

**DETECÇÃO DE MÚLTIPLOS PONTOS DE MUDANÇA EM SÉRIES TEMPORAIS  
ATRAVÉS DE UMA ABORDAGEM FUZZY/BAYESIANA****Marcos Flávio Silveira Vasconcelos D'Angelo**Departamento de Ciência da Computação – UNIMONTES  
Av. Rui Braga, s/n, Vila Mauriceia, Montes Claros-MG  
[marcos.dangelo@unimontes.br](mailto:marcos.dangelo@unimontes.br)**Reinaldo Martínez Palhares**Departamento de Engenharia Eletrônica – UFMG  
Av. Antônio Carlos, 6627 – Pampulha, 31270-901 – Belo Horizonte – MG  
[palhares@cpdee.ufmg.br](mailto:palhares@cpdee.ufmg.br)**Renato Dourado Maia**Departamento de Ciência da Computação – UNIMONTES  
Av. Rui Braga, s/n, Vila Mauriceia, Montes Claros-MG  
[renato.dourado@unimontes.br](mailto:renato.dourado@unimontes.br)**Camila Katheryne Santos Cangussu**Bacharelado em Engenharia de Sistemas – UNIMONTES  
[camila.katheryne@gmail.com](mailto:camila.katheryne@gmail.com)**Lucas Almeida Aguiar**Bacharelado em Engenharia de Sistemas – UNIMONTES  
[lucas.tamoios@gmail.com](mailto:lucas.tamoios@gmail.com)**RESUMO**

Este trabalho aborda o problema de detecção de vários pontos de mudança em séries temporais. Para lidar com esse problema, uma nova abordagem fuzzy/Bayesiana foi proposta. A principal vantagem dessa abordagem em relação às apresentadas na literatura é o fato de que ela não exige o conhecimento de informação prévia, como, por exemplo, o número de pontos de mudança ou a caracterização da distribuição probabilística dos dados.

**PALAVRAS CHAVE. Agrupamentos fuzzy. Metropolis-Hastings. Pontos de Mudança.****OA - Outras aplicações em PO.****ABSTRACT**

This paper addresses the problem of detecting multiple change points in time series. To handle this problem a new fuzzy/Bayesian approach is proposed which differs from previous approaches since it does not require prior information as: the number of change points or the characterization of the data probabilistic distribution.

**KEYWORDS. Fuzzy Clustering. Metropolis-Hastings. Change Points.****OA - Other applications in OR.**

## 1. Introdução

Métodos tradicionais de construção de modelos para séries temporais são baseados em técnicas estatísticas, objetivando a escolha de modelos que expliquem satisfatoriamente seu comportamento dinâmico. Porém, algumas questões podem ser levantadas:

- Existe mudança de regime na série?
- Pode um único modelo retratar essa dinâmica para todo o conjunto de dados?

Se existirem mudanças significativas na série temporal, parece natural encontrar pontos de mudança antes de modelar o processo inteiro.

Vários são os trabalhos que abordam o problema de detecção de pontos de mudança, como em séries financeiras (Oh, 2005), séries ecológicas (Beckage, 2007), série de índice de criminalidade (Loschi, 2005c), detecção de falhas (D'Angelo et al., 2011a), (D'Angelo et al., 2011b), etc. As principais técnicas para detecção de pontos de mudança apresentadas na literatura são baseadas em testes estatísticos e análise *bayesiana*. O teste estatístico mais comum no problema de detecção de pontos de mudança é o CUSUM e no contexto de análise *bayesiana* são largamente utilizados os métodos MCMC.

O teste CUSUM proposto por (Hinkey, 1971) é largamente utilizado na detecção de pontos de mudança, e aplicações desse método, bem como suas modificações e extensões, podem ser vistas em (Hadjiliadis, 2006), (Lee, 2006a), e (Lee, 2006b).

No contexto de análise *bayesiana*, tem-se como referência o trabalho de (Barry, 1993), que utiliza o modelo partição-produto (mpp) para a identificação de pontos de mudança na média de dados com distribuição normal. Em (Loschi, 2005b), a probabilidade *a posteriori* de um instante ser um ponto de mudança é utilizada como medida de evidência de que o comportamento de uma sequência de dados mudou em algum instante. Em (Loschi, 2005a), é avaliada a eficiência da medida na identificação de mudanças, proposta em (Loschi, 2005b), na taxa da distribuição Poisson, em dados sequencialmente observados, sendo feita a comparação com a medida proposta por (Hartigan, 1990).

Entretanto, todos esses trabalhos exigem algum conhecimento *a priori* do comportamento estatístico da série temporal, como, por exemplo, qual tipo de distribuição que melhor representa seu comportamento dinâmico. Neste trabalho será utilizada uma abordagem que independe desse conhecimento *a priori* sobre a série temporal, considerando a demonstração empírica apresentada (D'Angelo et al., 2011c) que séries temporais, após sofrerem uma transformação definida por operações *fuzzy*, resultam séries que podem ser adequadamente aproximadas por séries com distribuição *beta*. A parametrização da distribuição *beta*, portanto, é utilizada para substituir o conhecimento *a priori* a respeito da série temporal. Além disso, a metodologia foi estendida para detecção de múltiplos pontos de mudança, e não somente um ou dois pontos de mudança, como em (D'Angelo et al., 2011c) e (Moreira et al., 2010), respectivamente.

