
MODELO PARTIÇÃO PRODUTO COM “*CLUSTERS*” CORRELACIONADOS

JACQUELINE ALVES FERREIRA

Orientadora: Prof^a. Dr^a. Rosângela Helena Loschi
Co-orientador: Prof. Dr. Marcelo Azevedo Costa

Area de Concentração: Estatística Bayesiana

Dissertação submetida como requerimento parcial para obtenção do
grau de Mestre em Estatística pela Universidade Federal de Minas Gerais

Belo Horizonte, fevereiro de 2010

©Copyright by
JACQUELINE ALVES FERREIRA
2010
Todos os direitos reservados
Typeset by L^AT_EX

AGRADECIMENTOS

Primeiramente agradeço a Deus e aos meus pais e familiares que sempre me apoiaram nessa caminhada.

Agradeço a professora Rosângela pela valiosa orientação.

Agradeço também a todos os meus amigos que sempre me incentivaram e proporcionaram muitos momentos de alegria e descontração.

Agradeço as secretárias desse programa de Pós-Graduação pelo carinho, compreensão e presteza dos seus afazeres.

Agradeço a CAPES e a FAPEMIG pelo apoio financeiro.

RESUMO

Identificar pontos de mudança é extremamente útil em diversas áreas de conhecimento, como por exemplo, economia, engenharia, hidrologia, medicina, meteorologia entre outras. Neste trabalho, esse problema será tratado do ponto de vista bayesiano, através de uma extensão do Modelo Partição Produto (MPP). Apesar do MPP ser amplamente utilizado na literatura para análise de dados sequenciais, em geral, é feita a suposição de que os parâmetros comuns que indexam as observações nos diferentes blocos ou (“*clusters*”) são independentes. Esta dissertação propõe um modelo tipo partição produto, o qual inova por assumir que estes parâmetros comuns são correlacionados. Esta correlação é introduzida no modelo através da distribuição *a priori* para tais parâmetros, baseada na idéia de processos de Markov. Especificamente, considerar-se-á a identificação de pontos de mudança na média da distribuição normal, cuja variância é considerada desconhecida, porém fixa ao longo do tempo. O modelo proposto é avaliado no contexto de dados sequenciais e os resultados obtidos são comparados com os resultados do MPP original Barry & Hartigan (1993), e a extensão do MPP proposta por Monteiro, Loschi & Assunção (2008).

ABSTRACT

The identification of change points is extremely useful in several areas of knowledge, such as economics, engineering, hydrology, medicine, meteorology among many others. In this work, this problem is addressed by using the Bayesian approach. An extension of the product partition model (PPM) is introduced in order to model such problems. Although the PPM has been widely used in the literature for the analysis of clustering in sequential data, it assumes that the common parameters which index the observations into the same block or cluster are independent ones. The model presented here is innovative since it considers that such common parameters are correlated. The correlation is introduced in the model throughout the prior of the common parameters which is built based on the Markov process. The proposed model is applied to identify multiple change points in normal means. It is evaluated using simulated data sets and the results are compared to those obtained using the original PPM by Barry & Hartigan (1993), and its extension proposed by Monteiro, Loschi & Assunção (2008).

1	Introdução	1
2	Modelo Partição Produto para Identificação de Pontos de Mudança	4
2.1	Modelo Partição Produto (MPP)	5
2.1.1	Métodos Computacionais para MPP	8
2.1.2	MPP para média normal	10
2.2	Modelo Partição Produto para a média normal com Correlação Intra Blocos	12
2.2.1	Métodos Computacionais para MPP com Correlação Intra Blocos	14
3	Modelo tipo Partição Produto com “<i>Clusters</i>” Correlacionados: Introduzindo a Correlação Entre Blocos	16
3.1	Especificação do modelo para médias normais	16
3.2	Métodos Computacionais para MPP com Correlação Entre Blocos	18
3.2.1	Algoritmo MCMC com Saltos Reversíveis (MCMCSR)	21
4	Resultados e Discussões	24
4.1	Cenário 1 - Sem mudança na média	26
4.2	Cenário 2 - Dois pontos de mudança na média	28
4.3	Cenário 3 - Dois pontos de mudança na média: maior variabilidade	32
4.4	Cenário 4 - Quatro pontos de mudança na média	36
4.5	Cenário 5 - Cinco pontos de mudança na média: mudanças alternadas	38
4.6	Cenário 6 - Quatro pontos de mudança na média: série decrescente e crescente	42
5	Aplicações	47
5.1	Série do Preço Médio do Tomate	47
5.2	Cromossomo 11 - Coriell	51
6	Conclusões	54

Referências Bibliográficas	56
A Gráficos das distribuições <i>a posteriori</i>	58
B Gráficos de Convergência do Procedimento MCMC Implementado	64

CAPÍTULO 1

Introdução

Identificar pontos de mudança é extremamente útil em diversas áreas de conhecimento, como por exemplo, economia, engenharia, hidrologia, medicina, meteorologia entre outras. Neste trabalho, esse problema será tratado do ponto de vista bayesiano, através de uma extensão do Modelo Partição Produto (MPP).

O modelo partição produto, em sua forma original, foi proposto por Hartigan (1990). Barry & Hartigan (1992) aplicaram o MPP na identificação de pontos de mudança em dados sequencialmente observados onde apenas blocos contíguos são considerados. Barry & Hartigan (1993) aplicam o MPP para identificação de múltiplos pontos de mudança na média da distribuição normal com variância constante e desconhecida, e sugerem uma forma de implementar computacionalmente, via amostrador de Gibbs, o MPP.

O modelo proposto por Hartigan(1990) é bastante explorado na literatura, como alguns exemplos de autores que exploram o MPP pode-se citar: Crowley (1997), Tarantola, Consoni, & Dallaportas(2008), Quintana & Iglesias(2003) entre outros. Essa popularidade do PPM é justificada pela flexibilidade proporcionada por esse modelo para analisar o problema de ponto de mudança. Em

função dessa flexibilidade, na literatura existem várias extensões desse modelo, como por exemplo, Loschi et al. (2003) e Loschi & Cruz (2005) que propõem um método para obter as distribuições *a posteriori* para a posição e número de pontos de mudança, além da probabilidade de cada instante ser um ponto de mudança. Outra extensão do MPP é proposta por Fearnhead (2006) que utiliza algoritmo filtro de partículas para encontrar as distribuições *a posteriori*.

Crowley (1997) aplica o MPP para identificar múltiplos pontos de mudança em blocos não contíguos para média Normal, porém assume que a variância é conhecida, além disso, nesse contexto, essa autora propõe uma forma de amostrar das distribuições *a posteriori* via Gibbs, evitando o uso do MCMC Saltos Reversíveis (MCMCSR). Recentemente, Hegarty & Barry (2008) utilizam o MPP no contexto espacial, esses autores desenvolvem um modelo para estimar o risco relativo em cada área de uma região e agrupar essas áreas de acordo com os seus respectivos riscos relativos.

No seu artigo original em 1990, Hartigan faz duas suposições sobre o MPP, a primeira é que as observações dentro de um mesmo “*cluster*” (ou bloco) são identicamente distribuídas, e a segunda assume que os parâmetros em diferentes “*clusters*” temporais (ou segmentos da série) dado a partição são independentes. Para ficar mais claro para o leitor, neste contexto, entende-se por partição os blocos contíguos formados pelos índices dos pontos de mudança, ou seja, se uma série possui 60 observações e dois pontos de mudança, e esses pontos são respectivamente, 21 e 41, então a partição ρ é dada por $\{[1 - 20], [21 - 40], [41 - 60]\}$.

Monteiro, Loschi & Assunção (2008) apresentam uma extensão do MPP, na qual a primeira suposição imposta por Hartigan (1990) é tratada de forma mais flexível, ou seja, esses autores assumem que os parâmetros em diferentes blocos são independentes, porém consideram que observações em um mesmo segmento têm sua distribuição indexada por diferentes parâmetros. Uma grande vantagem desse modelo proposto por esses autores é que sob essas suposições a dimensão do espaço paramétrico se mantém fixa.

A segunda suposição imposta pelo MPP de que os parâmetros em diferentes “*clusters*” temporais dado a partição são mutuamente independentes é muito forte, uma vez que esse modelo é

usualmente utilizado no contexto de séries temporais, onde as observações são fortemente correlacionadas. A proposta desse trabalho é introduzir um modelo tipo partição produto, porém levando em consideração a estrutura de correlação entre os segmentos da série. Essa estrutura de correlação é introduzida no modelo através da especificação da distribuição *a priori* para as médias de cada segmento. No modelo proposto a dimensão do espaço paramétrico não é fixa, assim como se tinha em Barry & Hartigan (1993), porém ao assumir uma estrutura de correlação entre os blocos, o esquema via amostrador de Gibbs para gerar da partição proposto por Barry & Hartigan (1993) teve seu uso inviabilizado. Assim, fez-se necessário propor uma forma de gerar, via MCMCSR, da distribuição *a posteriori* da partição.

Para avaliar o desempenho do modelo proposto foram gerados diversos cenários e os resultados obtidos foram comparados com o MPP original, ou seja, sem levar em consideração as suas extensões propostas na literatura, e o modelo proposto por Monteiro, Loschi & Assunção (2008). Vale ressaltar que os resultados para o MPP foram obtidos utilizando o pacote bcp implementado no *software* R por Erdman & Emerson (2007). Para o modelo proposto por Monteiro, Loschi & Assunção (2008) os resultados foram obtidos utilizando a interface gráfica do programa implementado em C++, gentilmente, cedido pelos autores. Os resultados para o modelo proposto foram obtidos a partir do programa implementado na linguagem matricial Ox, Doornik & Ooms (2005).

A presente dissertação está dividida em seis capítulos como descrito a seguir. No Capítulo 2 é apresentada uma breve revisão dos modelos propostos por Barry & Hartigan (1992, 1993) e Monteiro, Loschi & Assunção (2008); adicionalmente, são apresentadas algumas extensões do PPM propostas por Loschi, Iglesias & Cruz (2003) e Loschi & Cruz (2005). No Capítulo 3 é apresentado o modelo proposto nessa dissertação para identificação de pontos de mudança, no contexto de dados sequencialmente observados com média Normal e variância desconhecida e fixa. Resultados de simulação e aplicações encontram-se nos Capítulos 4 e 5. Finalmente, o Capítulo 6 contém as principais conclusões desse trabalho.

Modelo Partição Produto para Identificação de Pontos de Mudança

Neste capítulo é feita uma breve revisão da metodologia proposta por Barry & Hartingan (1992 e 1993) e Monteiro, Loschi & Assunção (2008) para identificação de pontos de mudança utilizando modelos partição produto. No parágrafo a seguir, para padronizar a notação, são definidos alguns termos que serão utilizados ao longo de todo o texto.

Para simplificar a apresentação Y será utilizado tanto para denotar as observações feitas quanto para denotar variável aleatória. Sejam Y_1, \dots, Y_n observações a serem feitas ao longo do tempo. Admita que esta série experimentará mudanças em seu comportamento impondo uma partição nessa série em blocos ou subsequências contíguas no tempo.

Para estabelecer notação, chame de ρ a partição definida por esses pontos de mudança no conjunto $I = \{0, 1, \dots, n\}$. Assuma que um valor particular $\{i_0, \dots, i_b\}$ de ρ é tal que $0 = i_0 < i_1 < \dots < i_b = n$ a qual divide a série em $B = b$ blocos (“clusters”) contíguos. Um bloco temporal é denotado por $[i_j i_{j+1}] = \{i_j + 1, \dots, i_{j+1}\}$, onde $i_j \in I = \{0, 1, \dots, n\}$ e $j \in 0, 1, \dots, b - 1$, ou seja, $i_{j+1} + 1$ é o instante de um ponto de mudança. A partição em I , induz uma partição na série Y_1, \dots, Y_n . Dessa forma, ao bloco $[i_j i_{j+1}]$ está associado um bloco de observações $\mathbf{Y}_{[i_j i_{j+1}]} = [Y_{i_j+1}, \dots, Y_{i_{j+1}}]$.

2.1 Modelo Partição Produto (MPP)

O modelo partição produto foi proposto por Hartigan (1990) em sua forma geral, a qual permite a identificação de agrupamentos não contíguos num conjunto de dados. Barry & Hartigan (1992) aplicaram o MPP na identificação de pontos de mudança em dados sequencialmente observados obtendo as formas fechadas das distribuições *a posteriori* para ρ e para os parâmetros que indexam as distribuições das observações Y 's. Mais tarde, Barry & Hartigan (1993) aplicam o MPP para identificação de múltiplos pontos de mudança na média de dados normalmente distribuídos com variância comum desconhecida. Uma das grandes contribuições desse trabalho é que, esses autores, introduzem uma forma de implementar computacionalmente, via amostrador de Gibbs, o MPP apesar da dimensão do espaço paramétrico não ser conhecido. Nesta seção, o modelo partição produto, usualmente utilizado para identificar pontos de mudança e amplamente discutido na literatura, é brevemente descrito.

Denote por $C_{i_j i_{j+1}}$ a coesão *a priori* associada ao bloco $[i_j i_{j+1}] = \{i_j + 1, \dots, i_{j+1}\}$, para $i_j \in I$. A coesão $C_{i_j i_{j+1}}$ é um número não negativo que representa a percepção que se tem sobre a semelhança ou associação das observações dentro de um mesmo bloco, isto é, o grau de similaridade que se julga existir entre as observações em um mesmo bloco, desta forma é uma avaliação subjetiva. No caso de partições em blocos contíguos, as coesões podem ser interpretadas como as probabilidades de transição a um passo na cadeia de Markov definida por pontos de mudança, ver detalhes em Barry & Hartigan, 1992 e Loschi (1998).

Particularmente, será apresentado o MPP considerando as coesões propostas por Yao (1984). Seja p , $0 \leq p \leq 1$, a probabilidade de que uma mudança ocorra em um instante qualquer. A coesão *a priori* para o bloco $[i_j i_{j+1}]$ proposta por Yao(1984) é dada por:

$$C_{i_j i_{j+1}} = \begin{cases} p(1-p)^{i_{j+1}-i_j-1}, & i_{j+1} < n; \\ (1-p)^{i_{j+1}-i_j-1}, & i_{j+1} = n, \end{cases}$$

para todo $i_j \in I$. Essas coesões *a priori* implicam que condicional ao fato de que uma mudança ocorreu no instante $i_j + 1$ pontos de mudança ocorridos anteriormente a esse não são informativos sobre a ocorrência de um pontos de mudança no instante i_{j+1} . Essas coesões também serão

utilizadas nos modelos que serão vistos nas Seções 2.1.2, 2.2 e no modelo introduzido no Capítulo 3.

Seja $\theta_1, \dots, \theta_n$, uma sequência de parâmetros desconhecidos, condicionalmente nos quais, as variáveis aleatórias, Y_1, \dots, Y_n , são independentes e têm densidades condicionais marginais dadas por $f(Y_1|\theta_1), \dots, f(Y_n|\theta_n)$, respectivamente. A distribuição *a priori* de $\theta_1, \dots, \theta_n$ é construída como segue.

Dada uma partição $\rho = \{i_0, \dots, i_b\}$, para $b \in \{1, 2, \dots, n\}$, tem-se que $\theta_i = \theta_{[i_j i_{j+1}]}$, para todo $i_j < i \leq i_{j+1}$, $j = 0, \dots, b-1$, e que $\theta_{[i_0 i_1]}, \dots, \theta_{[i_{b-1} i_b]}$ são independentes entre si e de p . Assuma que $\theta_{[i_j i_{j+1}]}$ tem distribuição *a priori* $\pi_{[i_j i_{j+1}]}(\theta)$, $\theta \in \Theta_{[i_j i_{j+1}]}$, em que $\Theta_{[i_j i_{j+1}]}$ é o espaço paramétrico correspondente ao parâmetro comum $\theta_{[i_j i_{j+1}]}$ que indexa a densidade condicional das observações em $Y_{[i_j i_{j+1}]}$, ou seja, $\theta_{[i_j i_{j+1}]} = \theta_{i_j+1} = \dots = \theta_{i_{j+1}}$. Assim, Barry & Hartingan (1992) definem o modelo partição produto para as coesões *a priori* de Yao como segue:

i) dado p , a distribuição *a priori* de ρ tem a seguinte distribuição produto:

$$\pi(\rho = \{i_0, \dots, i_b\}) = p^{b-1}(1-p)^{n-b}, \quad \forall b = 1, 2, \dots, n; \quad (2.1)$$

ii) dado $\rho = \{i_0, \dots, i_b\}$ e p , a sequência Y_1, \dots, Y_n é independente de p e sua densidade conjunta é dada por:

$$f(Y_1, \dots, Y_n | \rho = \{i_0, \dots, i_b\}, p) = \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(\mathbf{Y}_{[i_j i_{j+1}]}), \quad (2.2)$$

onde $f_{[i_j i_{j+1}]}(\mathbf{Y}_{[i_j i_{j+1}]}) = \int_{\Theta_{[i_j i_{j+1}]}} f_{[i_j i_{j+1}]}(\mathbf{Y}_{[i_j i_{j+1}]} | \theta) \pi_{[i_j i_{j+1}]}(\theta) d\theta$ é a densidade conjunta do vetor aleatório $\mathbf{Y}_{[i_j i_{j+1}]}$.

A partir das suposições feitas sobre os pontos de mudança segue quase diretamente que a distribuição *a priori* de ρ tem a forma produto em 2.1. Note que a partição $\rho = \{i_0, i_1, \dots, i_b\}$ é equivalente ao vetor $(U_{i_0} = 1, \dots, U_{i_1-1} = 1, U_{i_1} = 0, U_{i_1+1} = 1, \dots, U_{i_2} = 0, \dots, U_b = 1)$, onde a quantidade aleatória U_i é definida da seguinte forma:

$$U_i = \begin{cases} 1, & \text{se } \theta_i = \theta_{i+1}; \\ 0, & \text{se } \theta_i \neq \theta_{i+1}, \end{cases}$$

para $i = 1, \dots, n-1$. Assim se p denota a probabilidade de uma mudança em um instante qualquer

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

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

$$E(\theta_i | Y_1, \dots, Y_n) = \sum_{i_j=0}^{i-1} \sum_{i_{j+1}=i}^n r_{[i_j i_{j+1}]}^* E(\theta_i | Y_{[i_j i_{j+1}]}),$$

onde $r_{[i_j i_{j+1}]}^* = P([i_j i_{j+1}] \in \rho | Y_1, \dots, Y_n)$ denota a relevância *a posteriori* para o bloco $[i_j i_{j+1}]$.

Observação 1. O modelo original de Barry & Hartigan (1992 e 1993) foi posteriormente estendido por Loschi, Iglesias & Cruz (2003) e Loschi & Cruz (2005) para obter as distribuições *a posteriori* dos parâmetros p , B , e para determinar a probabilidade *a posteriori* de cada ponto ser um ponto de mudança. Essa extensão é mostrada nos dois parágrafos a seguir.

