1} d\pi(p)}{f_{[xv]}(X_{[xv]}) f_{[ry]}(X_{[ry]}) \int_0^1 p^{b-1} (1-p)^{n-b} d\pi(p)}, \quad r = 1, \dots, n-1. \quad (9)$$

Então, cada amostra da distribuição *a posteriori* de p é gerada da seguinte distribuição:

$$p^s | X_1, \dots, X_n \sim p^{b^s-1} (1-p)^{n-b^s} \pi(p),$$

para $s \geq 1$, em que b^s é o número de blocos no s -ésimo vetor U e é obtido notando-se que o número de blocos em ρ é dado por:

$$B = 1 + \sum_{l=1}^{n-1} (1 - U_l).$$

As estimativas produto podem ser obtidas computando a relevância *a posteriori* do bloco $[ij]$, $i, j = 1, \dots, n$, $i < j$, como segue:

$$r^*_{[ij]} = \frac{M}{T}, \tag{10}$$

e a estimativa da probabilidade *a posteriori* de cada instante i ser um ponto de mudança utilizando os métodos computacionais citados acima é:

$$P(A_i) = \frac{N}{T} \tag{11}$$

para $i = 1, \dots, n-1$, em que N é o número de vetores U para os quais é observado que $U_{i-1} = 0$, M é o número de vetores U para os quais é observado $U_i = 0$, $U_{i+1} = \dots = U_{j-1} = 1$, e $U_j = 0$ e T é o número total de vetores U gerados. Ver detalhes em Loschi *et al.* (2003) e Loschi & Cruz (2005). Veja Gamerman (1997) para uma explicação geral sobre o amostrador de Gibbs.

3. Avaliação da Medida de Evidência de uma Mudança

Nesta seção comparam-se as probabilidades *a posteriori* de (i) cada instante ser ponto de mudança (Loschi & Cruz, 2004) e (ii) da partição aleatória formada pelos instantes de mudança (Hartigan, 1990). O objetivo é avaliar se a medida proposta por Loschi & Cruz (2005) tem a mesma eficiência que a medida proposta por Hartigan (1990) na identificação de pontos de mudança, sejam estas observações atípicas, sejam mudanças estruturais.

Serão simuladas várias seqüências de dados de tamanho 50 com nenhum, um, dois ou quatro pontos de mudanças. Os dados serão gerados, independentemente, de diferentes distribuições de Poisson. Como distribuição *a priori* de p para os casos em que se tem nenhuma ou uma mudança será considerada uma distribuição Beta(2,98); para os casos em que se tem duas mudanças, considera-se distribuição Beta(2,48) e, para os casos em que se tem quatro mudanças, distribuição Beta(2,23) será considerada. Estas distribuições *a priori* concentram a maior parte de suas massas em valores pequenos de p , indicando que, *a priori*, espera-se que a probabilidade de mudança em um instante qualquer seja baixa. Além disto, se $p \sim \text{Beta}(\alpha, \beta)$, o número esperado de blocos *a priori* é dado por:

$$B = (n-1) \frac{\alpha}{\alpha + \beta}.$$

Sendo assim, tem-se como consequência que, *a priori*, ocorre 1, 2 e 4 pontos de mudanças na seqüência de dados, respectivamente. Assume-se que a distribuição *a priori* para o parâmetro θ , em todos os casos, é a distribuição Gama(2,1) que tem esperança e variância iguais a 2. Esta distribuição foi escolhida de tal forma que seu valor médio fosse um valor intermediário entre os valores escolhidos para a taxa θ da qual geraram-se as distribuições Poisson e de tal forma que fosse mais próxima do valor da taxa da qual mais freqüentemente geraram-se os dados de uma mesma seqüência.

Em cada cenário, serão analisados três gráficos. No primeiro são apresentados a taxa real (linha cheia), a taxa estimada (linha tracejada) e os dados observados (pontos). No segundo são exibidas as distribuições *a priori* e *a posteriori* de p . No terceiro, mostra-se a partição mais provável (pontos) e a probabilidade de cada instante ser um ponto de mudança (linha cheia). Deve ser lembrado, neste terceiro gráfico, que um ponto no valor zero significa que uma mudança ocorreu. A Figura 1 apresenta os resultados para cenários sem mudança quando $\theta = 0,5$ e $\theta = 10$, respectivamente.