## 2. Quantização da Série Temporal

A teoria dos conjuntos *fuzzy*, proposta por (Zadeh, 1965), tem recebido muita atenção recentemente, não somente no âmbito de desenvolvimentos teóricos, mas também de aplicações. Uma de suas principais aplicações ocorre em métodos de agrupamentos, visto que os métodos clássicos de agrupamentos separam os dados em  $k$  categorias e, em muitos casos, alguns elementos não podem pertencer a uma categoria específica, pois pertencem a duas ou mais categorias simultaneamente. Utilizar métodos de agrupamentos *fuzzy* é uma boa maneira de resolver esse problema, pois, diferentemente do que ocorre na abordagem clássica, um elemento pode pertencer a mais de uma categoria simultaneamente.

Nesta seção é proposta uma alternativa de quantização da série temporal, através de agrupamento *fuzzy*, para ser utilizada no algoritmo Metropolis-Hastings, como mostrado a seguir.

**Definição (Agrupamento fuzzy):** Seja  $y(t)$  uma série temporal e considere um inteiro positivo  $k$ . Se o conjunto  $C = \{C_i \mid \min\{y(t)\} \leq C_i \leq \max\{y(t)\}, i = 1, \dots, k\}$  resolve o problema de minimização:

$$\min \sum_{i=1}^k \sum_{\mu_i(t) \in C_i} \|\mu_i(t) - C_i\|^2, \quad (1)$$

então  $C = \{C_i, i = 1, 2, \dots, k\}$  é o conjunto de centros da série temporal  $y(t)$ . Em (1),

$$\mu_i(t) \triangleq \left[ \sum_{j=1}^k \frac{\|y(t) - C_i\|^2}{\|y(t) - C_j\|^2} \right]^{-1} \quad (2)$$

é o grau de pertinência *fuzzy* de  $y(t)$  em relação a cada centro  $C_i$ . ■

Repare que, tendo em conta um conjunto  $C$  de centros, a tarefa de medir a distância de cada ponto da série temporal  $y(t)$  à cada centro  $C_i$  é simples e o problema é encontrar os centros que resolvam (1). Para resolver esse problema existem várias abordagens, por exemplo *K-means* (Kaufman, 1990), *C-means* (Bezdek, 1981) e rede de Kohonen (Kohonen, 2001), sendo que em todas essas abordagens é necessário o conhecimento prévio da quantidade de grupos, fato esse que justificou a utilização de um algoritmo adaptativo, em que essa informação não é necessária, como colocado na Seção 3.

Para fins de ilustração, no decorrer do trabalho será utilizada a série temporal (3):

$$f(t) = \begin{cases} p_1 + 0.1 * \mathcal{E}(t) - 0.1 * \mathcal{E}(t - 1), & \text{se } t \leq m_1; \\ p_2 + 0.1 * \mathcal{E}(t) - 0.1 * \mathcal{E}(t - 1), & \text{se } m_1 < t \leq m_2; \\ \vdots \\ p_k + 0.1 * \mathcal{E}(t) - 0.1 * \mathcal{E}(t - 1), & \text{se } t > m_{k-1}; \end{cases} \quad (3)$$

sendo  $p_1$  o primeiro ponto de operação,  $p_2$  o segundo ponto de operação,  $p_k$  o  $k$ -ésimo ponto de operação,  $\mathcal{E}(t)$  um ruído com distribuição  $\pi(\cdot)$  no instante  $t$  e  $m_i$  os pontos de mudança.

A Figura 1 ilustra a série temporal (3) com  $p_1 = 1$ ,  $p_2 = 2$ ,  $p_3 = 3$ ,  $p_4 = 4$ ,  $p_5 = 5$ ,  $m_1 = 30$ ,  $m_2 = 60$ ,  $m_3 = 90$ ,  $m_4 = 120$  e  $\mathcal{E}(t) \sim U(0,1)$  para 150 observações da série temporal, e os graus de pertinência  $\mu_i(t)$ . Nota-se que no primeiro e no último grau de pertinência tem-se uma única mudança, que será identificada utilizando a abordagem proposta por (D'Angelo, 2011c), e nos graus de pertinência intermediários tem-se duas mudanças, que serão identificadas na abordagem proposta por (Moreira, 2010).

A seguir é descrito o procedimento de quantização da série temporal:

1. Apresente a série temporal  $y(t)$ ;
2. Encontre um conjunto  $k$  de centros,  $C = \{C_i \mid \min\{y(t)\} \leq C_i \leq \max\{y(t)\}, i=1,2,\dots,k\}$ , que minimize a distância *euclidiana* conforme (1), considerando, por exemplo, a série temporal (3).
3. Calcule o grau de pertinência *fuzzy* como dado em (2), para cada amostra da série temporal,  $y(t)$ , em relação a cada centro  $C_i$  (como ilustrado na Figura 1 considerando, por exemplo, a série temporal (3)).