Assuma que p tem distribuição *a priori* $Beta(\alpha, \beta)$. Como consequência dessa consideração, segue que as distribuições *a posteriori* de ρ e B são dadas, respectivamente, por:

$$\pi(\rho = \{i_0, \dots, i_b\} | Y_1, \dots, Y_n) \propto \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + b - 1)\Gamma(n + \beta - b)}{\Gamma(n + \alpha + \beta - 1)} \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]}),$$

$$\pi(B = b | Y_1, \dots, Y_n) \propto \binom{n-1}{b-1} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + \beta - 1)\Gamma(n - \beta - b)}{\Gamma(n + \alpha + \beta - 1)} \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]}).$$

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

$$\pi(p | Y_1, \dots, Y_n) \propto \sum \prod_{j=0}^{b-1} f(Y_{[i_j i_{j+1}]}) p^{b+\alpha-2} (1-p)^{n-b+\beta-1},$$

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 a condição $0 = i_0 < i_1 < \dots < i_b = n$.

Na Seção 2.1.1, a seguir, os métodos computacionais para o MPP são apresentados baseados nessa extensão proposta por Loschi, Iglesias & Cruz (2003) e Loschi & Cruz (2005). Optou-se por apresentar tais métodos de forma mais geral, uma vez que o modelo proposto nessa dissertação e o

modelo proposto por Monteiro, Loschi & Assunção também permitem encontrar as distribuições *a posteriori* para p e B . Para o modelo originalmente proposto por Barry & Hartigan (1993) mais informações sobre o processo de estimação dos parâmetros serão dadas na Seção 2.1.2.

2.1.1 Métodos Computacionais para MPP

Para calcular a distribuição *a posteriori* de ρ, B e p , as estimativas produto de θ_k , alguns artigos na literatura sugerem utilizar o amostrador de Gibbs. Desta forma, é necessário encontrar a distribuição condicional completa de cada um dos parâmetros de interesse.

Seja $\mathbf{Y} = (Y_1, \dots, Y_n)$ e $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ e denote por $\boldsymbol{\theta}_{-l}$ o vetor $(\theta_1, \dots, \theta_{l-1}, \theta_{l+1}, \dots, \theta_n)$. Assuma que, dado ρ , $l \in [i_j + 1, i_j]$, ou seja, considere $\theta_l = \theta_{i_j+1} = \dots = \theta_{i_j}$, para $l = 1, \dots, n$ e $i_j \in I$. A distribuição condicional completa de p, ρ , e θ_l , para $l = 1, \dots, n$ são dadas, respectivamente, por:

$$\begin{aligned}\pi(p|\rho, \boldsymbol{\theta}, \mathbf{Y}) &\propto p^{b-1}(1-p)^{n-b}\pi(p); \\ \pi(\rho|p, \boldsymbol{\theta}, \mathbf{Y}) &\propto p^{b-1}(1-p)^{n-b} \prod_{j=1}^b f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]}) ; \\ \pi(\theta_l|\rho, p, \boldsymbol{\theta}_{-l}, \mathbf{Y}) &\propto f_{[i_j i_{j+1}]}(\theta_l|Y_{[i_j i_{j+1}]}),\end{aligned}$$

onde $\pi(p)$ denota a distribuição *a priori* de p .

Note que não é fácil amostrar diretamente da distribuição condicional completa de ρ , mas é possível definir uma quantidade aleatória auxiliar que torna mais fácil simular da distribuição condicional completa de ρ . Defina a quantidade aleatória auxiliar U_i , da seguinte forma:

$$U_i = \begin{cases} 1, & \text{se } \theta_i = \theta_{i+1}; \\ 0, & \text{se } \theta_i \neq \theta_{i+1}, \end{cases}$$

para $i = 1, \dots, n-1$. Note que a partição aleatória ρ é equivalente ao vetor aleatório $\mathbf{U} = (U_1, \dots, U_{n-1})$ dessas quantidades aleatórias, ou seja, como o zero indica um ponto de mudança, logo se o vetor aleatório \mathbf{U} é conhecido imediatamente a partição ρ também será conhecida.

Assumindo a equivalência entre ρ e \mathbf{U} , cada partição $(U_1^{(s)}, \dots, U_{n-1}^{(s)})$, $s \geq 1$, pode ser gerada usando o amostrador de Gibbs. Assim, a partir do valor inicial $(U_1^{(0)}, \dots, U_{n-1}^{(0)})$, no passo s , o r -ésimo elemento, $r = 1, \dots, n-1$, $U_r^{(s)}$ é gerado da distribuição condicional de U_r , dados $U_1^{(s)}, \dots, U_{r-1}^{(s)}, U_{r+1}^{(s-1)}, \dots, U_{n-1}^{(s-1)}, p^{(s-1)}, \boldsymbol{\theta}^{(s-1)}; \mathbf{Y}$.

Como cada coordenada U_r de \mathbf{U} só assume valores em $\{0, 1\}$ cada partição $(U_1^{(s)}, \dots, U_{n-1}^{(s)})$, $s \geq 1$, é gerada considerando a seguinte razão:

$$\begin{aligned} R_r &= \frac{P(U_r = 1 | U_1^{(s)}, \dots, U_{r-1}^{(s)}, U_{r+1}^{(s-1)}, \dots, U_{n-1}^{(s-1)}, p^{(s-1)}, \boldsymbol{\theta}^{(s-1)}; \mathbf{Y})}{P(U_r = 0 | U_1^{(s)}, \dots, U_{r-1}^{(s)}, U_{r+1}^{(s-1)}, \dots, U_{n-1}^{(s-1)}, p^{(s-1)}, \boldsymbol{\theta}^{(s-1)}; \mathbf{Y})} \\ &= \frac{f_{[xy]}(Y_{[xy]}) \int_0^1 p^{b-2} (1-p)^{n-b+1} d\pi(p)}{f_{[xr]}(Y_{[xr]}) f_{[ry]}(Y_{[ry]}) \int_0^1 p^{b-1} (1-p)^{n-b} d\pi(p)}, \end{aligned} \quad (2.3)$$

onde x é o último ponto de mudança antes de r e y é o próximo ponto de mudança depois de r .

O critério de escolha dos valores de $U_r^{(s)}$ é dada pela seguinte condição:

$$U_r^{(s)} = \begin{cases} 1, & \text{se } R_r \geq (1-u)/u; \\ 0, & \text{caso contrário,} \end{cases}$$

onde $r = 1, \dots, n-1$ e $u \sim \mathcal{U}(0, 1)$. O critério de escolha de U_r é baseado na forma de amostrar da distribuição de Bernoulli, veja Gamermam & Lopes (2006). Para maiores detalhes de como amostrar das distribuições *a posteriori* dos demais parâmetros veja Loschi & Cruz (2005).

Observação 2. Note que, neste contexto do MPP, a dimensão do espaço paramétrico muda dado a partição. A estratégia de amostragem sugerida por Barry & Hartigan (1993) apresenta uma solução para o problema de mudança de dimensão sem o uso do MCMCSR que só foi proposto posteriormente por Green (1995).

No que segue será abordado o problema de identificação de pontos de mudança na média de dados normalmente distribuídos com variância comum desconhecida. Brevemente, serão apresentados os modelos propostos por Barry & Hartigan (1993) e Monteiro, Loschi & Assunção (2008).

2.1.2 MPP para média normal

Nesta seção o modelo partição produto é aplicado ao caso normal, ou seja, Y_1, \dots, Y_n são variáveis aleatórias que, dados μ_1, \dots, μ_n , são independentes e tais que $Y_i \sim \mathcal{N}(\mu_i, \sigma_y^2)$.

Seja $\mu_{[i_j i_{j+1}]}$ o parâmetro comum que indexa a distribuição das observações no bloco $Y_{[i_j i_{j+1}]}$, assumamos uma distribuição *a priori* $\mathcal{N}(\mu_0, \sigma_0^2 / (i_{j+1} - i_j))$ para o parâmetro $\mu_{[i_j i_{j+1}]}$ do bloco $[i_j i_{j+1}]$, onde σ_0 é uma constante conhecida. Logo, a distribuição das observações $Y_{[i_j i_{j+1}]}$ para um dado bloco $[i_j i_{j+1}]$ é dada por:

$$f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]}) = \frac{1}{(2\pi\sigma_y^2)^{\frac{i_{j+1}-i_j}{2}}} \left(\frac{\sigma_y^2}{\sigma_0^2 + \sigma_y^2} \right)^{\frac{1}{2}} \exp(V_{[i_j i_{j+1}]}),$$

$$\text{onde } V_{[i_j i_{j+1}]} = -\frac{\sum_{l=i_j+1}^{i_{j+1}} (Y_l - \bar{Y}_{[i_j i_{j+1}]})^2}{2\sigma_y^2} - \frac{(i_{j+1}-i_j)(\bar{Y}_{[i_j i_{j+1}]} - \mu_0)^2}{2(\sigma_0^2 + \sigma_y^2)} \text{ e } \bar{Y}_{[i_j i_{j+1}]} = \sum_{l=i_j+1}^{i_{j+1}} \frac{Y_l}{(i_{j+1}-i_j)}.$$

Barry & Hartigan (1993) mostram que dado a partição ρ , a estimativa de μ_r para $r \in [i_j i_{j+1}] \in \rho$ é dada por:

$$\mu_r = (1 - w)\bar{Y}_{[i_j i_{j+1}]} + w\mu_0,$$

$$\text{onde } w = \sigma_y^2 / (\sigma_0^2 + \sigma_y^2).$$

Como na seção anterior assumamos que a distribuição *a priori* para ρ é a distribuição dada em (2.1). Para completar a especificação do modelo, assumamos as seguintes distribuições *a priori* utilizadas por Barry & Hartigan (1993)

$$\pi(\mu_0) = 1, \quad -\infty \leq \mu_0 \leq \infty,$$

$$\pi(\sigma_y^2) = \frac{1}{\sigma_y^2}, \quad 0 \leq \sigma_y^2 \leq \infty,$$

$$\pi(p) = \frac{1}{p_0}, \quad 0 \leq p \leq p_0,$$

$$\pi(w) = \frac{1}{w_0}, \quad 0 \leq w \leq w_0,$$

onde p_0 e w_0 são constantes conhecidas no intervalo $[0, 1]$. Barry & Hartigan (1993) assumem independência entre todos estes parâmetros.

Considerando essas especificações *a priori* Barry & Hartigan (1993) mostram que a distribuição *a posteriori* para ρ é dada por:

$$\pi(\rho|\mathbf{Y}) \propto \left[\int_0^{w_0} \frac{w^{(b-1)/2}}{(W+Sw)^{(n-1)/2}} dw \right] \left[\int_0^{p_0} p^{b-1}(1-p)^{n-b} dp \right],$$

note que essas integrais são betas incompletas, S é a soma de quadrados entre os blocos e W é a soma de quadrados dentro dos blocos. As estimativas para μ_r e σ_y^2 são dadas por:

$$E[\mu_r|\mathbf{Y}] = \sum_{\rho} [(1-w^*)\bar{Y}_{i_j i_{j+1}} + \bar{Y}w^*] \pi([\rho|\mathbf{Y}]), \quad (2.4)$$

$$E[\sigma_y^2|\rho, \mathbf{Y}] = \frac{1}{n-3} \frac{\int_0^{w_0} \frac{w^{(b-1)/2}}{[W+Sw]^{(n-3)/2}} dw}{\int_0^{w_0} \frac{w^{(b-1)/2}}{[W+Sw]^{(n-1)/2}} dw}, \quad (2.5)$$

onde $w^* = \frac{\int_0^{w_0} \frac{w^{(b+1)/2}}{[W+Sw]^{(n-1)/2}} dw}{\int_0^{w_0} \frac{w^{(b-1)/2}}{[W+Sw]^{(n-1)/2}} dw}$.

Métodos Computacionais para MPP: Barry & Hartigan (1993)

Nesta seção será descrito, brevemente, o procedimento computacional original de Barry & Hartigan (1993). Vale ressaltar, que todos os resultados obtidos nessa dissertação para o MPP foram obtidos baseados no modelo original de Barry e Hartigan (1993) utilizando o Pacote bcp implementado na linguagem computacional R por Erdman & Emerson (2007).

Para calcular a razão R_r dada em 2.3. Tem-se que:

$$\begin{aligned} R_r &= \frac{P(U_r = 1 | U_1^{(s)}, \dots, U_{r-1}^{(s)}, U_{r+1}^{(s-1)}, \dots, U_{n-1}^{(s-1)}, p^{(s-1)}, \boldsymbol{\theta}^{(s-1)}; \mathbf{Y})}{P(U_r = 0 | U_1^{(s)}, \dots, U_{r-1}^{(s)}, U_{r+1}^{(s-1)}, \dots, U_{n-1}^{(s-1)}, p^{(s-1)}, \boldsymbol{\theta}^{(s-1)}; \mathbf{Y})} \\ &= \frac{\left[\int_0^{w_0} \frac{w^{b/2}}{(W_1+S_1w)^{(n-1)/2}} dw \right] \left[\int_0^{p_0} p^b(1-p)^{n-b-1} dp \right]}{\left[\int_0^{w_0} \frac{w^{(b-1)/2}}{(W_0+S_0w)^{(n-1)/2}} dw \right] \left[\int_0^{p_0} p^{b-1}(1-p)^{n-b} dp \right]} \\ &= \left(\frac{W_0}{W_1} \right)^{\frac{n-b-2}{2}} \left(\frac{S_0}{S_1} \right)^{\frac{b+1}{2}} \sqrt{\frac{W_1}{S_1} \frac{\int_0^{\frac{S_1 w_0 / W_1}{1+S_1 w_0 / W_1}} p^{(b+2)/2} (1-p)^{(n-b-3)/2} dp \int_0^{w_0} p^b (1-p)^{n-b-1} dp}{\int_0^{\frac{S_0 w_0 / W_0}{1+S_0 w_0 / S_0}} p^{(b+1)/2} (1-p)^{(n-b-2)/2} dp \int_0^{w_0} p^{b-1} (1-p)^{n-b} dp}}, \end{aligned}$$

onde W_0, S_0, W_1 e S_1 são as somas de quadrados dentro e entre blocos obtidas quando $U_i = 0$ e $U_i = 1$, respectivamente.

Note que utilizando a idéia sugerida por Barry & Hartigan (1993), que foi discutida na Seção 2.1.1 para amostrar da partição ρ , o número de pontos de mudança em cada iteração é facilmente obtido contando o número de zeros em cada partição estimada, conseqüentemente, o número de blocos é o número de pontos de mudança acrescido de uma unidade. Também é possível obter a probabilidade de ocorrer uma mudança em cada ponto da série, utilizando-se do número de zeros em cada posição da partição estimada, ou seja, basta contar o número de zeros ao longo de todas as iterações e dividir o número de zeros obtidos em cada posição $i, i = \{1, \dots, n\}$ pelo tamanho da amostra final.

Finalmente, após ter atualizado a partição ρ , os μ'_i s são estimados utilizando (2.4), ao fim das iterações as estimativas de cada μ_i é dada pela média de todas as estimativas. A estimativa de σ_y^2 é calculada de forma equivalente.

A seguir será apresentado o modelo proposto por Monteiro, Loschi & Assunção (2008). Ao longo do presente texto o modelo proposto por esses autores será denominado de modelo partição produto com correlação intra blocos.

2.2 Modelo Partição Produto para a média normal com Correlação Intra Blocos

Nesta seção será exposto de forma, breve e sucinta o modelo proposto por Monteiro, Loschi & Assunção (2008). A principal diferença entre o modelo introduzido por Monteiro, Loschi & Assunção (2008) e a proposta feita por Barry & Hartigan (1993) é assumir que as médias dentro de cada bloco são diferentes e correlacionadas. A correlação é introduzida através da distribuição *a priori* para μ_1, \dots, μ_n . Ou seja, considerando uma distribuição *a priori* genérica π para os μ'_i s, assume-se que $\pi(\mu_{i+1}) = \pi(\mu_{i+1}|\mu_i)$, $i = 1, \dots, n - 1$. Sob estas suposições a dimensão do espaço paramétrico se mantém fixa, permitindo o uso do amostrador de Gibbs de forma mais natural.

Assuma que, dado μ_1, \dots, μ_n e τ_y , as variáveis aleatórias Y_1, \dots, Y_n são independentes e $Y_i|\mu_i, \tau_y \sim$

$\mathcal{N}(\mu_i, \frac{1}{\tau_y})$, isto é:

$$f(Y_i|\mu_i, \tau_y) = \frac{\tau_y^{(1/2)}}{\sqrt{2\pi}} \exp\left\{-\frac{\tau_y(Y_i - \mu_i)^2}{2}\right\}, \quad i = 1, \dots, n.$$

Assim como em Barry & Hartigan (1993), seja $\rho = \{i_0, \dots, i_b\}$ uma partição do conjunto de índice das observações $I = \{1, \dots, n\}$. Seja $p \in [0, 1]$ a probabilidade de que duas observações adjacentes estejam no mesmo “cluster” temporal. Conseqüentemente, introduzindo a estrutura de “cluster” e assumindo a partição aleatória ρ a verossimilhança é dada como segue:

$$f(\mathbf{Y}|\boldsymbol{\mu}, \tau_y, \rho = \{i_1, \dots, i_b\}) = \left(\frac{\tau_y}{2\pi}\right)^{\frac{n}{2}} \prod_{j=0}^{b-1} \prod_{k=i_j+1}^{i_{j+1}} \left[\exp\left\{-\frac{\tau_y(Y_k - \mu_k)^2}{2}\right\} \right].$$

Para controlar a similaridade dos parâmetros μ'_i s dentro do mesmo “cluster” assuma o hiperparâmetro $\tau_\mu \in \mathfrak{R}_+$. Nessa abordagem, a correlação temporal é introduzida no modelo através da distribuição *a priori* especificada para $\boldsymbol{\mu}$. Dados $\rho = \{i_0, \dots, i_b\}$ e o hiperparâmetro τ_μ , assuma que $\boldsymbol{\mu}$ é independente de p e que sua distribuição *a priori* é a seguinte distribuição de Gibbs, a qual contempla a correlação temporal dentro dos blocos:

$$\pi(\boldsymbol{\mu}|\rho = (i_0, \dots, i_b), \tau_\mu) = \tau_\mu^{\frac{n-b}{2}} \exp\left\{-\tau_\mu \sum_{i=1}^{n-1} \delta_i (\mu_i - \mu_{i+1})^2\right\},$$

onde δ_i é uma função indicadora:

$$\delta_i = \begin{cases} 0, & \text{se os parâmetros que indexam a distribuição das observações } i \text{ e } i+1 \\ & \text{não estão no mesmo “cluster” temporal;} \\ 1, & \text{se os parâmetros que indexam a distribuição das observações } i \text{ e } i+1 \\ & \text{estão no mesmo “cluster” temporal.} \end{cases}$$

Como distribuição *a priori* para ρ , assuma a distribuição *a priori* em (2.1). Note que tal distribuição é construída similarmente a partir da equivalência entre $\rho = \{i_0, \dots, i_b\}$ e $\{\delta_i, \dots, \delta_{i_1} = 1, \delta_{i_1+1} = 0, \dots, \delta_{i_2} = 1, \dots, \delta_{i_{c-1}} = 1, \dots, \delta_{n-1} = 0\}$. Como δ_i é uma variável aleatória com distribuição de Bernoulli com parâmetro p , a distribuição *a priori* de ρ é dada por:

$$\pi(\rho = \{i_1, \dots, i_b\}) = p^{b-1}(1-p)^{n-b},$$

onde $b = \sum_{i=1}^{n-1} \delta_i$ denota o número de “clusters” temporais no conjunto de dados. Completando a especificação do modelo, para τ_y e os hiperparâmetros τ_μ e p são assumidas as seguintes distribuições *a priori*:

$$\tau_y \sim \text{Gama}(t, u),$$

$$\tau_\mu \sim \text{Gama}(r, s),$$

$$p \sim \text{Beta}(a, b).$$

Logo, misturando as distribuições *a priori* e a verossimilhança via teorema de Bayes, têm-se que a distribuição *a posteriori* conjunta para os parâmetros deste modelo é:

$$\begin{aligned} \pi(\boldsymbol{\mu}, \tau_y^2, \rho, \tau_\mu^2, p | \mathbf{Y}) &\propto f(\mathbf{Y} | \boldsymbol{\mu}, \tau_y^2, \rho) \pi(\boldsymbol{\mu} | \rho, \tau_\mu^2) \pi(\rho | p) \pi(p) \pi(\tau_y^2) f(\tau_\mu^2) \\ &\propto \tau_y^{\frac{n+2t}{2}-1} \tau_\mu^{\frac{n-b+2r-2}{2}} p^{b+a-2} (1-p)^{n-b+c-1} \\ &\quad \exp \left\{ -\frac{\tau_y}{2} \left[\sum_{j=0}^{n-1} (Y_j - \mu_j)^2 + 2u \right] \right\} \\ &\quad \exp \left\{ -\frac{\tau_\mu}{2} \sum_{i=1}^{n-1} [\delta_i (\mu_i - \mu_{i+1})^2 + 2s] \right\}. \end{aligned}$$

Monteiro, Loschi & Assunção (2008) mostram que apesar da distribuição *a priori* escolhida para o vetor $\boldsymbol{\mu}$ ser imprópria a distribuição *a posteriori* é própria. A seguir são apresentados os métodos computacionais para o modelo proposto por esses autores.