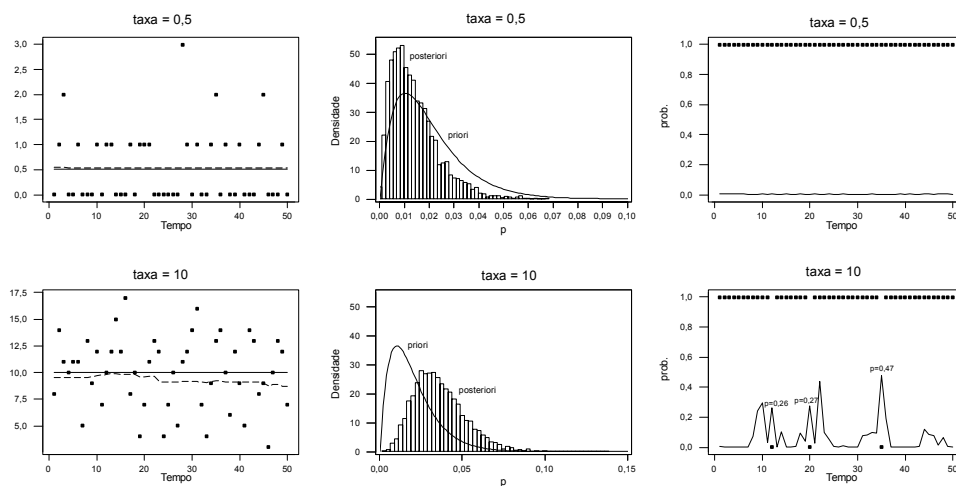


Figura 1 – Cenário sem mudança.

No caso em que $\theta = 0,5$ nota-se que as probabilidades de cada instante ser um ponto de mudança é muito baixa em todos os instantes, confirmando que θ não sofre mudanças. Percebe-se, também, que a partição mais provável em cada caso indica que o comportamento da taxa θ não sofre mudanças ao longo do tempo. Nota-se que tal partição tem alta probabilidade *a posteriori* de ocorrência (probabilidade 91%). No caso $\theta = 10$, a partição mais provável indicou a ocorrência de três mudanças, as quais são registradas nos instantes 12, 20 e 35. Porém, a probabilidade de ocorrência desta partição foi de apenas 14,8%. Além disso, nos instantes em que a partição detectou mudanças, as probabilidades *a posteriori* de tais instantes serem pontos de mudança são de apenas 26%, 28% e 47%, respectivamente, indicando para os dois primeiros uma probabilidade muito baixa de que sejam, de fato, pontos de mudança. Observa-se, ainda, que a probabilidade *a posteriori* de que o instante 22 seja um ponto de mudança foi 45% o que indica haver uma grande incerteza sobre este ponto ser ou não um ponto de mudança. Ressaltamos que este ponto não foi indicado pela partição

mais provável como sendo um ponto de mudança. Para os demais instantes observou-se uma probabilidade menor que 25% de ser ponto de mudança indicando forte evidência que tais instantes não são pontos de mudança. Nota-se ainda que as estimativas produto para a taxa, nos dois casos, são bem próximas do valor real, apresentando mudança apenas no caso em que a taxa (variabilidade) é alta. Vê-se também que as distribuições *a posteriori* de p indicam, com era de se esperar, pequena probabilidade *a posteriori* que uma mudança ocorra em um instante qualquer.

A Figura 2 mostra o resultado obtido para cenários contendo apenas uma mudança estrutural nos instantes 10 e 25, respectivamente. Neste caso, as observações antes da mudança são geradas de uma distribuição Poisson com parâmetro $\theta = 1$ e as observações após a mudança são geradas de uma distribuição com taxa $\theta = 4$.