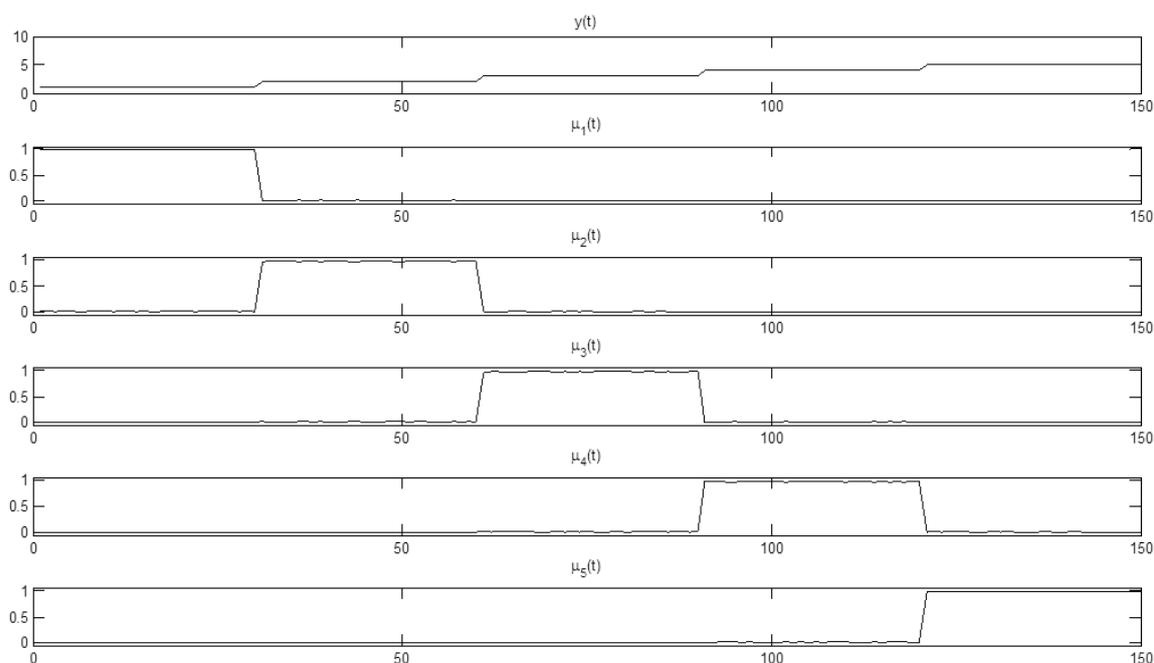


Figura 1 – Série Temporal com as respectivas pertinências para cada centro

### 3. Algoritmo de Agrupamento

Para resolver o problema de otimização proposto em (1), é necessária a utilização de um algoritmo de agrupamento adaptativo, isto é, que encontre, além dos centros dos grupos, a quantidade de grupos. Várias são as abordagens que tratam esse problema, como por exemplo o *Quality Threshold Clustering* (Heyer, 1999) ou o *K-means* Adaptativo, como proposto em (Chinrungrueng, 1995). No presente trabalho, a Largura Média de Silhueta, proposta inicialmente por (Rousseeuw 1987), foi utilizada para resolver esse problema, sendo que a mesma mede a qualidade de um agrupamento com base na proximidade entre os elementos de padrões semelhantes e na distância entre elementos de padrões distintos. Para resolver o problema de otimização dessa medida, deve-se escolher agrupamentos que preservem proximidade entre itens de um mesmo agrupamento e excluam de um mesmo grupo elementos com grandes distâncias entre si.

O algoritmo baseia-se nos seguintes passos:

1. Dado um conjunto de dados ordenados  $S = \{s_1, s_2, s_3, \dots, s_n\}$  espalhados em um espaço unidimensional, seja  $D = \{d_1, d_2, d_3, \dots, d_{n-1}\}$  o conjunto das distâncias entre pontos adjacentes de  $S$ , e  $\lambda$  o número máximo de grupos;
2. Encontre um ponto para segmentação do grupo  $S$  em dois grupos menores  $S^1$  e  $S^2$ . Esse ponto está entre os pontos adjacentes mais distantes entre si ainda não separados, sendo eles  $s_i$  e  $s_{i+1}$ , com  $i$  definido por  $d_i = \max\{D\}$ ;
3. Avalie o valor da largura média de silhueta do agrupamento obtido;
4. Repetem-se,  $\lambda$  vezes, os passos 2 e 3 e retorna-se o agrupamento com a maior largura média de silhueta.

O ponto de secção, por ser a maior distância de um elemento a outro adjacente, será a distância de um conjunto a outro, o que maximiza as distâncias entre agrupamentos, ao passo que

exclui-se a possibilidade de ter o  $\max\{D\}$  como distância de elementos de um mesmo grupo, o que diminuiria o valor da Largura Média de Silhueta (LMS).

Sendo assim, num agrupamento unidimensional, existe a hipótese de que os  $k$  agrupamentos de  $S$  que geram o melhor valor de LMS sejam marcados pelas  $k$  maiores distâncias entre elementos adjacentes. É imprescindível que esse número seja maior ou igual ao número de conjuntos, visto que a qualidade de agrupamento não segue a mesma ordem das iterações. A Figura 2 ilustra os centros encontrados para a série temporal apresentada na Figura 1, lembrando que os centros são unidimensionais e os valores na abcissa foram colocados com finalidade estética.