2.2.1 Métodos Computacionais para MPP com Correlação Intra Blocos

Observe que nesse modelo têm-se que dada a partição ρ , os μ'_i s dentro dos blocos não são considerados iguais, mas sim correlacionados, o que facilita na hora de encontrar as condicionais completas e calcular a razão R_r utilizada para amostrar da distribuição *a posteriori* de ρ . Como, para esse modelo, todas as condicionais completas possuem forma fechada, todo o processo de simulação é feito via amostrador de Gibbs. A seguir são dadas as condicionais completas de todos os parâmetros a serem estimados:

$$\pi(\rho | \boldsymbol{\mu}, \tau_\mu, p, \mathbf{Y}) \propto \exp \left\{ -\tau_\mu \sum_{i=1}^{n-1} \delta_i (\mu_i - \mu_{i+1})^2 \right\} p^{b-1} (1-p)^{n-b}$$

$$\mu_i | \mu_{i-1}, \mu_{i+1}, \tau_y, \tau_\mu, \mathbf{Y}, \sim \text{Normal}(m_i, v_i), \quad i = 2, \dots, n-1,$$

$$\mu_n | \mu_{n-1}, \tau_y, \tau_\mu, \mathbf{Y}, \sim \text{Normal}(m_n, v_n),$$

$$\mu_1 | \mu_2, \tau_y, \tau_\mu, \mathbf{Y} \sim \text{Normal}(m_1, v_1),$$

$$p | \boldsymbol{\mu}, \tau_\mu, \mathbf{Y} \sim \text{Beta}(a^*, c^*),$$

$$\tau_\mu | \boldsymbol{\mu}, \rho, \mathbf{Y} \sim \text{Gama}(r^*, s^*),$$

$$\tau_y | \boldsymbol{\mu}, \mathbf{Y} \sim \text{Gama}(t^*, u^*),$$

onde

$$m_i = Y_i + 2\tau_y^{-1}\tau_\mu(\delta_{i-1}\mu_{i-1} + \delta_i\mu_{i+1})[1 + 2\tau_y^{-1}\tau_\mu(\delta_{i-1} + \delta_i)]^{-1},$$

$$v_i = \tau_y^{-1}[1 + 2\tau_y^{-1}\tau_\mu(\delta_{i-1} + \delta_i)]^{-1},$$

$$m_n = [Y_n + 2\tau_y^{-1}\tau_\mu(\delta_{n-1}\mu_{i-1})] + [1 + 2\delta_{n-1}\tau_y^{-1}\tau_\mu]^{-1},$$

$$v_n = \tau_y^{-1}[1 + 2\delta_{n-1}\tau_y^{-1}\tau_\mu]^{-1},$$

$$m_1 = Y_1 + 2\tau_y^{-1}\tau_\mu(\delta_1\mu_2)[1 + 2\delta_1\tau_y^{-1}\tau_\mu]^{-1},$$

$$v_1 = \tau_y^{-1}[1 + 2\delta_1\tau_y^{-1}\tau_\mu]^{-1},$$

$$a^* = \sum_{i=1}^{n-1} \delta_i + a - 1, \quad c^* = n - \sum_{i=1}^{n-1} \delta_i + c,$$

$$r^* = \frac{n+2r-b}{2}, \quad s^* = \sum_{i=1}^{n-1} \delta_i(\mu_i - \mu_{i-1})^2 + s,$$

$$t^* = \frac{[n+2t]}{2}, \quad u^* = \frac{[\sum_{i=1}^{n-1}(Y_i - \mu_i)^2]}{2+u}.$$

Para amostrar da distribuição *a posteriori* da partição ρ , utiliza-se a mesma idéia introduzida por Barry & Hartigan (1993) que foi descrita na Seção 2.1.1. Neste caso, no entanto, a razão R_r é dada a seguir:

$$\begin{aligned} R_i &= \frac{\pi(\boldsymbol{\mu} | \delta_i^j = 0, \boldsymbol{\delta}_{-i}^j, \tau_\mu) P(\delta_i^j = 0 | p)}{\pi(\boldsymbol{\mu} | \delta_i^j = 1, \boldsymbol{\delta}_{-i}^j, \tau_\mu) P(\delta_i^j = 1 | p)} \\ &= \sqrt{\tau_\mu} \exp\{-\tau_\mu(\mu_{i+1} - \mu_i)^2\} \frac{1-p}{p}, \end{aligned}$$

onde $i = 1, \dots, n-1$ e $\boldsymbol{\delta}_{-i}^j = \{\delta_1^j, \dots, \delta_{i-1}^j, \delta_{i+1}^{j-1}, \dots, \delta_{n-1}^{j-1}\}$.

No próximo capítulo será introduzida uma extensão do MPP usual, porém ao invés de assumir correlação dentro dos blocos como Monteiro, Loschi & Assunção (2008) propõe, o modelo que será apresentado contempla a correlação entre as médias dos blocos.

Modelo tipo Partição Produto com “*Clusters*” Correlacionados:
Introduzindo a Correlação Entre Blocos

Neste capítulo é apresentada uma nova proposta de modelagem para identificação de ponto de mudança, no contexto de dados sequenciais com distribuição Normal. Na metodologia apresentada aqui, a suposição feita por Barry e Hartigan (1993) de que as observações dentro de um mesmo bloco são identicamente distribuídas é mantida. Porém, ao invés de considerar que os blocos são independentes assume-se que os mesmos são correlacionados. Essa correlação é introduzida através da distribuição *a priori* para as médias comuns dos blocos, ou seja, assume-se que $\mu_{i_j+1} = \dots = \mu_{i_{j+1}} = \mu_{[i_j i_{j+1}]}$, $j = 0, \dots, b - 1$, e considera-se que a distribuição *a priori* para $\mu_{[i_j i_{j+1}]}$ depende de $\mu_{[i_{j-1} i_j]}$. A seguir será mostrado de forma mais detalhada a construção desse modelo.

3.1 Especificação do modelo para médias normais

Assuma que, dados μ_1, \dots, μ_n e σ_y^2 , $\mu_i \in \mathbb{R}$ e $\sigma_y^2 \in \mathbb{R}^+$ a sequência de variáveis aleatórias $Y_1 \dots Y_n$ são independentes e tal que $Y_i \mid \mu_i, \sigma_y^2 \sim \mathcal{N}(\mu_i, \sigma_y^2)$. Para introduzir a estrutura de agrupamento

assuma, também, que dada uma partição $\rho = \{i_0, \dots, i_b\}$ de $I = \{0, 1, \dots, n\}$, tem-se que $\mu_{i_{j+1}} = \dots = \mu_{i_{j+1}} = \mu_{[i_j i_{j+1}]}, j = 0, \dots, b-1$. Em outras palavras, dado $\rho = \{i_0, \dots, i_b\}$, $Y_1 \dots Y_n$ são tais que:

- i) Observações num mesmo bloco são condicionalmente independentes e identicamente distribuídas;
- ii) Observações em blocos distintos são condicionalmente independentes.

Seja $\boldsymbol{\mu} = \mu_1, \dots, \mu_n$, desta forma, a verossimilhança é dada por:

$$\begin{aligned} f(\mathbf{Y} \mid \sigma_y^2, \boldsymbol{\mu}, \rho = \{i_0, \dots, i_b\}) &= \prod_{j=0}^{b-1} f(Y_{i_{j+1}}, \dots, Y_{i_j} \mid \mu_{[i_j i_{j+1}]}) \\ &= \prod_{j=0}^{b-1} \prod_{k=i_j+1}^{i_{j+1}} f(Y_k \mid \mu_{[i_j i_{j+1}]}) \\ &= \prod_{j=0}^{b-1} \prod_{k=i_j+1}^{i_{j+1}} \left(\frac{1}{\sqrt{2\pi\sigma_y^2}} \exp \left\{ -\frac{1}{2\sigma_y^2} (Y_k - \mu_{[i_j i_{j+1}]})^2 \right\} \right). \end{aligned}$$

A correlação entre blocos será introduzida no modelo através da distribuição *a priori* conjunta para $\boldsymbol{\mu}$. A forma como essa distribuição *a priori* é construída, contempla uma correlação dos μ 's entre blocos. Para a construção da distribuição *a priori* assumamos que $\pi(\mu_{[i_j i_{j+1}]} \mid \mu_{[i_{j-1} i_j]}) \sim \mathcal{N}(\mu_{[i_{j-1} i_j]}, \sigma_\mu^2)$, $j = 0, \dots, b-1$, de tal forma que:

$$\begin{aligned} \pi(\mu_1, \dots, \mu_n \mid \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2) &= \pi(\mu_{[i_0 i_1]}, \dots, \mu_{[i_{b-1} i_b]} \mid \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2) \\ &= \prod_{j=0}^{b-1} \pi(\mu_{[i_j i_{j+1}]} \mid \mu_{[i_{j-1} i_j]}) \\ &\propto \left(\frac{1}{\sigma_\mu^2} \right)^{\frac{b}{2}} \prod_{j=0}^{b-1} \exp \left\{ -\frac{1}{2\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_{[i_{j-1} i_j]})^2 \right\}, \end{aligned}$$

onde $\mu_{[i_{-1} i_0]}$ é uma constante, ou seja, o valor inicial para a média da distribuição que descreve o comportamento do parâmetro $\mu_{i_0 i_1}$ que indexa a distribuição do primeiro bloco. Em outras palavras, $\mu_{[i_0 i_1]} \sim \mathcal{N}(\mu_{[i_{-1} i_0]}, \sigma_\mu^2)$.

Como distribuição *a priori* para a partição aleatória ρ será considerada a mesma distribuição produto proposta Barry e Hartigan (1992), diga-se,

$$\pi(\rho = \{i_0, \dots, i_b\} \mid p) = p^{b-1} (1-p)^{n-b}.$$

Para os hiperparâmetros assume-se as seguintes distribuições *a priori*:

$$p \sim \text{Beta}(a, c), \quad a > 0, c > 0,$$

$$\sigma_y^2 \sim \text{Gama} - \text{Invertida}(t, u), \quad t > 0, u > 0,$$

$$\sigma_\mu^2 \sim \text{Gama} - \text{Invertida}(r, s), \quad r > 0, s > 0.$$

Diz-se que uma variável aleatória $v \sim \text{Gama} - \text{Invertida}(a, b)$ se sua densidade é:

$$f(v) = \frac{a^b}{\Gamma(a)} (v^2)^{-(a+1)} \exp\left\{-\frac{b}{v^2}\right\}, \quad v \geq 0, a > 0, b > 0.$$

Como consequência destas suposições segue que, a distribuição *a posteriori* de $f(\boldsymbol{\mu}, \sigma_y^2, \rho, \sigma_\mu^2, p, | \mathbf{Y})$ é dada por:

$$\begin{aligned} \pi(\boldsymbol{\mu}, \sigma_y^2, \rho, \sigma_\mu^2, p | \mathbf{Y}) &\propto f(\mathbf{Y} | \boldsymbol{\mu}, \sigma_y^2, \rho) \pi(\boldsymbol{\mu} | \rho, \sigma_\mu^2) \pi(\rho | p) \pi(p) \pi(\sigma_y^2) \pi(\sigma_\mu^2) \\ &\propto \left(\frac{1}{\sigma_y^2}\right)^{\frac{n+2(t+1)}{2}} \left(\frac{1}{\sigma_\mu^2}\right)^{\frac{b+2(r+1)}{2}} (1-p)^{n+a-b-1} p^{b+c-2} \\ &\quad \exp\left\{-\frac{1}{2\sigma_y^2} \sum_{j=0}^{b-1} \left[\sum_{k=i_j+1}^{i_{j+1}} (Y_k - \mu_{[i_j i_{j+1}]})^2 + \frac{2u}{b} \right]\right\} \\ &\quad \exp\left\{-\frac{1}{2\sigma_\mu^2} \sum_{j=0}^{b-1} \left[(\mu_{[i_j i_{j+1}]} - \mu_{[i_{j-1} i_j]})^2 + \frac{2s}{b} \right]\right\}. \end{aligned}$$

3.2 Métodos Computacionais para MPP com Correlação Entre Blocos

Em inferência Bayesiana um caminho natural a seguir quando não é possível obter a forma fechada da distribuição *a posteriori* é estimá-la utilizando métodos MCMC. O amostrador de Gibbs, por ser mais simples, é a primeira opção quando a forma fechada da distribuição condicional completa é conhecida. Para maiores detalhes sobre esse algoritmo ver Gamerman & Lopes (2006).

Suponha que $k \in [i_l i_{l+1}]$, $l = 0, \dots, n$. Então, No modelo proposto para identificação de pontos de mudança, levando em consideração a correlação entre “clusters”, as distribuições condicionais

completas são dadas por:

$$\begin{aligned} \pi(\mu_k | \boldsymbol{\mu}_{-k}, p, \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2, \sigma_y^2, \mathbf{Y}) &= \pi(\mu_{[i_l i_{l+1}]} | \boldsymbol{\mu}_{-[i_l i_{l+1}]}, p, \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2, \sigma_y^2, \mathbf{Y}) \\ &\propto f(\mathbf{Y}_{[i_l+1 i_{l+1}]} | \rho) \pi(\mu_{[i_{l+1} i_{l+2}]} | \mu_{[i_l i_{l+1}]}) \pi(\mu_{[i_l i_{l+1}]} | \mu_{[i_{l-1} i_l]}) \\ &\propto \exp \left\{ -\frac{2\sigma_y^2 + \sigma_\mu^2 l^*}{2\sigma_y^2 \sigma_\mu^2} \left[\mu_{[i_l i_{l+1}]} - \left(\frac{\sigma_\mu^2 \bar{Y}_l l^* + \sigma_y^2 (\mu_{[i_{l+1} i_{l+2}]} + \mu_{[i_{l-1} i_l]})}{2\sigma_y^2 + \sigma_\mu^2 l^*} \right) \right]^2 \right\}, \end{aligned}$$

$$\begin{aligned} \pi(\sigma_y^2 | \boldsymbol{\mu}, p, \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2, \mathbf{Y}) &\propto f(\mathbf{Y} | \boldsymbol{\mu}, \sigma_y^2, \rho) \pi(\sigma_y^2) \\ &\propto \left(\frac{1}{\sigma_y^2} \right)^{\frac{n+2(t+1)}{2}} \exp \left\{ -\frac{1}{\sigma_y^2} \left[\frac{1}{2} \sum_{j=0}^{b-1} \left(\sum_{k=[i_j]+1}^{[i_{j+1}]} (Y_k - \mu_{[i_j i_{j+1}]})^2 + \frac{2u}{b} \right) \right] \right\}, \end{aligned}$$

$$\begin{aligned} \pi(\sigma_\mu^2 | \boldsymbol{\mu}, p, \rho = \{i_0, \dots, i_b\}, \sigma_y^2, \mathbf{Y}) &\propto \pi(\boldsymbol{\mu} | \rho) \pi(\sigma_\mu^2) \\ &\propto \left(\frac{1}{\sigma_\mu^2} \right)^{\frac{b+2(r+1)}{2}} \exp \left\{ -\frac{1}{\sigma_\mu^2} \left[\frac{1}{2} \sum_{j=0}^{b-1} \left((\mu_{[i_j i_{j+1}]} - \mu_{[i_{j-1} i_j]})^2 + \frac{2s}{b} \right) \right] \right\}, \end{aligned}$$

$$\begin{aligned} \pi(\rho = \{i_0, \dots, i_b\} | \boldsymbol{\mu}, p, \sigma_\mu^2, \sigma_y^2, \mathbf{Y}) &\propto f(\mathbf{Y} | \boldsymbol{\mu}, \sigma_y^2, \rho) \pi(\boldsymbol{\mu} | \rho, \sigma_\mu^2) \pi(\rho | p) \\ &\propto p^{n-b} (1-p)^{b-1} \left(\frac{1}{\sigma_\mu^2} \right)^{\frac{b}{2}} \exp \left\{ \sum_{j=0}^{b-1} -\frac{1}{2\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_{[i_{j-1} i_j]})^2 \right\} \\ &\quad \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{j=0}^{b-1} \sum_{k=[i_j]+1}^{[i_{j+1}]} (Y_k - \mu_{[i_j i_{j+1}]})^2 \right\}, \end{aligned}$$

$$\begin{aligned} \pi(p | \boldsymbol{\mu}, \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2, \sigma_y^2, \mathbf{Y}) &\propto \pi(\rho | p) \pi(p) \\ &\propto p^{b+c-2} (1-p)^{n+a-b-1}, \end{aligned}$$

onde, $\mathbf{Y}_{i_l+1 i_{l+1}} = (Y_{i_l+1}, \dots, Y_{i_{l+1}})$, $\bar{Y}_l = \sum_{k=[i_l]+1}^{[i_{l+1}]} \frac{Y_k}{l^*}$ e $l^* = (i_{l+1} - i_l + 1)$.

Note que as distribuições condicionais completas de $p, \sigma_y^2, \sigma_\mu^2$ são distribuições conhecidas a saber:

$$\sigma_y^2 \sim \text{Gama - Invertida}(t^*, u^*),$$

$$\sigma_\mu^2 \sim \text{Gama - Invertida}(r^*, s^*),$$

$$p \sim \text{Beta}(a^*, c^*),$$

onde $t^* = \frac{n+2t}{2}$, $u^* = \frac{1}{2} \sum_{j=0}^{b-1} \left[\sum_{k=[i_j]+1}^{[i_{j+1}]} \left(Y_k - \mu_{[i_j i_{j+1}]} \right)^2 + \frac{2u}{b} \right]$, $r^* = \frac{b+2r}{2}$ e $s^* = \frac{1}{2} \sum_{j=0}^{b-1} \left[\left(\mu_{[i_j i_{j+1}]} - \mu_{[i_{j-1} i_j]} \right)^2 + \frac{2s}{b} \right]$, $a^* = (n + a - b - 1)$ e $c^* = (b + c - 2)$.