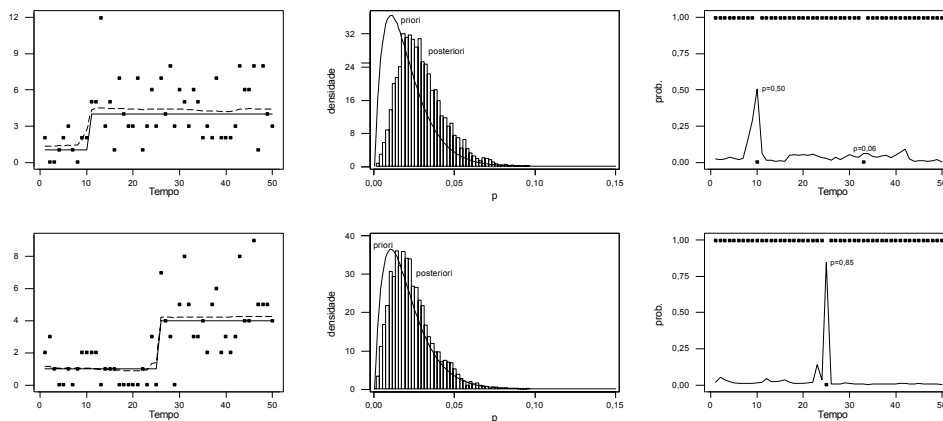


Figura 2 – Cenário com uma mudança estrutural.

No cenário em que θ sofre mudança no instante 10, a partição mais provável *a posteriori* indica erroneamente mudanças nos instantes 10 e 33. Porém tal partição tem probabilidade de ocorrência *a posteriori* de apenas 2%. No entanto, considerando a probabilidade *a posteriori* de cada instante ser ponto de mudança, nota-se que o instante 10 é um ponto de mudança com probabilidade 50%. Os demais instantes (inclusive o 33) têm probabilidade inferior 10% de ser ponto de mudança. Para o cenário com mudança no instante 25 nota-se que a partição mais provável *a posteriori* indica corretamente os instantes de mudança e esta partição ocorre com probabilidade *a posteriori* de 52%. Nota-se também que este instante tem probabilidade *a posteriori* de 85% de ser ponto de mudança e os demais instantes tem probabilidade *a posteriori* inferior a 15% de serem pontos de mudança. Ou seja, neste último cenário a distribuição *a posteriori* de p e a probabilidade *a posteriori* de cada instante ser ponto de mudança conduzem à mesma decisão. Porém, a probabilidade *a posteriori* de cada instante ser ponto de mudança fornece uma melhor informação. Percebe-se que quanto maior for a quantidade de observações antes das mudanças, mais eficientes na indicação de mudanças são os dois métodos considerados. Resultados similares (não mostrados) foram obtidos quando, após a mudança, geramos os dados de uma distribuição com maior variância.

A Figura 3 mostra o resultado obtido para cenários contendo duas mudanças estruturais nos instantes 10 e 20, e 30 e 40, respectivamente. Neste caso, as observações antes da primeira

mudança e após a segunda mudança são geradas de uma distribuição Poisson com parâmetro $\theta = 1$, e entre as mudanças são geradas de uma distribuição com taxa $\theta = 4$.

No primeiro caso, em que θ sofre mudança nos instantes 10 e 20, a partição mais provável *a posteriori* tem probabilidade 11% de ter ocorrido e indica mudanças nos instantes 9 e 20. Para os instantes 09 e 20 observe-se que suas probabilidades *a posteriori* de serem ponto de mudança são 57% e 79%, respectivamente. No instante 10 esta probabilidade é 23% e nos demais instantes são menores que 15%. No outro caso, a partição mais provável *a posteriori* indicou mudanças apenas nos instantes em que estas realmente ocorrem com probabilidade 19%. Neste caso, as probabilidades *a posteriori* de que os instantes 20 e 30 sejam pontos de mudança são 42% e 43%, respectivamente, e menor que 15% nos demais instantes. Nota-se ainda que, as estimativas produto para a taxa, em cada cenário, são bem razoáveis e, como era esperado, as distribuições *a posteriori* de p indicam que a probabilidade de que uma mudança ocorra em um instante qualquer é maior que nos casos anteriores.