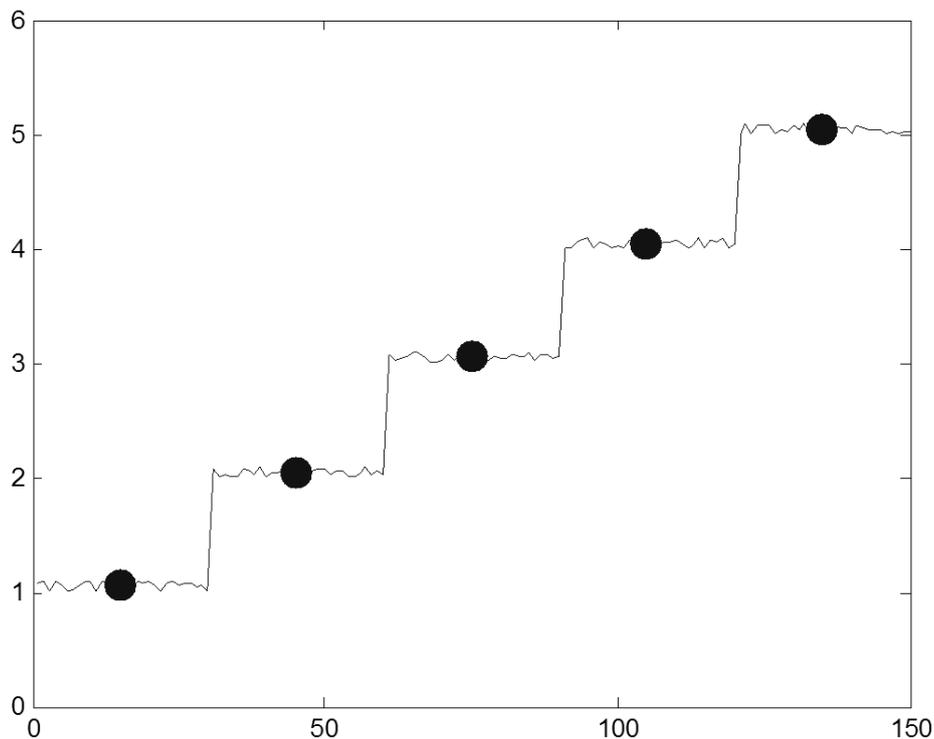


Figura 2 – Centros para Série Temporal da Figura 1

#### 4. Formulação do Algoritmo Metropolis-Hastings

O objetivo do algoritmo Metropolis-Hastings (Geman, 1997) é construir uma cadeia de Markov que tem uma determinada distribuição de equilíbrio  $\pi$ .

Defina uma cadeia de Markov como segue. Se  $X_{i-1} = x_{i-1}$ , então gere um valor candidato  $X^*$  a partir de uma distribuição com densidade  $f_{X^*|X}(y) = q(x_{i-1}, x^*)$ . A função  $q(\cdot)$  é conhecida como o núcleo de transição da cadeia de Markov. O valor candidato  $X^*$  é aceito ou rejeitado com probabilidade de aceitação:

$$\alpha(x, y) = \min\left(1, \frac{\pi(x^*)}{\pi(x_{i-1})} \frac{q(x_{i-1}, x^*)}{q(x^*, x_{i-1})}\right) \quad (4)$$

Se a candidata for aceita, faça  $X_i = Y$ , caso contrário, faça  $X_i = X_{i-1}$ . Assim, se a candidata for rejeitada, a cadeia de Markov repete a sequência. É possível mostrar que, em condições gerais, a sequência  $X_0, X_1, X_2, \dots$  é uma cadeia Markov com distribuição de equilíbrio  $\pi$ .

Em termos práticos, o algoritmo Metropolis-Hastings poder ser especificado pelos seguintes passos:

1. Escolha um valor inicial  $x_0$ , o número de simulações,  $R$ , e faça o contador de

- simulações  $r = 1$ ;
2. Gere um valor candidato  $y \sim q(x_i, \cdot)$ ;
  3. Calcule a probabilidade de aceitação como em (4) e gere  $u \sim U(0,1)$ ;
  4. Calcule o novo valor do estado atual:
 
$$x^{t+1} = \begin{cases} y, & \text{se } \alpha(x, y) > u \\ x^t, & \text{caso contrário} \end{cases}$$
  5. Se  $r < R$ , retorne ao passo 2. Caso contrário, pare.

Note que, como discutido em (D'Angelo, 2011c), a técnica de quantização gera uma série temporal transformada para o primeiro e o último grau de pertinência com a seguinte distribuição de probabilidade:

$$\begin{aligned} \mu_i(t) &\sim \text{Beta}(a, b), & \text{para } t = 1, \dots, m \\ \mu_i(t) &\sim \text{Beta}(c, d), & \text{para } t = m + 1, \dots, n \end{aligned}$$

Os parâmetros a serem estimados pelo algoritmo Metropolis-Hastings são  $a, b, c, d$  e o ponto de mudança  $m$ . Nesse tipo de algoritmo, normalmente, *prioris* pouco informativas são utilizadas, por exemplo:

$$\begin{aligned} a &\sim \text{Gamma}(0,1; 0,1) \\ b &\sim \text{Gamma}(0,1; 0,1) \\ c &\sim \text{Gamma}(0,1; 0,1) \\ d &\sim \text{Gamma}(0,1; 0,1) \\ m &\sim U[1, \dots, n], \text{ com } p(m) = 1/n \end{aligned}$$

A distribuição *Gamma* com parâmetros de forma e de escala iguais a 0,1 foi escolhida pelo fato de ser pouco informativa, com o objetivo de varrer todo o espaço paramétrico.