A distribuição condicional completa de μ_k , $k = 1, \dots, n$, é uma distribuição $\mathcal{N}(m^*, v^*)$, onde $m^* = \frac{\sigma_\mu^2 \bar{Y} l^* + \sigma_y^2 (\mu_{[i_{l+1} i_{l+2}]} + \mu_{[i_{l-1} i_l]})}{2\sigma_y^2 + \sigma_\mu^2 l^*}$ e $v^* = \frac{\sigma_y^2 \sigma_\mu^2}{2\sigma_y^2 + \sigma_\mu^2 l^*}$. Porém, essa distribuição depende das médias dos blocos vizinhos e como o processo de simulação é recursivo esses blocos podem não existir no passo anterior. Para contornar esse problema, será obtida uma amostra da distribuição condicional completa *a posteriori* conjunta de $\boldsymbol{\mu}$ que é dada por:

$$\begin{aligned} \pi(\boldsymbol{\mu} | p, \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2, \sigma_y^2, \mathbf{Y}) &\propto f(\mathbf{Y} | \boldsymbol{\mu}, p, \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2, \sigma_y^2) \pi(\boldsymbol{\mu} | p, \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2, \sigma_y^2) \\ &\propto \prod_{j=0}^{b-1} \prod_{k=[i_j]+1}^{i_{j+1}} \exp \left\{ -\frac{1}{2\sigma_y^2} (Y_k - \mu_{[i_j i_{j+1}]})^2 \right\} \\ &\quad \exp \left\{ -\frac{1}{2\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_{[i_{j-1} i_j]})^2 \right\}. \end{aligned}$$

Nesse caso será utilizado o algoritmo de Metropolis-Hasting para gerar da distribuição condicional completa conjunta de $\boldsymbol{\mu}$, uma vez que tal distribuição não tem forma fechada.

Para gerar da distribuição *a posteriori* de ρ o uso da razão R_r proposta por Barry & Hartigan (1992) é mais onerosa computacionalmente pois, como as médias são dependentes, para retirar o efeito das mesmas como foi feito em Barry & Hartigan (1993), seria necessário integrar para todos os μ 's. Assim, dependendo do número de blocos ter-se-ia quantidades diferentes a serem integradas em cada iteração do algoritmo. Para melhor exemplificar considere a razão R_r para o caso onde têm-se dois e três blocos, respectivamente, e quer-se amostrar da distribuição de ρ ter-se-ia:

i) A sequência possui dois blocos:

$$R_r = \frac{\int f(\mathbf{Y} | \mu_{i_0 i_1}, \rho) f(\mu_{i_0 i_1} | \rho) d\mu_{i_0 i_1} \int_0^1 p^{b-2} (1-p)^{n-b+1} d\pi(p)}{\int \int f(\mathbf{Y} | \mu_{i_0 i_1}, \mu_{i_1 i_2}, \rho) f(\mu_{i_0 i_1}, \mu_{i_1 i_2} | \rho) \partial \mu_{i_0 i_1} \partial \mu_{i_1 i_2} \int_0^1 p^{b-1} (1-p)^{n-b} d\pi(p)};$$

ii) A sequência possui três blocos:

$$R_r = \frac{\int \int f(\mathbf{Y} | \mu_{i_0 i_1}, \mu_{i_1 i_2}, \rho) f(\mu_{i_0 i_1}, \mu_{i_1 i_2} | \rho) \partial \mu_{i_0 i_1} \partial \mu_{i_1 i_2} \int_0^1 p^{b-2} (1-p)^{n-b+1} d\pi(p)}{\int \int \int f(\mathbf{Y} | \mu_{i_0 i_1}, \mu_{i_1 i_2}, \mu_{i_2 i_3}, \rho) f(\mu_{i_0 i_1}, \mu_{i_1 i_2}, \mu_{i_2 i_3} | \rho) \partial \mu_{i_0 i_1} \partial \mu_{i_1 i_2} \partial \mu_{i_2 i_3} \int_0^1 p^{b-1} (1-p)^{n-b} d\pi(p)}.$$

Desta forma, a razão R_r é inviável de ser calculada analiticamente e, numericamente, o custo computacional seria muito alto. Assim, será necessário o uso da razão R_r junto com o algoritmo MCMC com Saltos Reversíveis (MCMCSR) para amostrar da distribuição de ρ . Esse algoritmo é descrito na próxima seção.

3.2.1 Algoritmo MCMC com Saltos Reversíveis (MCMCSR)

O algoritmo MCMC com saltos reversíveis é uma extensão do algoritmo de Metropolis-Hastings e foi proposto por Green (1995). A grande vantagem desse algoritmo é a possibilidade de movimentos entre espaços (modelos) de dimensões diferentes. No contexto de ponto de mudança esse problema pode ser interpretado como a necessidade, por exemplo, da cadeia mover-se de uma partição com b blocos para uma partição contendo $b + 1$ blocos e vice-versa.

Em cada iteração o algoritmo envolve a atualização dos parâmetros dado o modelo utilizando Gibbs e Metropolis-Hastings, por exemplo, e a atualização da dimensão usando o seguinte procedimento.

Suponha que o estado atual é $(k, \boldsymbol{\theta})$, i.e. estamos no modelo k com parâmetros $\boldsymbol{\theta}$ e um novo modelo k' com parâmetros $\boldsymbol{\theta}'$ é proposto com probabilidade $r_{k,k'}$. Em geral isto significa incluir ou retirar parâmetros do modelo atual. Vamos assumir inicialmente que o modelo proposto tem dimensão maior, i.e. $n_{k'} > n_k$ e que $\boldsymbol{\theta}' = g(\boldsymbol{\theta}, \mathbf{u})$ para uma função determinística g e um vetor aleatório $\mathbf{u} \sim q(\mathbf{u})$ com dimensão $n_{k'} - n_k$ (Ehlers, 2007). Então o seguinte algoritmo é utilizado,

1. proponha $(k, \boldsymbol{\theta}) \rightarrow (k', \boldsymbol{\theta}')$ com probabilidade $r_{k,k'}$;
2. gere $\mathbf{u} \sim q(\mathbf{u})$ com dimensão $n_{k'} - n_k$;
3. faça $\boldsymbol{\theta}' = g(\boldsymbol{\theta}, \mathbf{u})$;
4. aceite $(k', \boldsymbol{\theta}')$ com probabilidade $\min(1, A)$ sendo

$$A_{k,k'} = \frac{\pi(k', \boldsymbol{\theta}')}{\pi(k, \boldsymbol{\theta})} \times \frac{r_{k',k}}{r_{k,k'} q(\mathbf{u})} \left| \frac{\partial g(\boldsymbol{\theta}, \mathbf{u})}{\partial(\boldsymbol{\theta}, \mathbf{u})} \right|.$$

MCMCSR para o Modelo com Correlação Entre Blocos

Como o MCMCSR permite esse movimento entre os espaços, nesta seção será mostrado como a proposta de Barry & Hartingan (1993) pode ser utilizada para gerar da partição ρ , porém não é necessário integrar em relação a todos os parâmetros para retirar a influência da mudança de dimensão. Note que, no contexto do MPP, a probabilidade do modelo k é equivalente a probabilidade da partição ρ , assim $A_{k,k'} = A_{\rho,\rho'}$.

Para facilitar o cálculo da razão $A_{\rho,\rho'}$ os saltos só ocorrem com diferença de uma unidade na dimensão, ou seja, a cada iteração “nasce” um novo ponto de mudança ou “morre” um ponto de mudança. Dessa forma, os movimentos entre modelos são vistos como: (1) passo de nascimento de ponto de mudança, neste caso propõem-se saltar de uma partição ρ com b blocos para uma partição ρ' contendo $b + 1$ blocos, com probabilidade $r_{\rho,\rho'}$; (2) passo de morte de um ponto de mudança, em que deseja-se saltar de uma partição ρ' com $b + 1$ blocos para uma partição ρ com b blocos. Logo, a partição $(U_1^s, \dots, U_{n-1}^s)$, $s \geq 1$, pode ser gerada usando o MCMCSR da seguinte forma:

1. Gere $v \sim \mathcal{U}(0, 1)$, se $v > 0,5$, então o nascimento de um novo ponto de mudança é proposto. Se $v < 0,5$, então propõem-se retirar um ponto de mudança;
2. (i) Se $v > 0,5$ escolha um dos $n - b$ pontos restantes para ser um candidato a novo ponto de mudança. Calcule a Razão $A_{\rho,\rho'}$ e se $\min(1, A) > w$, então tem-se um novo ponto de mudança. Se $\min(1, A) < w$ a partição continua a mesma.
(ii) Agora, se $v < 0,5$, então propõem-se retirar um ponto de mudança dos $b + 1$ pontos de mudança da partição ρ' e calcula-se a razão A^{-1} . Se o $\min(1, A^{-1}) > w$, então retira-se o ponto de mudança, caso contrário, o ponto de mudança permanece;
3. Depois de atualizada a partição atualiza-se os demais parâmetros;
4. Volte ao passo 1 e execute o algoritmo até atingir a convergência da cadeia.

Assim, a partir do valor inicial $(U_1^{(0)}, \dots, U_{n-1}^{(0)})$, no passo s a partição $(U_1^{(s)}, \dots, U_{n-1}^{(s)})$ é gerada da seguinte distribuição condicional:

$$U_r | U_1^{(s-1)}, \dots, U_{r-1}^{(s-1)}, U_{r+1}^{(s-1)}, \dots, U_{n-1}^{(s-1)}, p^{(s-1)}, \boldsymbol{\theta}^{(s-1)}; \mathbf{Y},$$

logo, cada partição $(U_1^{(s)}, \dots, U_{n-1}^{(s)})$, $s \geq 1$, é gerada considerando a seguinte razão:

$$A_{\rho, \rho'} = \frac{\pi(\rho', \boldsymbol{\theta}')}{\pi(\rho, \boldsymbol{\theta})} \times \frac{r_{\rho', \rho}}{r_{\rho, \rho'} q(u)} \left| \frac{\partial g(\boldsymbol{\theta}, u)}{\partial(\boldsymbol{\theta}, u)} \right|,$$

onde, neste contexto do modelo proposto para identificação de pontos de mudança, $\left| \frac{\partial g(\boldsymbol{\theta}, u)}{\partial(\boldsymbol{\theta}, u)} \right| = 1$, $q(u) \sim t_{(g,l)}(\mu, \sigma^2)$ e $r_{\rho', \rho} = r_{\rho, \rho'} = 0.5$. Note que:

$$\pi(\rho, \boldsymbol{\theta}) \propto L_\rho(\mathbf{Y} | \boldsymbol{\theta}) \pi_\rho(\boldsymbol{\theta}) p(\rho).$$

Neste caso, $L_\rho(\mathbf{Y}, \boldsymbol{\theta})$ é a função de verossimilhança sob o modelo ρ dado os parâmetros do modelo ρ , $\pi(\boldsymbol{\theta})$ é a distribuição *a priori* dos parâmetros do modelo ρ e $p(\rho)$ é a probabilidade *a priori* do modelo ρ . Desta forma, fazendo $\boldsymbol{\theta} = (p, \mu, \sigma_\mu^2, \sigma_y^2)$ têm-se que:

$$\pi(\rho, p, \boldsymbol{\mu}, \sigma_\mu^2, \sigma_y^2) \propto L_\rho(\mathbf{Y} | \boldsymbol{\mu}, \sigma_\mu^2, \sigma_y^2) \pi_\rho(\boldsymbol{\mu} | \sigma_\mu^2) \pi_\rho(\sigma_\mu^2) \pi_\rho(\sigma_y^2) \int \pi_\rho(\rho | p) \pi_\rho(p) dp.$$

Assim, se $v > 0,5$, ou seja, está sendo proposto um novo ponto de mudança, o critério de escolha dos valores de $U_r^{(s)}$ é dada pela seguinte condição:

$$U_r^{(s)} = \begin{cases} 1, & \text{se } \min(1, A) > w; \\ 0, & \text{caso contrário,} \end{cases}$$

Conseqüentemente, se $v < 0,5$, ou seja, está sendo proposto retirar um ponto de mudança, $U_r^{(s)}$ é dada pela seguinte condição:

$$U_r^{(s)} = \begin{cases} 0, & \text{se } \min(1, A^{-1}) > w; \\ 1, & \text{caso contrário,} \end{cases}$$

onde $r = 1, \dots, n$, $v \sim \mathcal{U}(0, 1)$ e $w \sim \mathcal{U}(0, 1)$.

O capítulo a seguir reporta os resultados obtidos pelos três modelos discutidos nessa dissertação. Os resultados referem-se a diferentes cenários gerados. Esses cenários levam em consideração diversas características de uma série temporal como, por exemplo, se a série é crescente ou decrescente e se as médias dentro dos blocos são identicamente distribuídas.

Resultados e Discussões

Neste capítulo serão apresentados vários estudos de simulação para comparar os modelos introduzidos por Barry & Hartigan (1993) e Monteiro, Loschi & Assunção (2008) com o modelo proposto nessa dissertação. Os cenários escolhidos favorecem tanto o modelo com a correlação entre blocos, quanto os outros dois modelos citados.

As séries foram simuladas utilizando o gerador de números aleatórios de George Marsaglia implementado na linguagem de programação Ox. Para identificação dos pontos de mudança através do MPP proposto por Barry & Hartigan (1993) foi utilizado o pacote `bcp` implementado no R por Erdman & Emerson (2007). Vale ressaltar, novamente, que nessa abordagem original do MPP utilizada nesse trabalho é possível obter somente as estimativas pontuais para μ e σ_y , além disso, também não é possível obter a distribuição *a posteriori* para p . A identificação de pontos de mudança via MPP com correlação intra blocos proposta por Monteiro, Loschi & Assunção (2008) foi feita utilizando o programa em C++ fornecido pelos autores. E por fim, o MPP com correlação entre blocos foi implementado na linguagem matricial Ox. Os gráficos necessários para melhor visualização dos resultados foram feitos no *software* livre R 2.9.1 R, Development Core Team (2009).

Para implementação dos métodos MCMC utilizados foram geradas cadeias com 100 mil iterações, com um período de *burn-in* de 50 mil iterações. Os gráficos indicando a convergência encontram-se no Apêndice B. Para evitar a correlação entre as observações foi utilizado *lag* 50, resultando em uma amostra final de tamanho 1.000.

Todos os cenários foram analisados assumindo-se as mesmas distribuições *a priori*. As Tabelas 4.1, 4.2, 4.3 a seguir mostram as especificações *a priori* para cada modelo.

Tabela 4.1: Especificações *a priori* para MPP com correlação entre blocos

Parâmetros	Distribuições <i>a priori</i>	Média	Des. Padrão
σ_μ^2	<i>Gama – Invertida</i> (3; 30)	15,000	15,000
σ_y^2	<i>Gama – Invertida</i> (0,001; 0,001)	-	-
p	<i>Beta</i> (1, 1; 1, 1)	0,500	0,279

Tabela 4.2: Especificações *a priori* para MPP com correlação intra blocos

Parâmetros	Distribuições <i>a priori</i>	Média	Des. Padrão
τ_μ	<i>Gama</i> (800; 100)	8,000	0,283
τ_y	<i>Gama</i> (1; 1)	1,000	1,000
p	<i>Beta</i> (5; 1)	0,833	0,141

Tabela 4.3: Distribuições *a priori* para MPP

Parâmetros	Distribuições <i>a priori</i>
μ_0	1, $-\infty \leq \mu_0 \leq \infty$
σ_y^2	$1/\sigma_y^2$, $0 \leq \sigma_y^2 \leq \infty$
p	$1/p_0$, $0 \leq p \leq p_0$
w	$1/w_0$, $0 \leq w \leq w_0$

Na Tabela 4.3 assumiu-se que $w_0 = 0.2$ e $p_0 = 0.2$, esses valores são os mesmos sugeridos no artigo de Barry & Hartingan (1993). Na Tabela 4.2 nota-se que a distribuição *a priori* para τ_μ é informativa, uma vez que a variância é muito pequena. Essa escolha é feita, pois no MPP com correlação intra blocos assume-se que os μ 's são fortemente correlacionados. Dessa forma, a escolha da distribuição *a priori* com uma variância pequena faz com que o MPP com correlação intra blocos não se afaste demasiadamente da suposição, feita pelo MPP original, de que as μ 's dentro de um mesmo bloco são identicamente distribuídas. Para o MPP com correlação entre blocos a suposição de correlação entre as μ 's é incorporada ao modelo de uma forma menos concentrada, como pode-se observar na

Tabela 4.1. Nas seções a seguir serão descritos cada um dos cenários considerados nesse trabalho, além da discussão dos resultados obtidos.

4.1 Cenário 1 - Sem mudança na média

Nesse cenário os dados foram gerados sem ponto de mudança, ou seja, as observações $y_i | \mu_i, \sigma_y^2 \sim \mathcal{N}(5; 1)$, $1 \leq i \leq 60$. A Figura 4.1 apresenta as estimativas das médias *a posteriori* para os três modelos considerados nessa dissertação.

Observa-se na Figura 4.1 que os modelos MPP e MPP com correlação entre blocos apresentam estimativas idênticas para as médias, como era esperado, pois a diferença entre esses modelos é a correlação entre blocos, logo se há apenas um bloco os modelos são equivalentes. Além disso, nota-se que o modelo com correlação intra blocos apresenta estimativas menos suaves para a média que os demais modelos, indicando que a média ao longo do tempo não é constante. Na Figura 4.2 que mostra a probabilidade de mudança em cada instante e o número de blocos *a posteriori* constata-se que todos os modelos indicam corretamente que não há pontos de mudança na série.

A Tabela 4.4 indica as partições mais prováveis e suas respectivas probabilidades. Nessa tabela nota-se que os modelos MPP e MPP com correlação entre blocos identificam a partição corretamente com probabilidade 1. Já o MPP com correlação intra blocos identifica a partição correta com probabilidade 0,90.