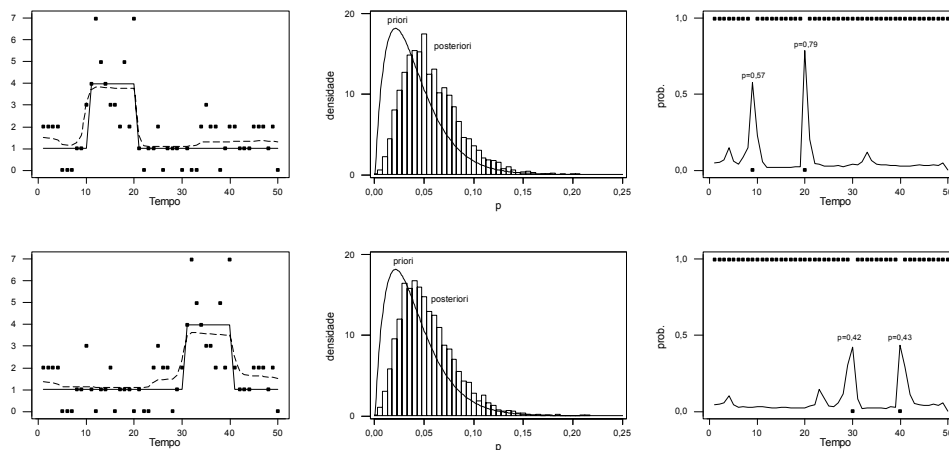


Figura 3 – Cenário com duas mudanças estruturais.

A Figura 4 mostra o resultado obtido para cenários contendo dois pontos atípicos nos instantes 24 e 26, e 21 e 29, respectivamente. Os pontos atípicos são gerados de uma distribuição Poisson com parâmetro $\theta = 10$ e as demais observações são geradas considerando taxa $\theta = 1$.

No caso em que os pontos atípicos estão nos instantes 24 e 26, a partição mais provável *a posteriori* indicou mudanças apenas nos instantes 23 e 26, com probabilidade 9%, e observou-se que as probabilidades *a posteriori* de que estes instantes sejam pontos de mudança são de 79% e 98%, respectivamente. Nos instantes 24 e 25, quando também ocorrem mudanças, as probabilidades *a posteriori* de mudança foram 3,1% e 2,7%, respectivamente, e, para os demais instantes, esta probabilidade foi sempre menor que 20%. Perceba que, neste caso, as estimativas produto dão uma idéia de que houve uma mudança estrutural e não duas observações atípicas muito próximas. No outro caso a partição, mais provável *a posteriori* indicou mudanças exatamente nos instantes em que elas ocorrem, com probabilidade 1%, e a probabilidade *a posteriori* de mudança nestes instantes foi sempre maior que 61%. Observa-se então que foram necessárias pelo menos três observações entre

os pontos atípicos para detectar-se todas as mudanças, tanto quando consideramos a partição mais provável *a posteriori*, como quando utilizamos a probabilidade *a posteriori* do instante ser ponto de mudança. Perceba também que a probabilidade *a posteriori* de que uma mudança ocorra em um instante qualquer é de cerca de 10%.

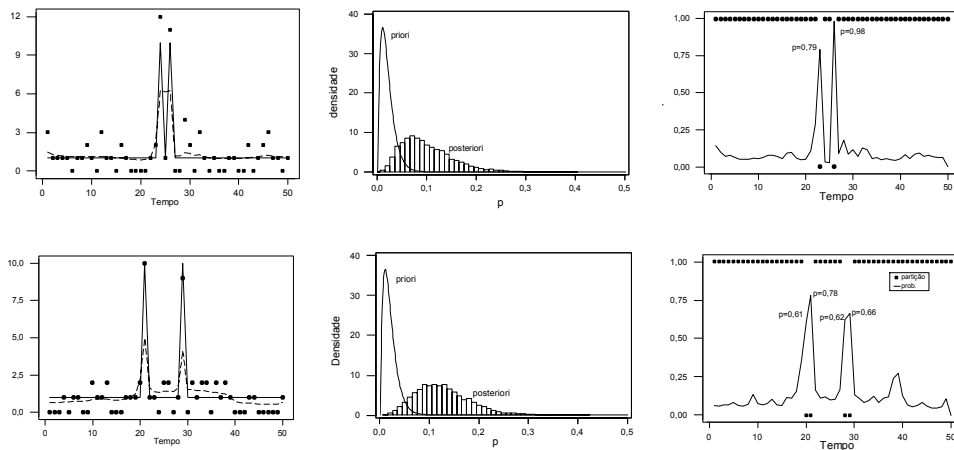


Figura 4 – Cenário com dois pontos atípicos.

A Figura 5 mostra o resultado obtido para cenários contendo duas mudanças estruturais nos instantes 15 e 30, considerando valores crescentes para a taxa. No primeiro (segundo) cenário as observações antes da primeira mudança são geradas de uma distribuição Poisson com parâmetro $\theta = 1$ ($\theta = 1$), após a primeira mudança são geradas de uma distribuição com taxa $\theta = 4$ ($\theta = 6$) e após a segunda mudança são geradas de uma distribuição com $\theta = 7$ ($\theta = 12$).