Considerando os graus de pertinência intermediários, tem-se as seguintes distribuições de probabilidade:

$$\begin{aligned} \mu_i(t) &\sim \text{Beta}(a, b), & \text{para } t = 1, \dots, m_{i-1} \\ \mu_i(t) &\sim \text{Beta}(c, d), & \text{para } t = m_{i-1} + 1, \dots, m_i \\ \mu_i(t) &\sim \text{Beta}(e, f), & \text{para } t = m_i + 1, \dots, n \end{aligned}$$

Os parâmetros a serem estimados pelo algoritmo Metropolis-Hastings são  $a, b, c, d, e, f$  e o segundo ponto de mudança,  $m_i$ , visto que o primeiro ponto de mudança,  $m_{i-1}$ , é identificado no passo anterior. Nesse tipo de algoritmo, normalmente são utilizadas *prioris* pouco informativas, por exemplo:

$$\begin{aligned} a &\sim \text{Gamma}(0,1; 0,1) \\ b &\sim \text{Gamma}(0,1; 0,1) \\ c &\sim \text{Gamma}(0,1; 0,1) \\ d &\sim \text{Gamma}(0,1; 0,1) \\ e &\sim \text{Gamma}(0,1; 0,1) \\ f &\sim \text{Gamma}(0,1; 0,1) \\ m_i &\sim U[m_{i-1}, \dots, n], \text{ com } p(m) = 1/(n - m_i) \end{aligned}$$

Os núcleos de transição da cadeia de Markov para o modelo com um único ponto de mudança e para o modelo com dois pontos de mudança são apresentados nos apêndices A e B, respectivamente.

## 5. Resultados

Para ilustrar a eficiência da metodologia proposta para detecção de pontos de mudança, foram feitos 100 testes para diferentes valores dos parâmetros:  $k$  (número de pontos de mudança),

$\mathcal{E}(t)$  (aqui foram utilizadas as distribuições uniforme, normal e *t*'students) e a porcentagem da amplitude do ruído em relação à mudança, e para fins de ilustração a mudança ocorre de uma unidade, e o primeiro ponto de operação é  $p_1 = 1$ . A metodologia proposta apresentou 100% de acerto na detecção de pontos de mudança, conforme Tabela 1. A Figura 3 mostra as probabilidades dos pontos de mudança para 1.000 simulações de Monte Carlo, para a série proposta em (3), e ilustrada na Figura 1. Nota-se que metodologia proposta foi capaz de identificar todos os pontos de mudança.

**Tabela 1 – Resultados de Detecção de Pontos de Mudança**

$p_1$	$k$	Amplitude do Ruído	% de acerto para $\mathcal{E}(t) \sim U(0,1)$	% de acerto para $\mathcal{E}(t) \sim N(0,1)$	% de acerto para $\mathcal{E}(t) \sim t(5)$
1	2	0,1	100	100	100
1	2	0,2	100	100	100
1	2	0,3	100	100	100
1	2	0,4	100	100	100
1	3	0,1	100	100	100
1	3	0,2	100	100	100
1	3	0,3	100	100	100
1	3	0,4	100	100	100
1	4	0,1	100	100	100
1	4	0,2	100	100	100
1	4	0,3	100	100	100
1	4	0,4	100	100	100

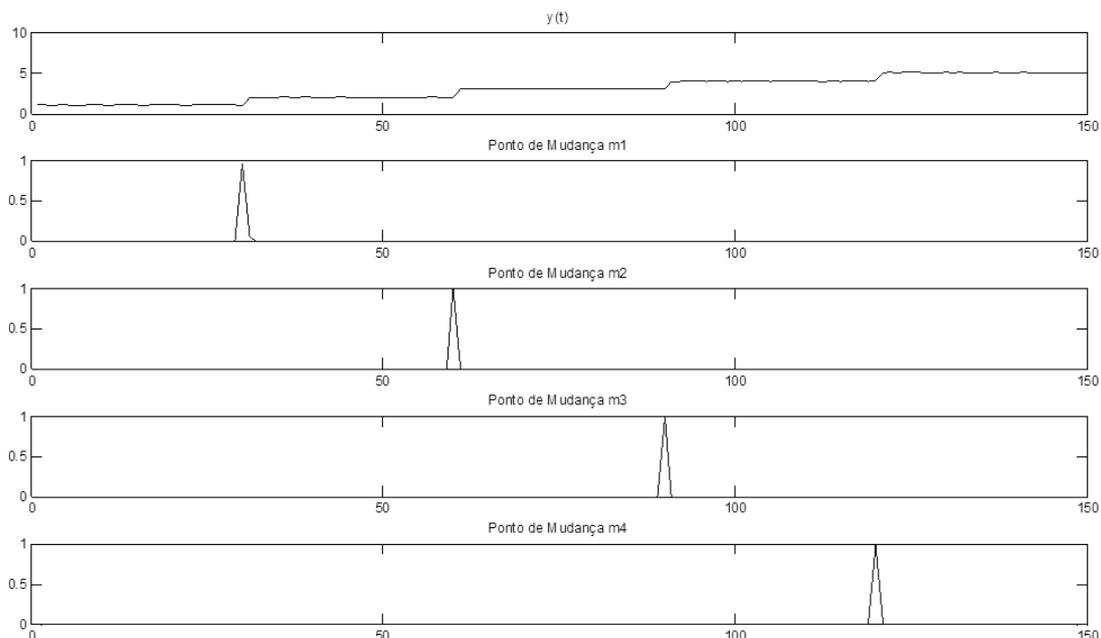


Figura 3 – Resultado da metodologia proposta para Série Temporal da Figura 1

## 6. Conclusões

Neste trabalho foi apresentada uma metodologia alternativa para detecção de pontos de mudança em séries temporais. Através de estudos de simulação mostrou-se a eficiência da

utilização de técnicas da teoria dos conjuntos *fuzzy* associadas a uma técnica *bayesiana* no problema de detecção de pontos de mudança em séries temporais. Os estudos experimentais sugerem que uma vantagem da metodologia adotada é o fato de não requerer nenhum conhecimento prévio sobre a estrutura dos dados, podendo ser utilizada em muitas aplicações práticas nas quais não haja conhecimento *a priori* de dados, tais como a função de distribuição de probabilidades das variáveis envolvidas.