A Tabela 4.5 mostra o resumo das distribuições *a posteriori* dos parâmetros σ_μ^2 , σ_y^2 , p e B para os modelos MPP com correlação entre e intra blocos. Para o MPP, como foi utilizado o modelo original implementado no R, só foi possível obter a distribuição *a posteriori* para o número de blocos e a probabilidade de mudança em cada instante do tempo. Pois, Erdman & Emerson (2007) ao implementar o MPP no R utilizam o artifício mencionado na Seção 2.1.2 de contar os zeros da partição. Observa-se, nessa tabela, que o MPP com correlação entre blocos apresenta estimativas pontuais mais próximas dos verdadeiros parâmetros. Os gráficos das distribuições *a posteriori* dos

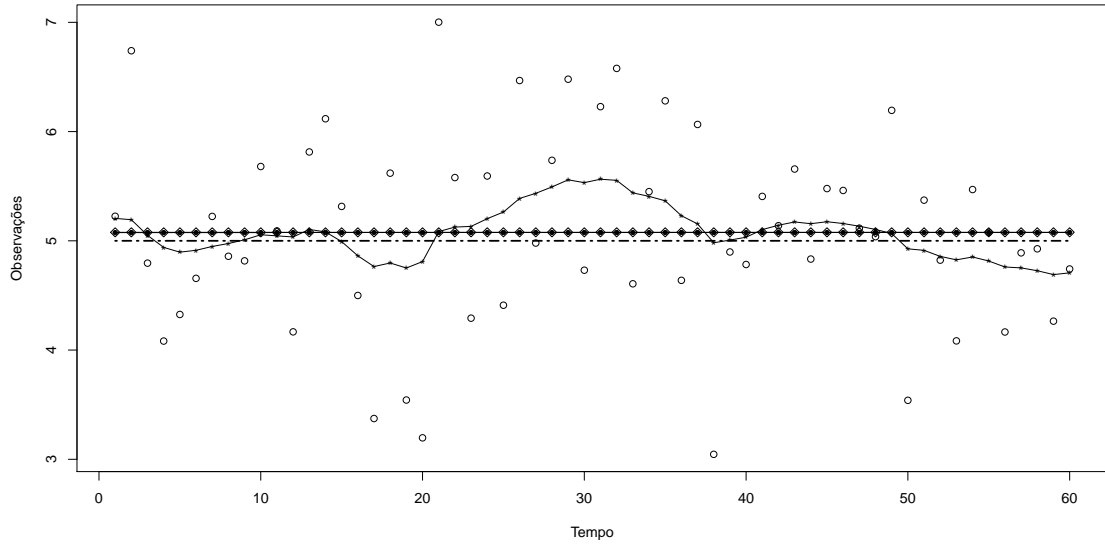


Figura 4.1: Estimativas das médias *a posteriori*. O símbolo \bullet indica o MPP com correlação entre blocos, \diamond são as estimativas via PPM, $*$ são estimativas via MPP com correlação intra blocos, a linha tracejada indica o parâmetro verdadeiro e o símbolo \circ indica os valores simulados da distribuição normal.

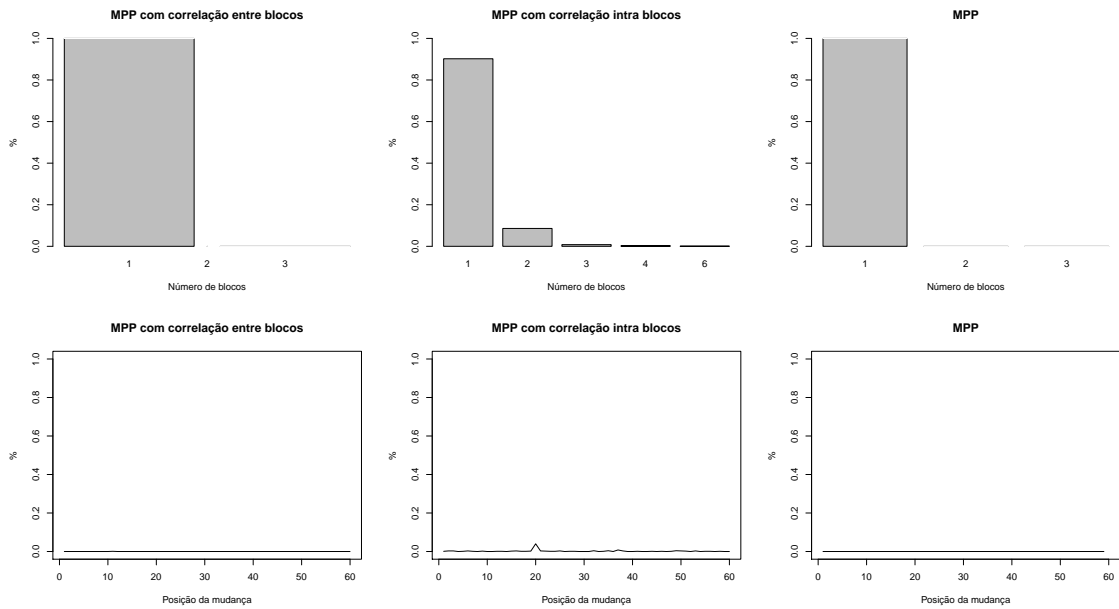


Figura 4.2: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto.

parâmetros σ_μ^2 , σ_y^2 e p obtidos através dos modelos MPP com correlação intra e entre blocos para todos os cenários considerados nessa dissertação encontram-se no Apêndice A.

Observação 3. Note que o modelo com correlação intra blocos elicia as distribuições *a priori* para a precisão, ou seja, para τ_μ e τ_y . Desta forma, tem-se que $\sigma_y^2 = \tau_y^{-1}$ e $\sigma_\mu^2 = \tau_\mu^{-1}$.

Tabela 4.4: Partições mais prováveis *a posteriori* - Cenário 1.

Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 60]\}) = 1$
MPP com correlação intra blocos	$P(\rho = \{[0 - 60]\}) = 0.902$ $P(\rho = \{[0 - 20], [21 - 60]\}) = 0.033$
MPP	$P(\rho = \{[0 - 60]\}) = 1$

Tabela 4.5: Resumo das distribuições *a posteriori* - Cenário 1

Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	20,850	19,207	[4,331; 50,533]
	σ_y^2	0,786	0,150	[0,522; 1,087]
	p	0,998	0,005	[0,991; 1,000]
	B	1,001	0,037	[1,000; 1,000]
MPP com correlação intra blocos	σ_μ^2	0,125	0,004	[0,117; 0,133]
	σ_y^2	0,721	0,143	[0,482; 1,005]
	p	0,982	0,017	[0,949; 0,999]
	B	1,116	0,396	[1,000; 2,000]

4.2 Cenário 2 - Dois pontos de mudança na média

Nesta seção têm-se dois casos a serem considerados, um assume que dentro dos blocos as observações são identicamente distribuídas, o outro caso considera também dois pontos de mudança, porém assume que dentro dos blocos as observações não são identicamente distribuídas. Desta forma, o cenário 2(a) favorece o MPP e o MPP com correlação entre blocos e o cenário 2(b) favorece o modelo com correlação intra blocos. Assim, têm-se os seguintes cenários:

$$\mathbf{Y} = \begin{cases} y_i \sim \mathcal{N}(10; 1), & 1 \leq i \leq 20; \\ y_i \sim \mathcal{N}(5; 1), & 21 \leq i \leq 40; \\ y_i \sim \mathcal{N}(8; 1), & 41 \leq i \leq 60. \end{cases} \quad \mathbf{Y}|\boldsymbol{\epsilon} = \begin{cases} y_i \sim \mathcal{N}(10 + \epsilon_i; 1), & 1 \leq i \leq 20; \\ y_i \sim \mathcal{N}(5 + \epsilon_i; 1), & 21 \leq i \leq 40; \\ y_i \sim \mathcal{N}(8 + \epsilon_i; 1), & 41 \leq i \leq 60. \end{cases}$$

Cenário 2(a)

Cenário 2(b)

onde $\epsilon_i \sim \mathcal{U}(-0,5; 0,5)$. A Figura 4.3 mostra as estimativas das médias *a posteriori* para ambos os cenários.

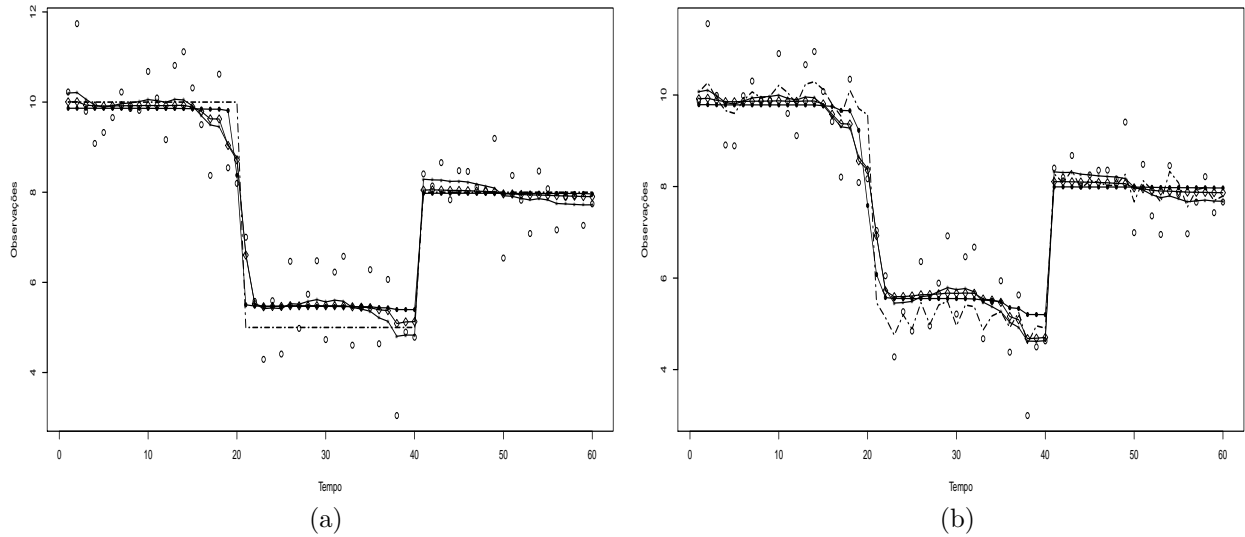


Figura 4.3: Estimativas das médias *a posteriori*. O símbolo \bullet indica o MPP com correlação entre blocos, \diamond são as estimativas via PPM, $*$ são estimativas via MPP com correlação intra blocos, a linha tracejada indica o parâmetro verdadeiro e o símbolo \circ indica os valores simulados da distribuição normal.

Em termos de estimativas para a média dos “clusters”, pode-se observar na Figura 4.3 que no cenário 2(a) o MPP com correlação entre blocos e o MPP apresentam estimativas mais suaves que o MPP com correlação intra blocos. Já no cenário 2(b), como era esperado, o MPP com correlação intra blocos apresenta estimativas ligeiramente melhores que os outros dois modelos. Nas Figuras 4.4 e 4.5 observa-se que o MPP superestima o número de blocos. Nota-se, também, que em ambos os cenários, o MPP com correlação entre blocos apresenta estimativas mais precisas em relação a posição da mudança.

A Tabela 4.6 que apresenta as partições mais prováveis para o cenário 2(a) mostra claramente que o MPP com correlação entre blocos indica a verdadeira partição com maior probabilidade quando comparado aos outros dois modelos. No cenário 2(b), cenário que favorece o MPP com correlação intra blocos, pode-se observar que esse modelo indica corretamente a partição com probabilidade 0,248. O modelo proposto nessa dissertação indica a partição $\rho = \{[0 - 19], [20 - 40], [41 - 60]\}$

com probabilidade 0,279. Note que apesar do cenário não ser favorável ao MPP com correlação entre blocos a partição estimada está muito próxima da verdadeira partição simulada.

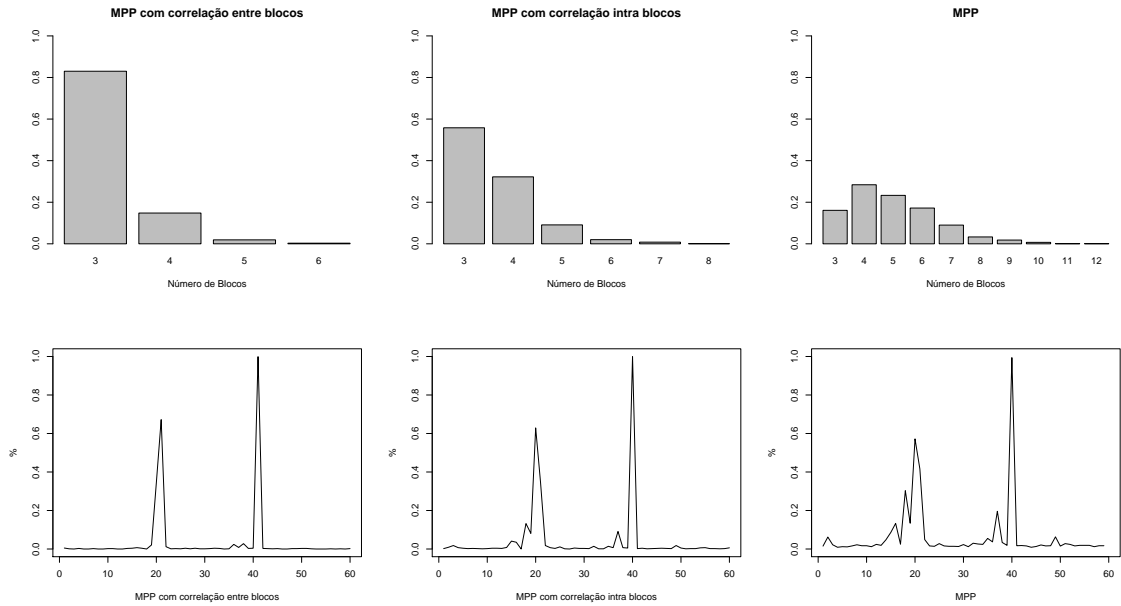


Figura 4.4: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto do cenário 2(a).

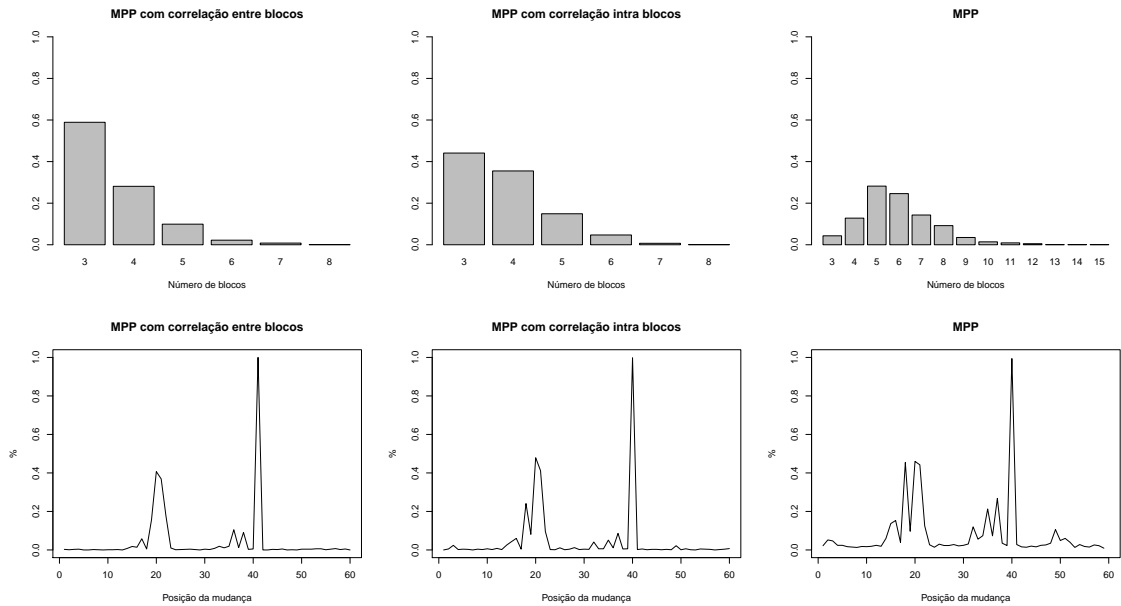


Figura 4.5: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto do cenário 2(b).

Tabela 4.6: Partições mais prováveis *a posteriori*.

Cenário 2(a)	
Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,551$
	$P(\rho = \{[0 - 19], [20 - 40], [41 - 60]\}) = 0,278$
MPP com correlação intra blocos	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,400$
	$P(\rho = \{[0 - 21], [22 - 40], [41 - 60]\}) = 0,130$
MPP	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,139$
	$P(\rho = \{[0 - 18], [19 - 21], [22 - 40], [41 - 60]\}) = 0,068$
Cenário 2(b)	
Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 19], [20 - 40], [41 - 60]\}) = 0,279$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,259$
MPP com correlação intra blocos	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,248$
	$P(\rho = \{[0 - 21], [22 - 40], [41 - 60]\}) = 0,145$
MPP	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,023$
	$P(\rho = \{[0 - 18], [19 - 21], [22 - 35], [36 - 40], [41 - 60]\}) = 0,023$

Tabela 4.7: Resumo das distribuições *a posteriori*.

Cenário 2(a)				
Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	29,260	19,103	[7,065; 65,860]
	σ_y^2	0,782	0,155	[0,482; 1,050]
	p	0,962	0,025	[0,913; 0,998]
	B	3,195	0,462	[3; 4]
MPP com correlação intra blocos	σ_μ^2	0,125	0,004	[0,117; 0,133]
	σ_y^2	0,646	0,142	[0,404; 0,931]
	p	0,945	0,031	[0,881; 0,990]
	B	3,601	0,811	[3; 5]
Cenário 2(b)				
Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	27,588	18,765	[7,169; 57,793]
	σ_y^2	0,826	0,180	[0,522; 1,205]
	p	0,956	0,029	[0,900; 0,998]
	B	3,582	0,831	[3; 5]
MPP com correlação intra blocos	σ_μ^2	0,125	0,004	[0,117; 0,134]
	σ_y^2	0,617	0,142	[0,371; 0,875]
	p	0,942	0,031	[0,881; 0,994]
	B	3,827	0,910	[3; 6]

Observa-se na Tabela 4.7, a qual mostra os resumos das distribuições *a posteriori*, dos cenários 2(a) e 2(b), respectivamente, que o MPP com correlação intra blocos em ambos os cenários subestima o valor de σ_y^2 , observa-se também que os intervalos HPD não contém o valor 1, ou seja, os intervalos HPD não contém o verdadeiro valor simulado de σ_y^2 .