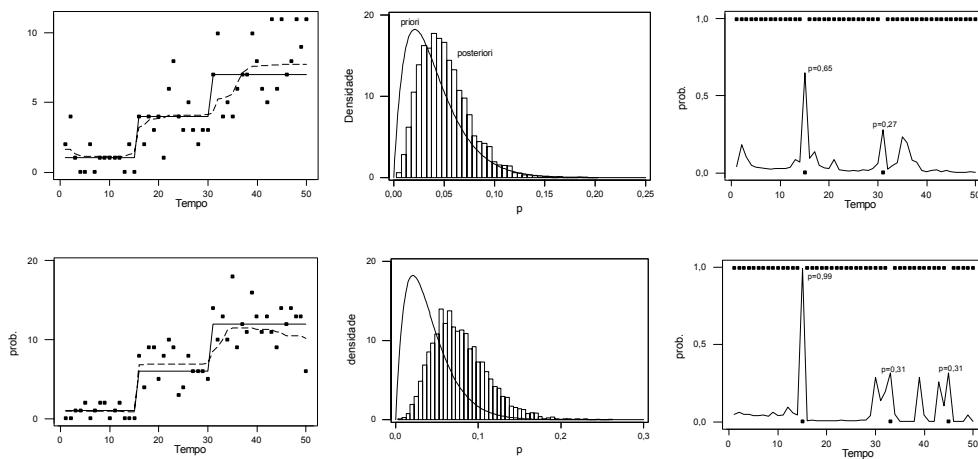


Figura 5 – Cenário com duas mudanças e taxa crescente.

No primeiro caso a partição mais provável *a posteriori* indicou mudanças nos instantes 15 e 31 com probabilidade 7%. A probabilidade *a posteriori* de que estes instantes sejam pontos de mudança é 65% e 27%, respectivamente. No instante 30 esta probabilidade é 6% e é menor que 25% para os demais instantes. Nota-se que a partição mais provável *a posteriori* detectou a segunda mudança com um pequeno atraso e a probabilidade *a posteriori* de mudança neste instante não foi alta (27%), significando forte evidência de que o instante 30 não seja um ponto de mudança. Isto deve ter ocorrido pelo fato de as mudanças sofridas por θ terem sido pequenas (de 1 para 4 e de 4 para 7). No segundo caso, a partição mais provável *a posteriori* indica mudanças nos instantes 15, 33 e 46 e ocorre com probabilidade 13%. As probabilidades *a posteriori* de que estes instantes sejam pontos de mudança são, respectivamente, 99%, 31% e 31%. Observa-se também que o instante 30 é ponto de mudança com probabilidade de apenas 29%. Neste caso, a probabilidade da primeira mudança ocorrer foi bem maior que no primeiro caso, porém, além da segunda mudança ter sido detectada com atraso, foi detectada uma terceira mudança que não ocorreu na realidade. Nota-se que para os casos em que θ sofre dois aumentos, a segunda mudança foi detectada com atraso pela partição mais provável *a posteriori* e teve probabilidade *a posteriori* relativamente baixa de ser um ponto de mudança. Perceba, ainda que, as estimativas produto para a taxa são bem razoáveis e captam relativamente bem as mudanças.

Concluindo, pode-se observar através do estudo dos diferentes cenários simulados que a medida proposta teve um bom desempenho geral na detecção de pontos de mudança. Notou-se que na maioria dos instantes em que ocorreu mudança, a probabilidade de tal instante ser um ponto de mudança foi maior que 0,50. Percebe-se ainda que as duas medidas de evidência de uma mudança são úteis na análise de pontos de mudança sendo que a medida proposta teve desempenho melhor ou igual ao da partição mais provável *a posteriori*. Na próxima seção, aplica-se o MPP a dados reais relacionados com a criminalidade na cidade de Belo Horizonte.