## 7. Agradecimentos

Os autores agradecem o suporte financeiro da FAPEMIG e CNPq.

## Referências

- Barry, D. e Hartigan, J. A. (1993). A Bayesian Analysis for Change Point Problems. *Journal of the American Statistical Association*, 88 (421), 309-319.
- Beckage, B., Joseph, L., Belisle, P., Wolfson, D. B. e Platt, W. J. (2007). Bayesian change-point analyses in ecology. *New Phytologist*, 174, 456-467.
- Bezdek, J. C., *Pattern recognition with fuzzy objective function algorithms*. Plenum Press, 1981.
- Chinrungrueng, C. (1995). *Optimal adaptive k-means algorithm with dynamic adjustment of learning rate*, *IEEE Transactions on Neural Networks*, 6 (1), 157 – 169.
- D'Angelo M.S.F.V., Palhares R.M., Takahashi R.H.C. e Loschi R.H. (2011a). *Fuzzy/Bayesian change point detection approach to incipient fault detection*. *IET Control Theory and Applications*, 5(5), 539–551.
- D'Angelo M.S.F.V., Palhares R.M., Takahashi R.H.C., Loschi R.H., Baccarini L.M.R. e Caminhas W.M. (2011b). *Incipient fault detection in induction machine stator-winding using a fuzzy-Bayesian change point detection approach*. *Applied Soft Computing*, 11(1), 179–192.
- D'Angelo M.S.F.V., Palhares R.M., Takahashi R.H.C. e Loschi R.H. (2011c). *A fuzzy/Bayesian approach for the time series change point detection problem*. *Pesquisa Operacional*, 31, 217-234.
- Gamerman, D., *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman & Hall, 1997.
- Hadjiliadis, O. e Moustakides, V. (2006). *Optimal and Asymptotically Optimal CUSUM Rules for Change Point Detection in the Brownian Motion Model with Multiple Alternatives*. *Theory of Probability and its Applications*, 50 (1), 75–85.
- Hartigan, J. A. (1990). *Partition Models*. *Communication in Statistics-Theory and Methods*, 19 (8), 2745–2756.
- Heyer, L. J., Krugylak, S. e Yooseph, S., (1999). *Exploring expression data: Identification and analysis of coexpressed genes*. *Genome Research*, 9 (11), 1106–1115.
- Hinkey, D. V. (1971). *Inference About the Change Point from Cumulative Sum Test*. *Biometria*, 26, 279–284.
- Kaufman, L. e Rousseeuw, P. J., *Finding groups in data: An introduction to cluster analysis*. John Wiley & Sons, 1990.

Kohonen, T., *Self-organizing maps*. Springer, 2001.

Lee, S., Nishiyama, Y. e Yoshida, N. (2006). *Test for Parameter Change in Diffusion Processes by Cusum Statistics Based on One-step Estimators*. *Annals of the Institute of Statistical Mathematics* , 58 (2), 211-222.

Lee, S., Park, S., Maekawa, K. e Kawai, K. (2006). *Test for Parameter Change in ARIMA Models*. *Communications in Statistics: Simulation and Computation* , 35 (2), 429-439.

Loschi, R. H. e Cruz, F. R. B. (2005). *Bayesian Identification of Multiple Change Points in Poisson Data*. *Advances in Complex Systems* , 8, 465-482.

Loschi, R. H. e Cruz, F. R. B. (2005). *Extension to the product partition model: computing the probability of a change*. *Computational Statistics and Data Analysis* , 48 (2), 255-268.

Loschi, R. H., Gonçalves, F. B. e Cruz, F. B. R. (2005). *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*. *Pesquisa Operacional* , 25 (3), 459-463.

Moreira F.S., D'angelo M.F.S.V., Palhares R.M. e Caminhas W.M. (2010). *Incipient fault detection in induction machine stator-winding using a fuzzy-Bayesian two change points detection approach*. In *Proceedings of the 9th IEEE/IAS International Conference on Industry Applications (INDUSCON)*, 1–6.

Oh, K. J., Roh, T. H. e Moon, M. S. (2005). *Developing time-based clustering neural networks to use change-point detection: Application to financial time series*. *Asia-Pacific Journal Of Operational Research* , 22 (1), 51-70.

Rousseeuw, P. J. (1987). *Silhouettes: a graphical aid to the interpretation and validation of cluster analysis*. *Journal of Computational and Applied Mathematics*, 20, 53–65.

Zadeh, L. A. (1965). *Fuzzy Sets*. *Information and Control* , 8 (3), 338–353.