4.3 Cenário 3 - Dois pontos de mudança na média: maior variabilidade

Nesta seção será investigado o comportamento dos três modelos discutidos nessa dissertação quando a série possui variância maior que 1. Especificamente foram geradas observações de uma distribuição Normal com variância 9. O esquema a seguir especifica a série analisada.

$$\mathbf{Y} = \begin{cases} y_i \sim \mathcal{N}(10; 9), & 1 \leq i \leq 30; \\ y_i \sim \mathcal{N}(13; 9), & 31 \leq i \leq 60; \\ y_i \sim \mathcal{N}(7; 9), & 61 \leq i \leq 90. \end{cases}$$

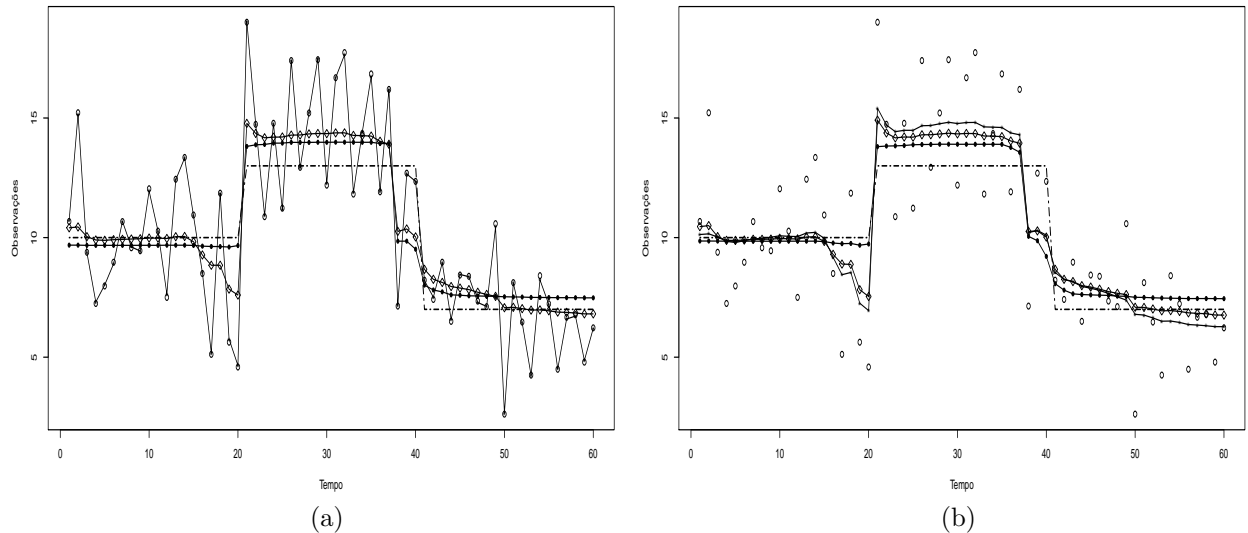


Figura 4.6: Estimativas das médias *a posteriori*. O símbolo \bullet indica o MPP com correlação entre blocos, \diamond são as estimativas via PPM, $*$ são estimativas via MPP com correlação intra blocos, a linha tracejada indica o parâmetro verdadeiro e o símbolo \circ indica os valores simulados da distribuição normal.

Somente nessa seção os cenários considerados não se referem a duas séries diferentes a serem analisa-

das. Aqui os cenários indicam se as distribuições *a priori* informativas para τ_μ^2 e σ_μ^2 são informativas ou não, ou seja, no cenário 3 (a) $\tau_\mu \sim Gama(800; 100)$ e $\sigma_\mu \sim Gama-Invertida(3; 30)$, já no cenário 3 (b) $\tau_\mu \sim Gama(0, 001; 0, 001)$ e $\sigma_\mu \sim Gama - Invertida(0, 1; 0, 1)$. Para o MPP as distribuições *a priori* foram as mesmas utilizadas nas seções anteriores.

Na Figura 4.6 (a) observa-se que as estimativas obtidas via MPP com correlação intra blocos são bem diferentes das obtidas via MPP e o modelo proposto nesse trabalho. Na Figura 4.6 (b) nota-se que o MPP com correlação entre blocos apresenta, de uma forma geral, estimativas mais suaves para as médias dos segmentos da série, seguido pelo MPP e o modelo proposto por Monteiro, Loschi & Assunção (2008).

Ao analisar a Figura 4.7 observa-se que o MPP com correlação intra blocos indica que há 90 pontos de mudança, enquanto que os modelos MPP e MPP com correlação entre blocos indicam 5 e 3 blocos, respectivamente. Na Figura 4.8 nota-se que mesmo com uma distribuição *a priori* não informativa o MPP com correlação intra blocos indica que existem 5 blocos quando na verdade existem 3. Assim, em virtude dos resultados obtidos no Cenário 3(a) e 3(b), existem evidências de que quando a série possui uma maior variabilidade deve-se optar por elicitar uma distribuição *a priori* para τ_μ menos informativa, caso contrário o modelo pode indicar erroneamente que todos os pontos são de mudança.

A Tabela 4.8 mostra que todos os modelos, em ambos os cenários, indicam a mesma partição como a mais provável, exceto o MPP com correlação intra blocos que no Cenário 3(a) indica que todos os pontos são pontos de mudança. Na Tabela 4.9 observa-se que ao utilizar uma distribuição *a priori* não informativa para σ_μ , a estimativa desse parâmetro foi muito alta. Porém, como pode-se observar através dos outros resultados discutidos, essa estimativa não comprometeu o desempenho do modelo proposto indicando que o mesmo não é sensível a escolha da distribuição *a priori* para σ_μ .

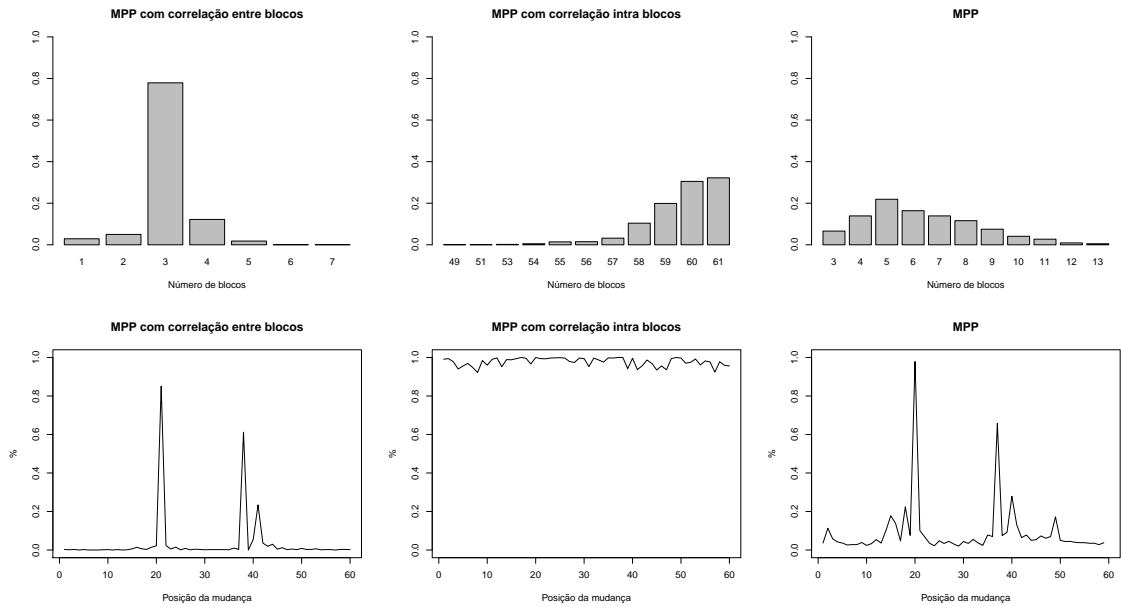


Figura 4.7: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto - Cenário 3(a).

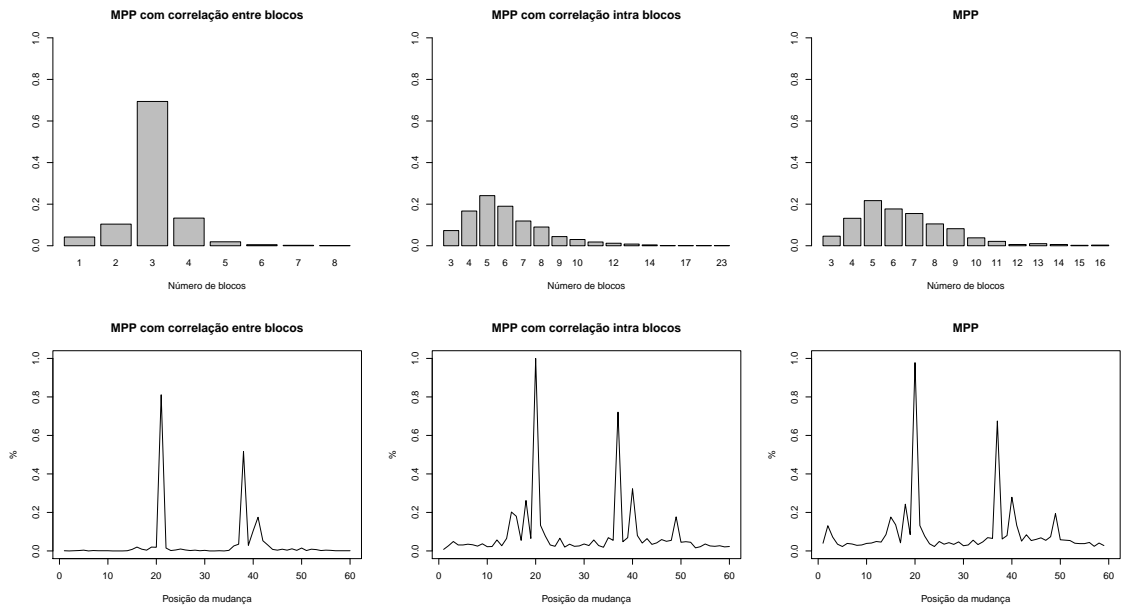


Figura 4.8: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto - Cenário 3(b).

Tabela 4.8: Partições mais prováveis *a posteriori*.

Cenário 3(a)	
Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 20], [21 - 37], [38 - 60]\}) = 0,456$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,177$
MPP com correlação intra blocos	$P(\rho = \{[1], [2], \dots, [60]\}) = 0,322$
	$P(\rho = \{[1], [2], \dots, [5], [6 - 7], [8], \dots, [59], [60]\}) = 0,018$
MPP	$P(\rho = \{[0 - 20], [21 - 37], [38 - 60]\}) = 0,037$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,017$
Cenário 3(b)	
Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 20], [21 - 37], [38 - 60]\}) = 0,381$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,136$
MPP com correlação intra blocos	$P(\rho = \{[0 - 20], [21 - 37], [38 - 60]\}) = 0,053$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,019$
MPP	$P(\rho = \{[0 - 20], [21 - 37], [38 - 60]\}) = 0,032$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60]\}) = 0,016$

Tabela 4.9: Resumo das distribuições *a posteriori*.

Cenário 3(a)				
Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	37,340	34,252	[8,221; 83,685]
	σ_y^2	7,333	2,203	[3,868; 11,107]
	p	0,965	0,025	[0,916; 1,000]
	B	3,057	0,618	[2; 4]
MPP com correlação intra blocos	σ_μ^2	0,125	0,004	[0,116; 0,133]
	σ_y^2	3,016	8,127	[0,159; 9,684]
	p	0,096	0,043	[0,023; 0,178]
	B	59,625	1,469	[57; 61]
Cenário 3(b)				
Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	2,325e+18	6,171e+19	[12,461; 3521,223]
	σ_y^2	7,654	2,771	[4,152; 13,524]
	p	0,966	0,026	[0,917; 1,000]
	B	3,011	0,764	[1; 4]
MPP com correlação intra blocos	σ_μ^2	0,128	0,046	[0,055; 0,214]
	σ_y^2	5,544	1,169	[3,592; 7,811]
	p	0,908	0,049	[0,813; 0,990]
	B	6,057	2,260	[3; 10]

4.4 Cenário 4 - Quatro pontos de mudança na média

Nesta seção é considerada uma série com 5 “clusters”. O padrão considerado nesse cenário é irregular, ou seja, ora a média apresenta pequenas diferenças entre um bloco e outro, ora a mudança é mais brusca. A série gerada é descrita a seguir:

$$\mathbf{Y} = \begin{cases} y_i \sim \mathcal{N}(5; 1), & 1 \leq i \leq 20; \\ y_i \sim \mathcal{N}(6, 5; 1), & 21 \leq i \leq 40; \\ y_i \sim \mathcal{N}(9; 1), & 41 \leq i \leq 60; \\ y_i \sim \mathcal{N}(4; 1), & 61 \leq i \leq 80; \\ y_i \sim \mathcal{N}(1; 1), & 81 \leq i \leq 100. \end{cases}$$

Observa-se na Figura 4.9 que as estimativas via MPP e MPP com correlação entre blocos são muito próximas, quando a diferença entre as médias é grande, como é o caso do bloco 3 para o bloco 4. Nota-se que o modelo proposto apresenta as melhores estimativas na maioria dos blocos e que o modelo proposto por Monteiro, Loschi & Assunção é o que apresenta estimativas levemente piores.

Em relação ao número de “clusters”, a Figura 4.10 mostra que o modelo proposto é o que indica maior probabilidade *a posteriori* de haver cinco blocos. Já o MPP indica que o número mais provável de blocos é seis.

A Tabela 4.10 mostra que todos os modelos identificaram corretamente a partição, porém o MPP com correlação entre blocos indica a verdadeira partição com maior probabilidade, $P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 80], [81 - 100]\}) = 0,682$. Na Tabela 4.11 observa-se que o MPP com correlação intra blocos mais uma vez subestima o valor de σ_y^2 . Nessa mesma tabela observa-se que para os parâmetros p e B , o modelo proposto nesse trabalho produz intervalos HPD menores que os intervalos obtidos pelo modelo proposto por Monteiro, Loschi & Assunção (2008).

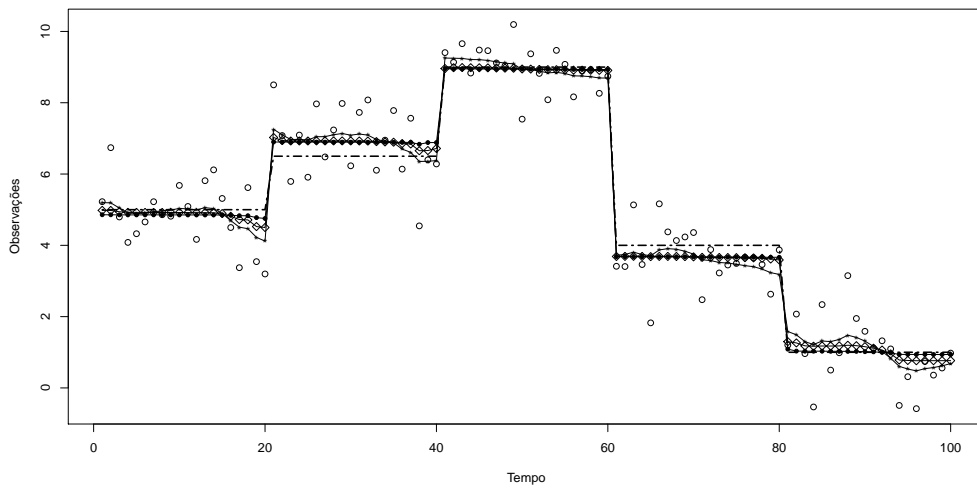


Figura 4.9: Estimativas das médias *a posteriori*. O símbolo \bullet indica o MPP com correlação entre blocos, \diamond são as estimativas via PPM, $*$ são estimativas via MPP com correlação intra blocos, a linha tracejada indica o parâmetro verdadeiro e o símbolo \circ indica os valores simulados da distribuição normal.

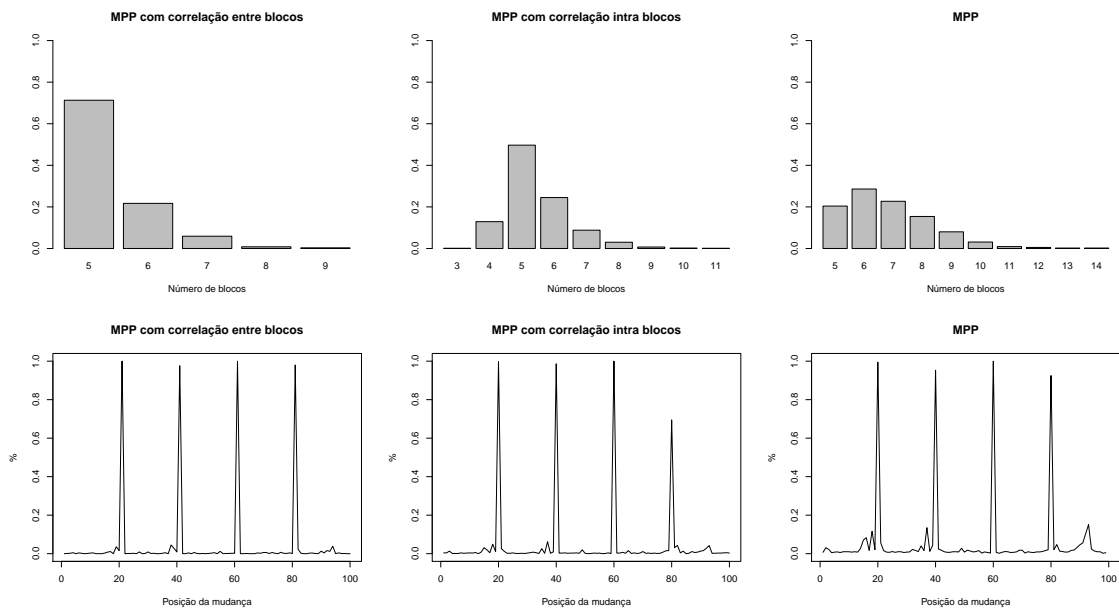


Figura 4.10: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto.

Tabela 4.10: Partições mais prováveis *a posteriori* - Cenário 4.

Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 80], [81 - 100]\}) = 0,682$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 80], [81 - 93], [94 - 100]\}) = 0,025$
MPP com correlação intra blocos	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 80], [81 - 100]\}) = 0,381$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 100]\}) = 0,126$
MPP	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 80], [81 - 100]\}) = 0,175$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 80], [81 - 93], [94 - 100]\}) = 0,030$

Tabela 4.11: Resumo das distribuições *a posteriori* - Cenário 4.

Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	15,296	8,815	[5,071; 31,641]
	σ_y^2	0,768	0,118	[0,562; 1,006]
	p	0,956	0,021	[0,915; 0,991]
	B	5,371	0,660	[5; 7]
MPP com correlação intra blocos	σ_μ^2	0,125	0,004	[0,117; 0,134]
	σ_y^2	0,671	0,113	[0,468; 0,898]
	p	0,950	0,023	[0,904; 0,987]
	B	5,424	1,010	[4; 7]

4.5 Cenário 5 - Cinco pontos de mudança na média: mudanças alternadas

Neste caso foram considerados dois cenários com as mesmas médias, porém no cenário 5(a) os blocos possuem 20 observações e no cenário 5(b) o tamanho dos “clusters” é de 30 observações. A seguir as séries consideradas são especificadas:

$$\mathbf{Y} = \begin{cases} y_i \sim \mathcal{N}(0; 1), & 1 \leq i \leq 20; \\ y_i \sim \mathcal{N}(3; 1), & 21 \leq i \leq 40; \\ y_i \sim \mathcal{N}(0; 1), & 41 \leq i \leq 60; \\ y_i \sim \mathcal{N}(3; 1), & 61 \leq i \leq 80; \\ y_i \sim \mathcal{N}(0; 1), & 81 \leq i \leq 100; \\ y_i \sim \mathcal{N}(3; 1), & 101 \leq i \leq 120. \end{cases}$$

Cenário 5(a)

$$\mathbf{Y} = \begin{cases} y_i \sim \mathcal{N}(0; 1), & 1 \leq i \leq 30; \\ y_i \sim \mathcal{N}(3; 1), & 31 \leq i \leq 60; \\ y_i \sim \mathcal{N}(0; 1), & 61 \leq i \leq 90; \\ y_i \sim \mathcal{N}(3; 1), & 91 \leq i \leq 120; \\ y_i \sim \mathcal{N}(0; 1), & 121 \leq i \leq 150; \\ y_i \sim \mathcal{N}(3; 1), & 151 \leq i \leq 180. \end{cases}$$

Cenário 5(b)

A Figura 4.11 (a) mostra que quando o tamanho do bloco é de 20 observações, o modelo proposto não identifica corretamente os pontos de mudança, além disso o MPP apresenta estimativas mais suaves para as médias dos blocos que o modelo com correlação intra blocos. Na Figura 4.11 (b) nota-se que a medida que o tamanho dos blocos aumenta o modelo com correlação entre blocos é capaz de identificar melhor os pontos de mudança e apresenta estimativas mais suaves para as médias quando comparado aos demais modelos.