4. Análise de Dados Reais

Nesta seção utiliza-se a metodologia apresentada na seção 2 para analisar dados referentes ao número de crimes violentos notificados mês a mês no período de janeiro de 1998 a setembro de 2001 na 17ª Cia. da Polícia Militar da cidade de Belo Horizonte, que é responsável pelo registro dos crimes cometidos no campus da UFMG e em alguns bairros próximos. O objetivo principal é verificar se o programa “Policinando com Resultados”, introduzido pelo Comando da Polícia Militar do Estado no final de 1999 (Ward, 2000), produziu uma redução na taxa de criminalidade nesta região.

Uma vez que o único evento do qual se tinha conhecimento que poderia ter provocado uma mudança na taxa de criminalidade era a implementação do programa “Policinando com Resultados”, escolheu-se como distribuição *a priori* para p (probabilidade de ocorrer uma mudança em um instante qualquer) uma distribuição Beta com parâmetros $\alpha=1,5$ e $\beta=28,5$. Esta distribuição reflete que, *a priori*, a probabilidade de ocorrer uma mudança em um instante qualquer é pequena (média = 0,05 e moda = 0,0178) e tem-se uma grande certeza sobre isto (variância = 0,0015). Conseqüentemente, assume-se que, *a priori*, o número médio de pontos de mudança na série analisada é 1. Assume-se também que, dado a taxa de criminalidade θ , o número de crimes cometidos nessa região em cada mês possui uma distribuição de Poisson. Desde que não se tem informação prévia sobre a taxa de criminalidade em Belo Horizonte, adota-se que *a priori* θ tem uma distribuição Gama,

pouco informativa, com parâmetros $\tau = 0,001$ e $\phi = 1,01$. Conseqüentemente, estima-se que, *a priori*, a taxa média de crimes violentos cometidos nesta região de Belo Horizonte é de 1010 crimes por mês. A Figura 6 mostra as estimativas produto da taxa de crimes violentos registrados na 17ª Cia. Nota-se que as estimativas produto para θ apresentam comportamento crescente no período de janeiro de 1998 a junho de 2000, atingindo seu máximo neste último mês (82 crimes). No período que vai de junho de 1998 a fevereiro de 1999, a taxa de criminalidade apresenta maior estabilidade e é de aproximadamente 20 crimes por mês. Após junho de 2000 percebe-se uma redução na velocidade de crescimento da taxa de criminalidade, a qual poderia ser uma conseqüência do programa “Policinando com Resultados”. Em junho de 2000 a taxa de criminalidade sofre uma queda significativa caindo de 82 para 58 crimes por mês. Em março de 2001 sofre um forte aumento passando de 50 para 76 crimes por mês. A partir de maio de 2001, a taxa começa a decrescer atingindo a taxa de 45 crimes por mês em setembro desse mesmo ano.

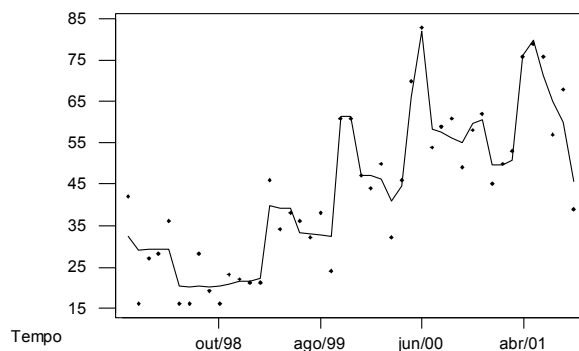


Figura 6 – Estimativas produto da taxa de criminalidade.

A Figura 7 apresenta a probabilidade *a posteriori* de cada mês ser um ponto de mudança (linha contínua) juntamente com a partição mais provável *a posteriori*. A partição mais provável *a posteriori* é indicada através de pontos em que um valor zero representa uma mudança naquele mês. Da Figura 7 percebe-se que, no período junho de 1998 a fevereiro de 1999, a probabilidade de cada instante ser um ponto de mudança é no máximo 26%. Mudanças importantes na taxa podem ser observadas em maio de 1998, com probabilidade 89%, fevereiro, setembro e novembro de 1999 com probabilidades 98%, 100% e 100%, respectivamente, abril, maio, junho e dezembro de 2000 com probabilidades 71%, 100%, 87% e 98%, respectivamente, março, abril, maio, e agosto de 2001, com probabilidades 93%, 100%, 100% e 59%, respectivamente. Observa-se que nos instantes em que grandes mudanças nas estimativas produto da taxa de criminalidade (Figura 6) são percebidas, as probabilidades destes instantes serem pontos de mudança são bem altas. Já no período em que se observa estabilidade nas estimativas da taxa, tais probabilidades são baixas (menores que 26%). Ou seja, a medida proposta está em acordo com o que é observado no gráfico das estimativas produto para a taxa de crimes violentos. A partição mais provável *a posteriori* ocorreu com probabilidade 1,9% e a maioria dos meses indicados como instantes de mudança por esta partição coincide com os meses cuja probabilidade *a posteriori* de ser ponto de mudança é maior do que 50%. Ou seja, as duas medidas levaria a conclusões similares sobre os instantes em que as mudanças ocorreram.