## A. Apêndice A

Este apêndice tem como finalidade mostra as probabilidades de aceitação para as distribuições *a posteriori* dos parâmetros  $a$ ,  $b$ ,  $c$ ,  $d$  e  $m$  descritas no algoritmo *Metropolis-Hastings* para o modelo com uma única mudança. As distribuições de referências usadas aqui são as próprias distribuições *a priori*, que neste trabalho foram escolhidas distribuições *Gamma(0.1,0.1)* por serem pouco informativas.

1. Para o parâmetro  $a$ :

$$\frac{\pi(a^*) q(a^*, a_{i-1})}{\pi(a_{i-1}) q(a_{i-1}, a^*)} = \frac{\left[ \frac{\Gamma(a^* + b^{i-1})}{\Gamma(a^*)} \right]^{m^{i-1}} \prod_{j=1}^{m^{i-1}} y_j^{a^*-1} \left\{ \left[ \frac{a^{i-1}}{a^*} \right]^{0.9} e^{-0.1(a^*-a^{i-1})} \right\}^2}{\left[ \frac{\Gamma(a^{i-1} + b^{i-1})}{\Gamma(a^{i-1})} \right]^{m^{i-1}} \prod_{j=1}^{m^{i-1}} y_j^{a^{i-1}-1}}$$

2. Para o parâmetro  $b$ :

$$\frac{\pi(b^*) q(b^*, b_{i-1})}{\pi(b_{i-1}) q(b_{i-1}, b^*)} = \frac{\left[ \frac{\Gamma(a^i + b^*)}{\Gamma(b^*)} \right]^{m^{i-1}} \prod_{j=1}^{m^{i-1}} (1 - y_j)^{b^* - 1} \left\{ \left[ \frac{b^{i-1}}{b^*} \right]^{0.9} e^{-0.1(b^* - b^{i-1})} \right\}^2}{\left[ \frac{\Gamma(a^{i-1} + b^{i-1})}{\Gamma(b^{i-1})} \right]^{m^{i-1}} \prod_{j=1}^{m^{i-1}} (1 - y_j)^{b^{i-1} - 1}}$$

3. Para o parâmetro  $c$ :

$$\frac{\pi(c^*) q(c^*, c_{i-1})}{\pi(c_{i-1}) q(c_{i-1}, c^*)} = \frac{\left[ \frac{\Gamma(c^* + d^{i-1})}{\Gamma(c^*)} \right]^{n - m^{i-1}} \prod_{j=m^{i-1}+1}^n y_j^{c^* - 1} \left\{ \left[ \frac{c^{i-1}}{c^*} \right]^{0.9} e^{-0.1(c^* - c^{i-1})} \right\}^2}{\left[ \frac{\Gamma(c^{i-1} + d^{i-1})}{\Gamma(c^{i-1})} \right]^{n - m^{i-1}} \prod_{j=m^{i-1}+1}^n y_j^{c^{i-1} - 1}}$$

4. Para o parâmetro  $d$ :

$$\frac{\pi(d^*) q(d^*, d_{i-1})}{\pi(d_{i-1}) q(d_{i-1}, d^*)} = \frac{\left[ \frac{\Gamma(c^i + d^*)}{\Gamma(d^*)} \right]^{n - m^{i-1}} \prod_{j=m^{i-1}+1}^n [(1 - y_j)^{d^* - 1}] \left\{ \left[ \frac{d^{i-1}}{d^*} \right]^{0.9} e^{-0.1(d^* - d^{i-1})} \right\}^2}{\left[ \frac{\Gamma(c^i + d^{i-1})}{\Gamma(d^{i-1})} \right]^{n - m^{i-1}} \prod_{j=m^{i-1}+1}^n [(1 - y_j)^{d^{i-1} - 1}]}$$

5. Para o parâmetro  $m$ :

$$\frac{\pi(m^*) q(m^*, m)}{\pi(m_{i-1}) q(m_{i-1}, m^*)} = \frac{\left[ \frac{\Gamma(a^i + b^i)}{\Gamma(a^i)\Gamma(b^i)} \right]^{m^*} \left[ \frac{\Gamma(c^i + d^i)}{\Gamma(c^i)\Gamma(d^i)} \right]^{n - m^*} \prod_{j=1}^{m^*} [y_j^{a^i - 1} (1 - y_j)^{b^i - 1}] \prod_{j=m^*+1}^n [y_j^{c^i - 1} (1 - y_j)^{d^i - 1}]}{\left[ \frac{\Gamma(a^i + b^i)}{\Gamma(a^i)\Gamma(b^i)} \right]^{m^{i-1}} \left[ \frac{\Gamma(c^i + d^i)}{\Gamma(c^i)\Gamma(d^i)} \right]^{n - m^{i-1}} \prod_{j=1}^{m^{i-1}} [y_j^{a^i - 1} (1 - y_j)^{b^i - 1}] \prod_{j=m^{i-1}+1}^n [y_j^{c^i - 1} (1 - y_j)^{d^i - 1}]}$$

## B. Apêndice B

Este apêndice tem como finalidade mostra as probabilidades de aceitação para as distribuições  $a$  *posteriori* dos parâmetros  $a, b, c, d, e, f, m_1$  e  $m_2$  descritas no algoritmo *Metropolis-Hastings* para o modelo com duas mudanças e lembrando que o ponto  $m_1$  será obtido sempre através da série anterior, não sendo necessária então que seja estimado. As distribuições de referências usadas aqui são as próprias distribuições *a priori*, que neste trabalho foram escolhidas distribuições  $Gamma(0.1, 0.1)$  por serem pouco informativas.