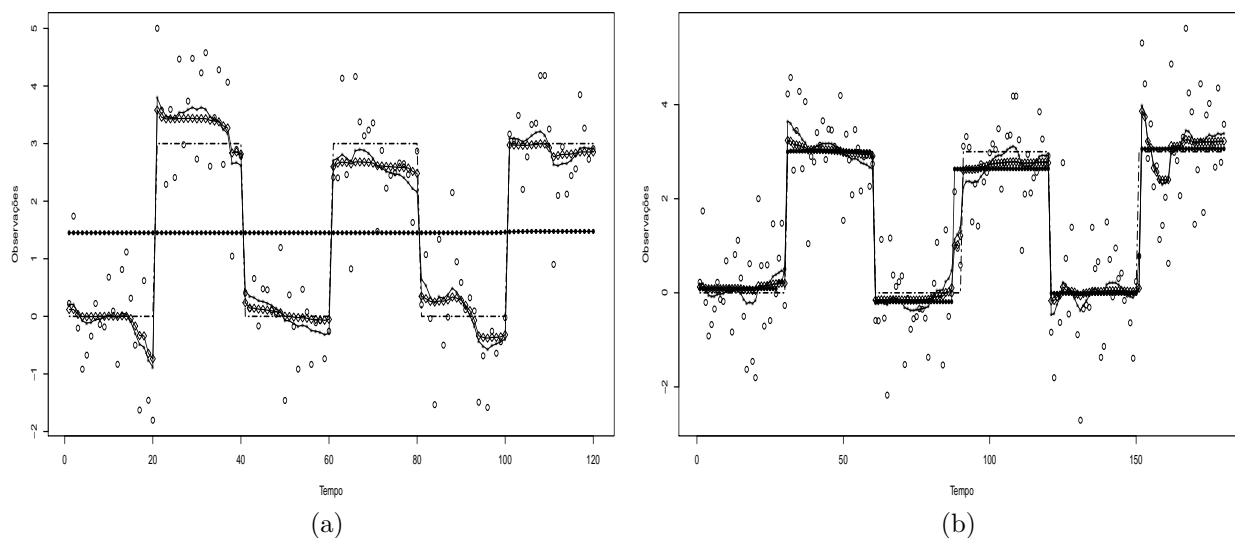


Figura 4.11: Estimativas das médias *a posteriori*. O símbolo ● indica o MPP com correlação entre blocos, ◇ são as estimativas via PPM, * são estimativas via MPP com correlação intra blocos, a linha tracejada indica o parâmetro verdadeiro e o símbolo ○ indica os valores simulados da distribuição normal.

Nas Figuras 4.12 e 4.13 observa-se que no cenário 5(a) o modelo com correlação intra blocos foi o único que identificou corretamente o número de “clusters”. Já no cenário 5(b) o MPP com cor-

relação entre blocos foi o único modelo que identificou corretamente o número de blocos.

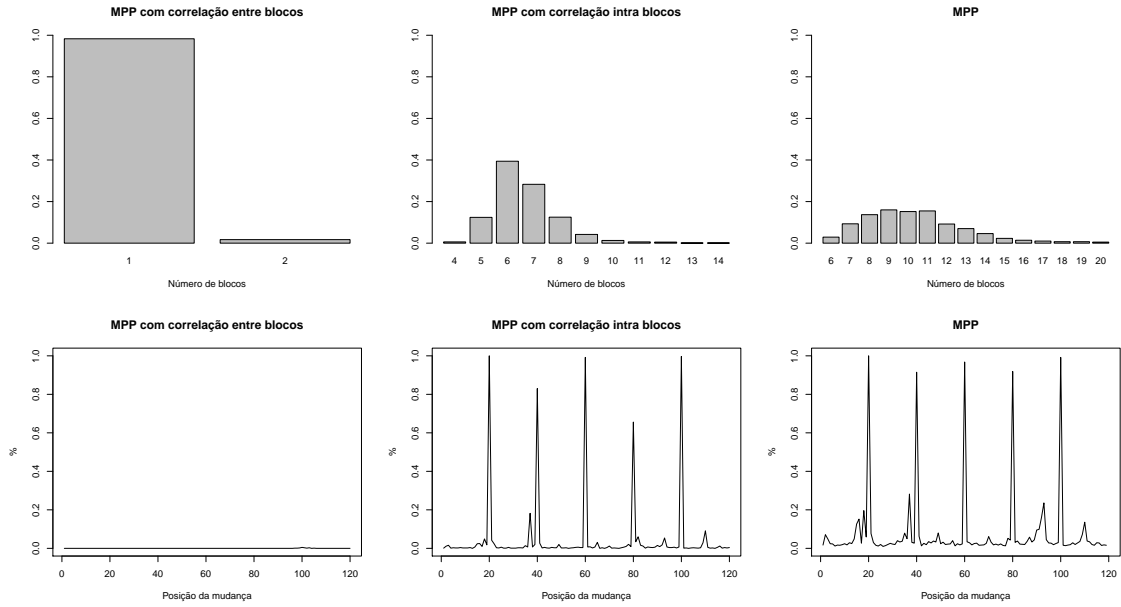


Figura 4.12: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto do cenário 5(a).

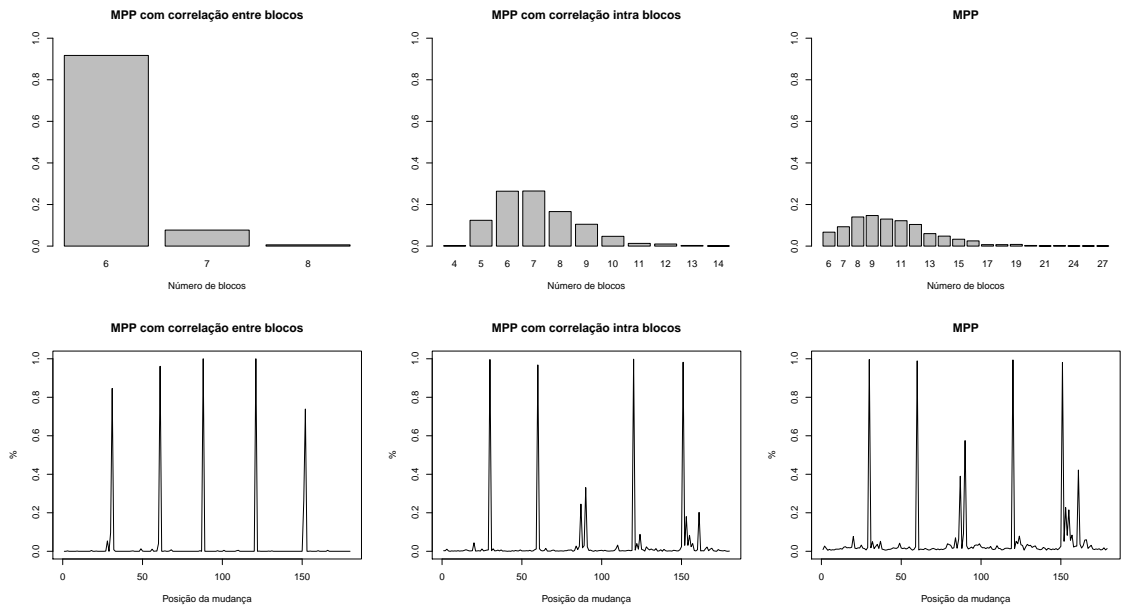


Figura 4.13: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto do cenário 5(b).

Tabela 4.12: Partições mais prováveis *a posteriori*.

Cenário 5(a)	
Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 120]\}) = 0,983$ $P(\rho = \{[0 - 100], [101 - 120]\}) = 0,005$
MPP com correlação intra blocos	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 80], [81 - 100], [101 - 120]\}) = 0,235$ $P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 100], [101 - 120]\}) = 0,095$
MPP	$P(\rho = \{[1 - 20], [21 - 40], [41 - 60], [61 - 80], [80 - 101], [101 - 120]\}) = 0,026$ $P(\rho = \{[1 - 20], [21 - 40], [41 - 60], [61 - 80], [80 - 93], [94 - 100], [101 - 120]\}) = 0,010$
Cenário 5(b)	
Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 30], [31 - 60], [61 - 87], [88 - 120], [121 - 151], [152 - 180]\}) = 0,690$ $P(\rho = \{[0 - 30], [31 - 60], [62 - 87], [88 - 120], [121 - 150], [151 - 180]\}) = 0,089$
MPP com correlação intra blocos	$P(\rho = \{[0 - 30], [31 - 60], [61 - 120], [121 - 151], [152 - 180]\}) = 0,115$ $P(\rho = \{[0 - 30], [31 - 60], [61 - 90], [91 - 120], [121 - 151], [152 - 180]\}) = 0,064$
MPP	$P(\rho = \{[0, 30], [31, 60], [61, 90], [91, 120], [121, 151], [152, 180]\}) = 0,038$ $P(\rho = \{[0, 30], [31, 60], [61, 90], [91, 120], [121, 150], [151, 180]\}) = 0,021$

Na Tabela 4.12 observa-se uma inversão no comportamento dos modelos MPP com correlação entre e MPP com correlação intra blocos. Nota-se que no cenário 5(a) o MPP com correlação entre blocos indica que não há mudança na média, porém no cenário 5(b) esse modelo identifica uma partição muito próxima da verdadeira. Já o MPP com correlação intra blocos no cenário 5(b) indica que a partição mais provável não contém os blocos $[61 - 90]$, $[91 - 120]$.

Analisando a Tabela 4.13 nota-se que as estimativas do MPP com correlação entre blocos não foram satisfatórias no cenário 5(a), no entanto o MPP com correlação intra blocos também deixou a desejar na estimativa da variância da série. Já no cenário 5(b) observa-se que o MPP com correlação entre blocos apresenta intervalos HPD menores que o MPP com correlação intra blocos.

Tabela 4.13: Resumo das distribuições *a posteriori*.

Cenário 5(a)				
Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	15,200	13,090	[3,256; 40,112]
	σ_y^2	3,123	0,402	[2,427; 3,929]
	p	0,999	0,004	[0,994; 1,000]
	B	1,017	0,129	[1; 1]
MPP com correlação intra blocos	σ_μ^2	0,125	0,004	[0,117; 0,133]
	σ_y^2	0,644	0,097	[0,471; 0,842]
	p	0,947	0,021	[0,910; 0,986]
	B	6,650	1,241	[5; 9]
Cenário 5(b)				
Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	11,173	5,690	[3,496; 20,489]
	σ_y^2	0,999	0,115	[0,789; 1,221]
	p	0,971	0,013	[0,945; 0,992]
	B	6,089	0,305	[6; 7]
MPP com correlação intra blocos	σ_μ^2	0,125	0,004	[0,116; 0,133]
	σ_y^2	0,780	0,248	[0,050; 1,020]
	p	0,962	0,016	[0,931; 0,990]
	B	7,126	1,546	[5; 10]

4.6 Cenário 6 - Quatro pontos de mudança na média: série decrescente e crescente

Neste estudo de caso foram levados em consideração dois aspectos: a diferença entre as médias de duas unidades e o fato da série ser crescente ou decrescente. O esquema a seguir especifica as séries geradas:

$$\mathbf{Y} = \begin{cases} y_i \sim \mathcal{N}(10; 1), & 1 \leq i \leq 20; \\ y_i \sim \mathcal{N}(8; 1), & 21 \leq i \leq 40; \\ y_i \sim \mathcal{N}(6; 1), & 41 \leq i \leq 60; \\ y_i \sim \mathcal{N}(4; 1), & 61 \leq i \leq 80; \\ y_i \sim \mathcal{N}(2; 1), & 81 \leq i \leq 100. \end{cases}$$

Cenário 6(a)

$$\mathbf{Y} = \begin{cases} y_i \sim \mathcal{N}(2; 1), & 1 \leq i \leq 20; \\ y_i \sim \mathcal{N}(4; 1), & 21 \leq i \leq 40; \\ y_i \sim \mathcal{N}(6; 1), & 41 \leq i \leq 60; \\ y_i \sim \mathcal{N}(8; 1), & 61 \leq i \leq 80; \\ y_i \sim \mathcal{N}(10; 1), & 81 \leq i \leq 100. \end{cases}$$

Cenário 6(b)

Através da Figura 4.14 observa-se que no cenário 6(a) o MPP com correlação intra blocos tende a suavizar demais as estimativas das médias a ponto de não identificar mudança nas médias, esse

cenário também foi considerado para blocos com 30 observações e distribuição *a priori* menos informativa para τ_μ , ou seja $\tau_\mu \sim Gama(0,001;0,001)$, e os resultados obtidos foram similares aos apresentados aqui. No cenário 6(b), no geral, todos os modelos apresentaram estimativas muito próximas, com uma leve vantagem no primeiro e no último bloco para o modelo proposto, pois apresentou estimativas mais próximas do verdadeiro valor da média do bloco.

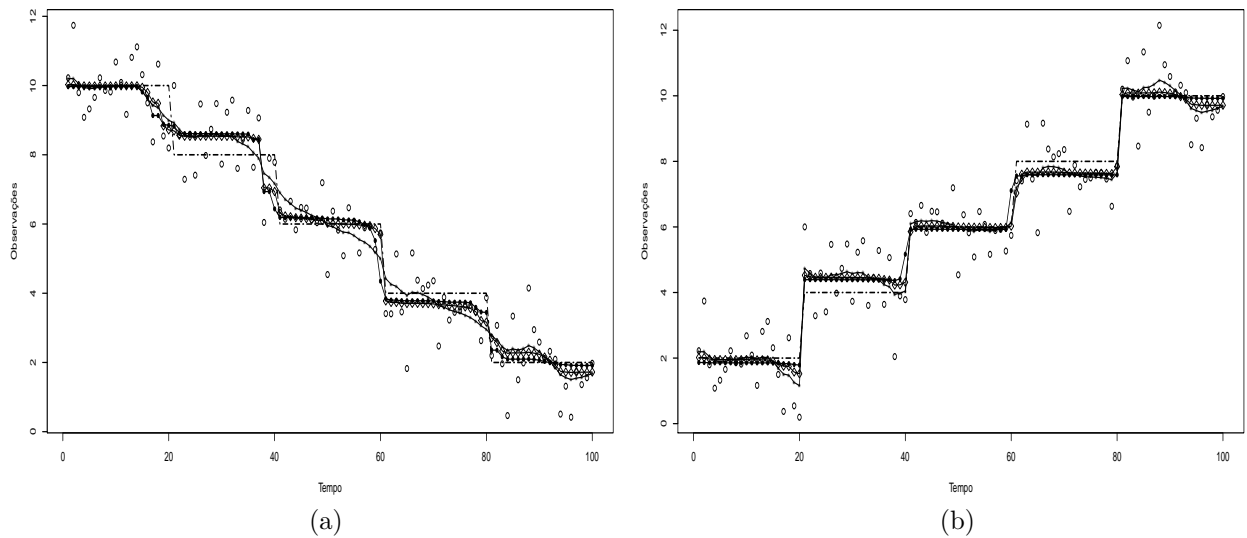


Figura 4.14: Estimativas das médias *a posteriori*. O símbolo \bullet indica o MPP com correlação entre blocos, \diamond são as estimativas via PPM, $*$ são estimativas via MPP com correlação intra blocos, a linha tracejada indica o parâmetro verdadeiro e o símbolo \circ indica os valores simulados da distribuição normal.

Nas Figuras 4.15 e 4.16 pode-se observar que no cenário 6(a) o modelo proposto indica corretamente o número de blocos. O MPP original superestima o número de blocos e o MPP com correlação intra bloco indica que não há pontos de mudança na série. No cenário 6(b) o MPP indica que há seis blocos, o MPP com correlação intra blocos e o modelo proposto indicam corretamente o número de blocos. Porém, vale ressaltar, que o modelo com correlação entre blocos apresenta maior probabilidade *a posteriori* na hora de indicar o número correto de blocos, ou seja, o modelo proposto indica que $P(B = 5) \approx 0,8$ e o MPP com correlação intra blocos indica que $P(B = 5) \approx 0,4$.

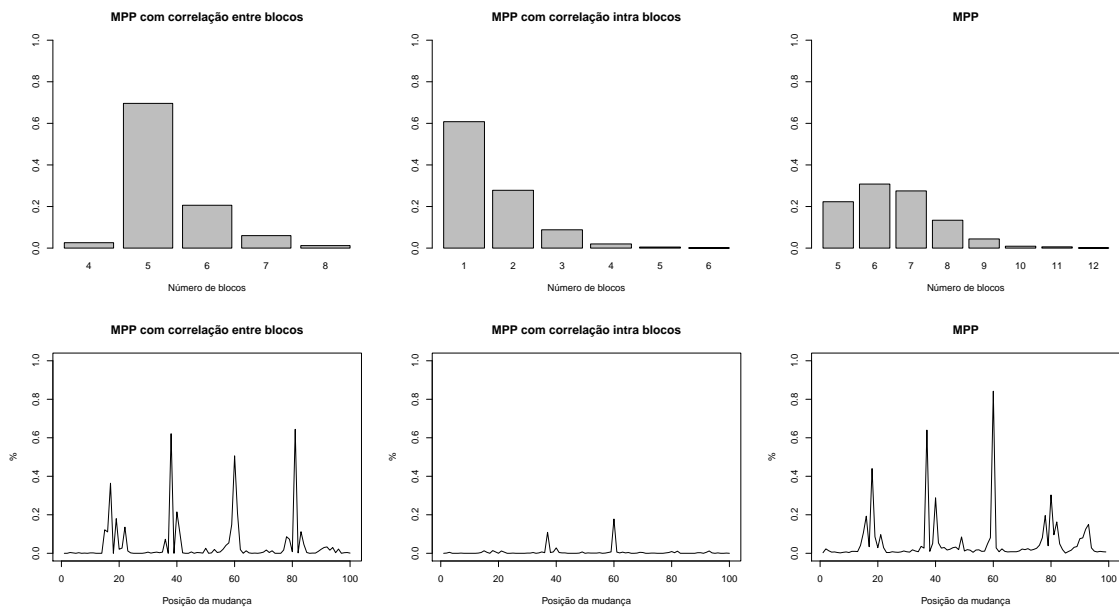


Figura 4.15: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto do cenário 6(a).