Percebe-se da Figura 8 que a probabilidade de que uma mudança ocorra em um instante qualquer é, *a posteriori*, maior do que se acreditava *a priori*. Observe que a distribuição *a posteriori* de p está concentrada em valores mais altos que a distribuição *a priori*. Note que, *a posteriori*, em média, p é 26%, o que significa dizer que a taxa de criminalidade está mais propensa a mudanças do que se acreditava inicialmente. A distribuição *a posteriori* do número de blocos (figura não apresentada) indicou a presença de mais pontos de mudança do que se esperava *a priori*. Notou-se que, em média, *a posteriori* há 16,55 pontos de mudança com desvio padrão igual a 0,549.

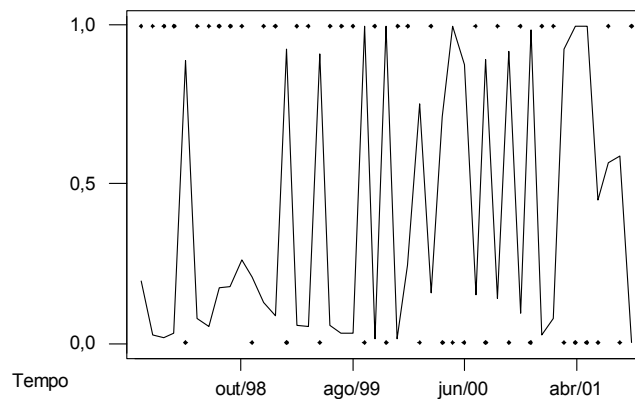


Figura 7 – Probabilidade de uma mudança e partição mais provável, *a posteriori*.

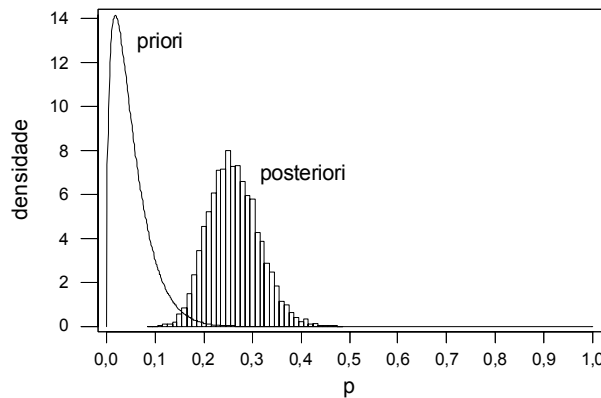


Figura 8 – Distribuições *a priori* e *a posteriori* de p .

Concluindo, sob nosso ponto de vista, o programa “Policinando com Resultados” foi pouco eficiente para a redução da criminalidade na região estudada uma vez que, apesar de observamos que a taxa de criminalidade de meados de 2000 em diante cresce com menor velocidade, a taxa de crimes violentos na região ainda está muito alta.

5. Conclusão

Neste trabalho estendeu-se para o modelo Poisson o estudo da eficiência da probabilidade *a posteriori* de cada instante ser ponto de mudança, como uma medida de evidência de que a mudança, de fato, ocorreu. Esta medida, proposta por Loschi & Cruz (2005), foi comparada com a partição mais provável *a posteriori*, medida proposta originalmente por Hartigan (1990) e Barry & Hartigan (1992). A avaliação foi feita considerando seqüências de dados Poisson gerados independentemente, com e sem a presença de mudanças. Para cenários com pontos de mudança, avaliou-se o comportamento da medida de Loschi & Cruz (2005) na presença tanto de observações atípicas, quanto de quebras estruturais. Após este estudo, apresentou-se uma análise para dados de criminalidade da cidade de Belo Horizonte.