1. Para o parâmetro  $a$ :

$$\frac{\prod(a^*) q(a^*, a)}{\prod(a) q(a, a^*)} = \frac{\{0.1^{0.1} [\Gamma(0.1)]^{-1} a^{*0.1-1} e^{-0.1a^*}\}^2 \prod_{i=1}^{m_1} \frac{\Gamma(a^* + b)}{\Gamma(a^*)\Gamma(b)} y_i^{a^*-1}}{\{0.1^{0.1} [\Gamma(0.1)]^{-1} a^{0.1-1} e^{-0.1a}\}^2 \prod_{i=1}^{m_1} \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} y_i^{a-1}}$$

2. Para o parâmetro  $b$ :

$$\frac{\prod(b^*) q(b^*, b)}{\prod(b) q(b, b^*)} = \frac{\{0.1^{0.1} [\Gamma(0.1)]^{-1} b^{*0.1-1} e^{-0.1b^*}\}^2 \prod_{i=1}^{m_1} \frac{\Gamma(a + b^*)}{\Gamma(a)\Gamma(b^*)} (1 - y_i)^{b^*-1}}{\{0.1^{0.1} [\Gamma(0.1)]^{-1} b^{0.1-1} e^{-0.1b}\}^2 \prod_{i=1}^{m_1} \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} (1 - y_i)^{b-1}}$$

3. Para o parâmetro  $c$ :

$$\frac{\prod(c^*) q(c^*, c)}{\prod(c) q(c, c^*)} = \frac{\{0.1^{0.1} [\Gamma(0.1)]^{-1} c^{*0.1-1} e^{-0.1c^*}\}^2 \prod_{i=m_1+1}^{m_2} \frac{\Gamma(c^* + d)}{\Gamma(c^*)\Gamma(d)} y_i^{c^*-1}}{\{0.1^{0.1} [\Gamma(0.1)]^{-1} c^{0.1-1} e^{-0.1c}\}^2 \prod_{i=m_1+1}^{m_2} \frac{\Gamma(c + d)}{\Gamma(c)\Gamma(d)} y_i^{c-1}}$$

4. Para o parâmetro  $d$ :

$$\frac{\prod(d^*) q(d^*, d)}{\prod(d) q(d, d^*)} = \frac{\{0.1^{0.1} [\Gamma(0.1)]^{-1} d^{*0.1-1} e^{-0.1d^*}\}^2 \prod_{i=m_1+1}^{m_2} \frac{\Gamma(c + d^*)}{\Gamma(c)\Gamma(d^*)} (1 - y_i)^{d^*-1}}{\{0.1^{0.1} [\Gamma(0.1)]^{-1} d^{0.1-1} e^{-0.1d}\}^2 \prod_{i=m_1+1}^{m_2} \frac{\Gamma(c + d)}{\Gamma(c)\Gamma(d)} (1 - y_i)^{d-1}}$$

5. Para o parâmetro  $e$ :

$$\frac{\prod(e^*) q(e^*, e)}{\prod(e) q(e, e^*)} = \frac{\{0.1^{0.1} [\Gamma(0.1)]^{-1} e^{*0.1-1} e^{-0.1e^*}\}^2 \prod_{i=m_2+1}^n \frac{\Gamma(e^* + f)}{\Gamma(e^*)\Gamma(f)} y_i^{e^*-1}}{\{0.1^{0.1} [\Gamma(0.1)]^{-1} e^{0.1-1} e^{-0.1e}\}^2 \prod_{i=m_2+1}^n \frac{\Gamma(e + f)}{\Gamma(e)\Gamma(f)} y_i^{e-1}}$$

6. Para o parâmetro  $f$ :

$$\frac{\prod(f^*) q(f^*, f)}{\prod(f) q(f, f^*)} =$$

$$\frac{\{0.1^{0.1} [\Gamma(0.1)]^{-1} f^{*0.1-1} e^{-0.1f^*}\}^2 \prod_{i=m_2+1}^n \frac{\Gamma(e+f^*)}{\Gamma(e)\Gamma(f^*)} (1-y_i)^{f^*-1}}{\{0.1^{0.1} [\Gamma(0.1)]^{-1} f^{0.1-1} e^{-0.1f}\}^2 \prod_{i=m_2+1}^n \frac{\Gamma(e+f)}{\Gamma(e)\Gamma(f)} (1-y_i)^{f-1}}$$

7. Para o parâmetro  $m_2$ :

$$\frac{\prod(m_2^*) q(m_2^*, m_1)}{\prod(m_2) q(m_2, m_2^*)} = \frac{\left[\frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)}\right]^{m_2^*-m_1} \left[\frac{\Gamma(e+f)}{\Gamma(e)\Gamma(f)}\right]^{n-m_2^*}}{\left[\frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)}\right]^{m_2-m_1} \left[\frac{\Gamma(e+f)}{\Gamma(e)\Gamma(f)}\right]^{n-m_2}}$$

$$\times \frac{\prod_{i=m_1+1}^{m_2^*} y_i^{c-1} (1-y_i)^{d-1} \prod_{i=m_2^*+1}^n y_i^{e-1} (1-y_i)^{f-1}}{\prod_{i=m_1+1}^{m_2} y_i^{c-1} (1-y_i)^{d-1} \prod_{i=m_2+1}^n y_i^{e-1} (1-y_i)^{f-1}}$$