Notou-se que, em geral, a medida de Loschi & Cruz (2005) teve desempenho não inferior para identificar a presença de pontos de mudança. Percebeu-se que na maioria dos instantes em que ocorreu mudança, a probabilidade *a posteriori* de tal instante ser um ponto de mudança foi maior que 0,50. Observou-se ainda que, quanto maior for a quantidade de observações antes das mudanças, mais eficientes na identificação de pontos de mudança são tanto a partição mais provável *a posteriori*, quanto a probabilidade *a posteriori* de um instante ser ponto de mudança. Em seqüência de dados em que observações atípicas ocorrem muito próximas umas das outras, as mudanças podem não ser bem identificadas por qualquer das duas medidas. Além disto, nestes cenários, as estimativas produto podem dar a idéia errônea de que ocorreu uma mudança estrutural. Notou-se também que para cenários em que a taxa θ sofre mudanças estruturais aumentando o seu valor em cada mudança, nenhum dos dois métodos teve bom desempenho. Nestes casos, notou-se que as estimativas produto nem sempre estão próximas do valor real da taxa.

Na análise da série de número de crimes violentos, registrados mensalmente, cometidos em uma região da cidade de Belo Horizonte, observou-se que a taxa de criminalidade teve tendência ao crescimento de janeiro de 1998 até junho de 2000. Após esta data, notou-se que a taxa de criminalidade passou a crescer com menor velocidade. Esta diminuição na velocidade de crescimento pode ter sido conseqüência do programa “Policinando com Resultados” introduzido pelo Comando da Polícia Militar de Minas Gerais no final de 1999. Contrário ao que se esperava na avaliação *a priori*, a taxa de criminalidade na região estudada experimentou muitas mudanças no período estudado e continuou alta mesmo após o programa “Policinando com Resultados” ter sido implantado. Conclui-se que a metodologia é eficiente na identificação de pontos de mudança.

Agradecimento

Os autores agradecem ao CRISP-UFMG na pessoa do Prof. Renato Martins Assunção (EST-UFMG) pelo fornecimento dos dados utilizados no trabalho. Rosângela H. Loschi agradece ao CNPq (processos 3000325/2003-7 e 472066/2004-8) e PRPq-UFMG (processo 3893-UFMG/RTR/FUNDO/PRPq/RECEMDOUTORES/00), por parcialmente financiar sua pesquisa. Flávio B. Gonçalves é bolsista do Programa de IC do CNPq. Frederico R. B. Cruz tem sua pesquisa financiada por CNPq (processos 301809/96-8, 472066/2004-8 e 201046/94-6), FAPEMIG (processos CEX-289/98 e CEX-855/98) e PRPq-UFMG (processo 4081-UFMG/RTR/FUNDO/PRPq/99).

Referências Bibliográficas

- (1) Barry, D. & Hartigan, J.A. (1993). A Bayesian analysis for change point problem. *Journal of the American Statistical Association*, **88**(421), 309-319.
- (2) Barry, D. & Hartigan, J.A. (1992). Product partition models for change point problems. *The Annals of Statistics*, **20**(1), 260-279.
- (3) Crowley, E.M. (1997). Product partition models for normal means. *Journal of the American Statistical Association*, **93**(437), 192-198.
- (4) Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall, London, UK.
- (5) Hartigan, J.A. (1990). Partition Models. *Communication in Statistics-Theory & Methods*, **19**(8), 2745-2756
- (6) Hawkins, D.M. (2001). Fitting multiple change-point models to data. *Computational Statistics & Data Analysis*, **37**(3), 323-341.
- (7) 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**(2), 255-268.
- (8) Loschi R.H.; Cruz, F.R.B; Iglesia, 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**(3), 463-482.
- (9) Quintana, F. & Iglesias, P.L. (2003). Bayesian Clustering and product partition models. *Journal of the Royal Statistical Society B*, **65**(2), 557-574.
- (10) Ward, H.H. (2000). Poling reform in Latin America: Current efforts in Argentina, Brazil and Chile. *Woodrow Wilson Center for Scholars. Latin America Program*, Dezembro de 2000, 29 pp. (Manuscrito não publicado).
- (11) Yao, Y. (1984). Estimation of a noise discrete-time step function: Bayes and empirical Bayes approaches. *The Annals of Statistics*, **12**(4), 1434-1447.