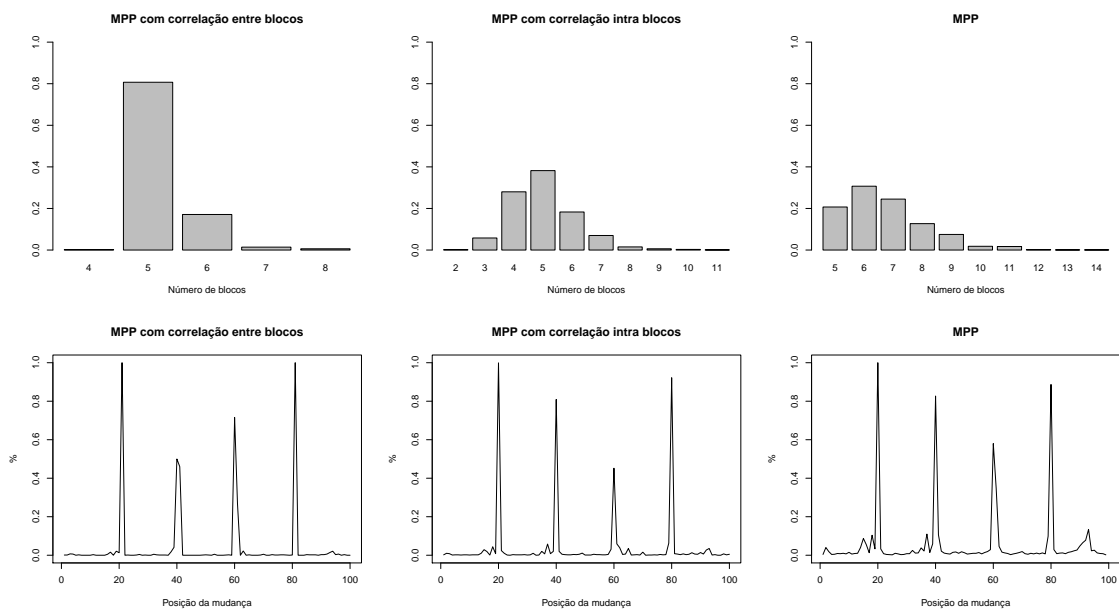


Figura 4.16: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto do cenário 6(b).

Tabela 4.14: Partições mais prováveis *a posteriori*.

Cenário 6(a)	
Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 16], [17 - 39], [40 - 59], [60 - 80], [81 - 100]\}) = 0,169$
	$P(\rho = \{[0 - 18], [19 - 37], [38 - 60], [61 - 80], [81 - 100]\}) = 0,090$
MPP com correlação intra blocos	$P(\rho = \{[0 - 100]\}) = 0,600$
	$P(\rho = \{[0 - 60], [61 - 100]\}) = 0,098$
MPP	$P(\rho = \{[0 - 18], [19 - 37], [38 - 60], [61 - 80], [81 - 100]\}) = 0,016$
	$P(\rho = \{[0 - 18], [19 - 37], [38 - 60], [61 - 78], [79 - 100]\}) = 0,014$
Cenário 6(b)	
Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[0 - 20], [21 - 39], [40 - 59], [60 - 80], [81 - 100]\}) = 0,351$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 59], [60 - 80], [81 - 100]\}) = 0,202$
MPP com correlação intra blocos	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 80], [81 - 100]\}) = 0,192$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 80], [81 - 100]\}) = 0,183$
MPP	$P(\rho = \{[0 - 20], [21 - 40], [41 - 60], [61 - 80], [81 - 100]\}) = 0,071$
	$P(\rho = \{[0 - 20], [21 - 40], [41 - 61], [62 - 80], [81 - 100]\}) = 0,054$

Tabela 4.15: Resumo das distribuições *a posteriori*.

Cenário 6(a)				
Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	9,764	5,026	[3,242; 19,999]
	σ_y^2	0,827	0,135	[0,573; 1,072]
	p	0,957	0,021	[0,917; 0,993]
	B	5,215	0,487	[5; 6]
MPP com correlação intra blocos	σ_μ^2	0,125	0,005	[0,116; 0,134]
	σ_y^2	0,697	0,124	[0,484; 0,939]
	p	0,953	0,024	[0,908; 0,993]
	B	5,011	1,156	[3; 7]
Cenário 6(b)				
Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	23,857	13,530	[7,02425; 50,383]
	σ_y^2	0,781	0,140	[0,540; 1,078]
	p	0,958	0,0199	[0,919; 0,990]
	B	5,274	0,776	[4; 7]
MPP com correlação intra blocos	σ_μ^2	0,125	0,004	[0,117; 0,134]
	σ_y^2	0,672	0,114	[0,459; 0,880]
	p	0,988	0,012	[0,964; 0,999]
	B	1,300	0,590	[1; 2]

Neste capítulo são apresentadas e discutidas duas aplicações a dados reais dos modelos descritos nas Seções 2.1.2, 2.2 e 3.1. A primeira aplicação utiliza a série temporal do preço médio do quilo do tomate no período que se estende de 1992 a 2009. A segunda aplicação refere-se ao conjunto de dados denominado cromossomo 11 - coriell, esse conjunto de observações faz parte do banco de dados do software R e foi utilizado por Erdman & Emerson (2007) como uma aplicação do MPP.

Ambos os conjuntos de dados foram analisados sob as mesmas distribuições *a priori* eliciadas no capítulo anterior, exceto os parâmetros σ_μ^2 e τ_μ cujo as distribuições *a priori* utilizadas foram *Gama – Invertida*(3; 10) e *Gama*(10; 1), respectivamente.

5.1 Série do Preço Médio do Tomate

Nesta seção serão apresentados os resultados para a série do preço médio mensal do quilo do tomate (recebido pelo produtor brasileiro) no período de janeiro de 1992 a setembro de 2009.

Optou-se por realizar essa aplicação a série do preço médio do tomate, pois identificar se a série

possui mudança na média segundo Perron (1990) é importante, uma vez que a presença de quebra estrutural no nível da série pode levar a inferir erroneamente sobre a presença de uma raiz unitária, quando de fato a série em estudo não apresenta esta característica.

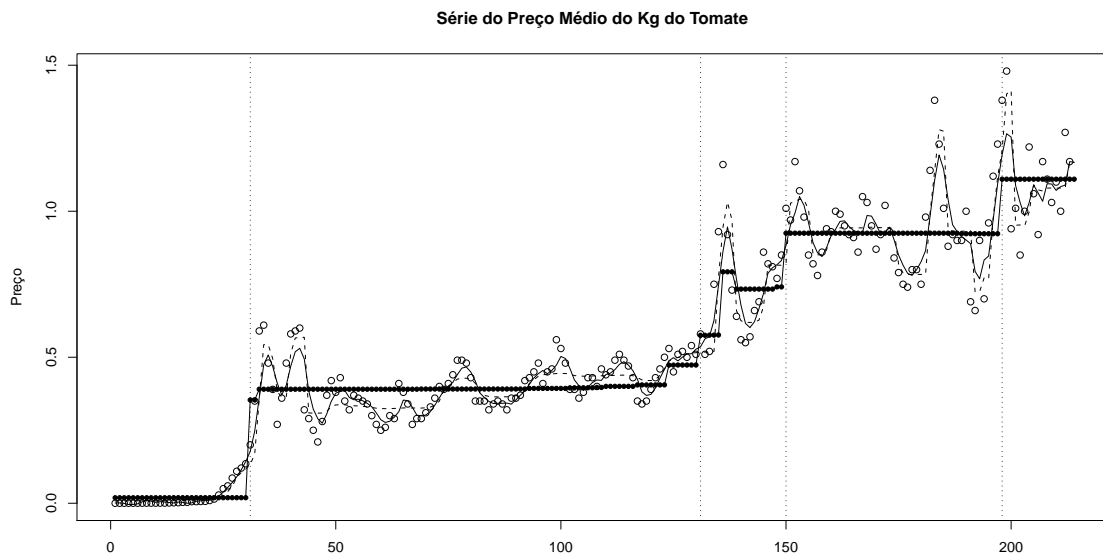


Figura 5.1: Estimativas das médias *a posteriori* para a série do preço médio do Kg do tomate. O símbolo ● indica o MPP com correlação entre blocos, a linha tracejada indica as estimativas via PPM, a linha contínua são as estimativas via MPP com correlação intra blocos e o símbolo ○ indica os valores observados. As linhas tracejadas verticais indicam a partição mais provável obtida via o MPP com correlação entre blocos.

Ao analisar a Figura 5.1 percebe-se que o modelo proposto apresenta estimativas mais suaves para as médias que os demais modelos. A Figura 5.2 mostra que o modelo proposto na presente dissertação apresenta estimativas mais realistas em relação ao número de blocos, uma vez que o MPP com correlação intra blocos tende a subestimar o número de blocos e o MPP tende a superestimar o número de segmentos na série.

A Tabela 5.1 indica que o MPP com correlação entre blocos apresenta uma melhor estimativa da partição mais provável, uma vez que o MPP com correlação intra blocos indica que não há pontos de mudança na série e o MPP identifica muitas possíveis partições com a mesma probabilidade, indicando uma grande incerteza em relação a partição ρ . A partição mais provável indicada pelo

modelo proposto nessa dissertação é ilustrada na Figura 5.1.

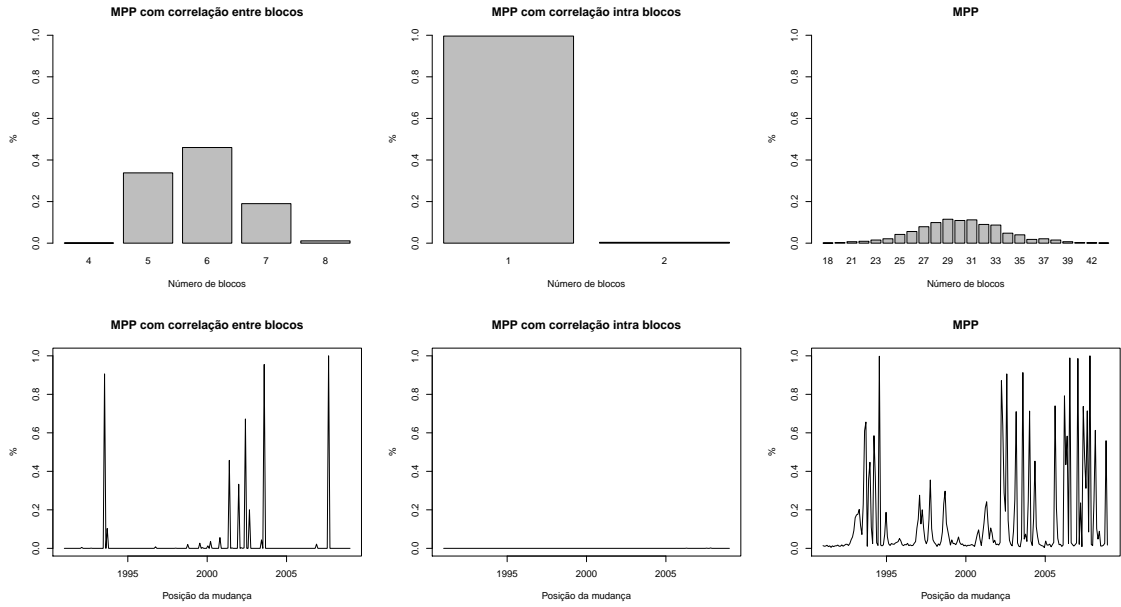


Figura 5.2: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto da série do preço médio do Kg do tomate

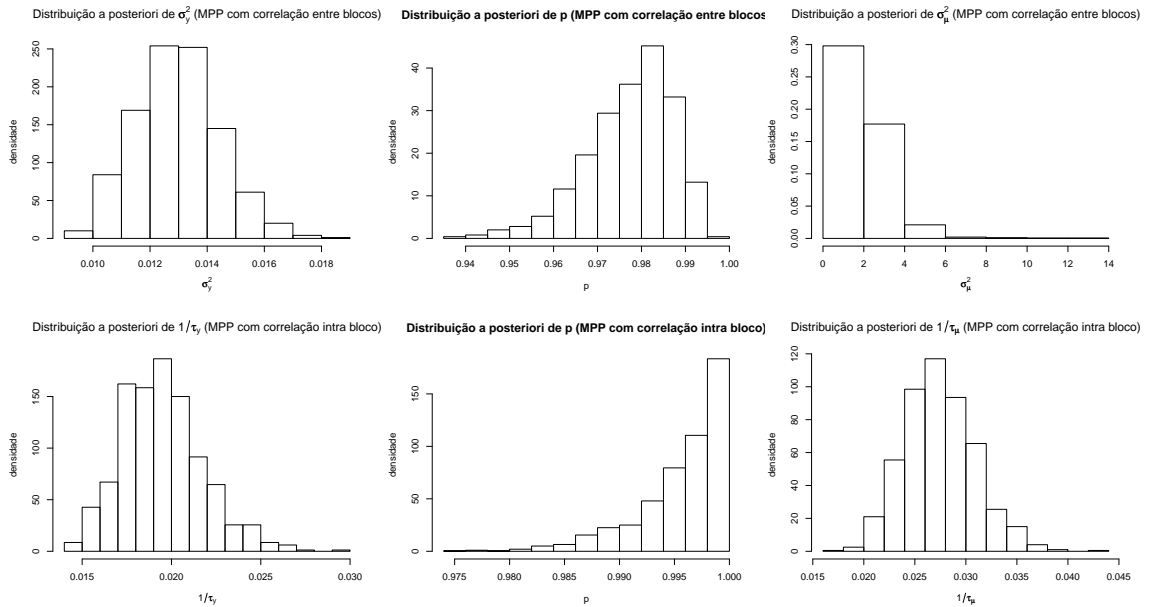


Figura 5.3: Distribuições *a posteriori* dos parâmetros σ_y^2 , p e σ_μ^2 para a série do preço médio do Kg do tomate

Nota-se na Tabela 5.2 que mesmo utilizando uma distribuição *a priori* menos informativa que nos cenários gerados, as estimativas *a posteriori* para a média do parâmetro σ_μ^2 , em ambos os modelos, é menor que a média *a priori* (a média *a priori* para σ_μ^2 no MPP com correlação entre blocos é 5 e para o MPP com correlação intra blocos a média é 0,11). Na Figura 5.3 observa-se que em ambos os modelos as distribuições *a posteriori* para todos os parâmetros são assimétricas e possuem uma única moda.

Tabela 5.1: Partições mais prováveis *a posteriori* - Série do Preço Médio do Kg do Tomate.

Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[1 - 30], [31 - 130], [131 - 149], [150 - 197], [198 - 214]\}) = 0,256$ $P(\rho = \{[1 - 30], [31 - 123], [124 - 135], [136 - 149], [150 - 197], [198 - 214]\}) = 0,250$
MPP com correlação intra blocos	$P(\rho = \{[1 - 214]\}) = 0,996$ $P(\rho = \{[1 - 200], [201 - 214]\}) = 0,002$
MPP	Muitas partições com a mesma probabilidade

Tabela 5.2: Resumo das distribuições *a posteriori* - Série do Preço Médio do Kg do Tomate.

Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	2,054	1,090	[0,626; 4,018]
	σ_y^2	0,013	0,001	[0,010; 0,016]
	p	0,977	0,010	[0,958; 0,994]
	B	5,872	0,748	[5; 7]
MPP com correlação intra blocos	σ_μ^2	0,027	0,003	[0,021; 0,034]
	σ_y^2	0,019	0,002	[0,015; 0,024]
	p	0,996	0,004	[0,988; 0,999]
	B	1,004	0,063	[1; 1]

Como os resultados do Cenário 3 indicam que as estimativas obtidas via MPP com correlação intra blocos são afetadas pela escolha da distribuição *a priori* do parâmetro τ_μ , a análise desse conjunto de dados também foi feita utilizando uma distribuição *a priori* menos informativa para esse parâmetro, ou seja, $\tau_\mu \sim Gama(0,001; 0,001)$, no entanto os resultados obtidos foram similares.

Em virtude dos resultados apresentados nessa seção nota-se que, apesar da análise visual da série indicar que a variância não é constante ao longo do tempo, o modelo proposto apresentou um

comportamento satisfatório para a identificação de pontos de mudança.

5.2 Cromossomo 11 - Coriell

Erdman & Emerson (2007) ilustram uma aplicação do MPP através do conjunto de dados coriell, nessa seção esse mesmo conjunto de dados será analisado através dos três modelos discutidos na presente dissertação. As observações são referentes ao cromossomo 11 do arquivo coriell e podem ser encontradas no pacote `bcp` implementado no R por esses mesmos autores. O arquivo coriell é referente a um estudo do *Coriell Institute for Medical Research*, sobre a Hibridação Genômica Comparativa (HGC). A HGC foi desenvolvida como um método para a detecção e mapeamento de aberrações cromossômicas no genoma.

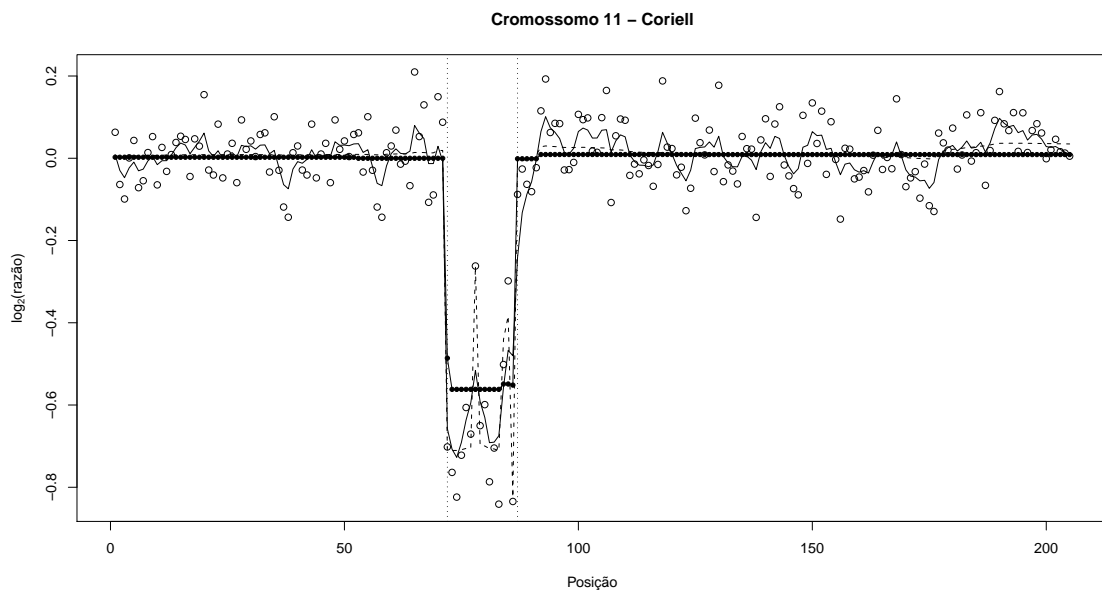


Figura 5.4: Estimativas das médias *a posteriori* para o cromossomo 11 - coriell. O símbolo \bullet indica o MPP com correlação entre blocos, a linha tracejada indica as estimativas via PPM, a linha contínua são as estimativas via MPP com correlação intra blocos e o símbolo \circ indica os valores observados. As linhas tracejadas verticais indicam a partição mais provável obtida via o MPP com correlação entre blocos.

Ao analisar a Figura 5.4 nota-se que as estimativas obtidas via MPP com correlação entre blocos são mais suaves que a estimativas fornecidas pelo MPP, indicando que o modelo proposto é menos

sensível a influência de observações atípicas. As estimativas obtidas para as médias via o MPP com correlação intra blocos indicam que, mesmo dentro dos dois maiores segmentos, a média não é constante ao longo do tempo. A Figura 5.5 indica, mais uma vez, a tendência do MPP em superestimar o número de blocos, como foi visto em alguns cenários do capítulo anterior.

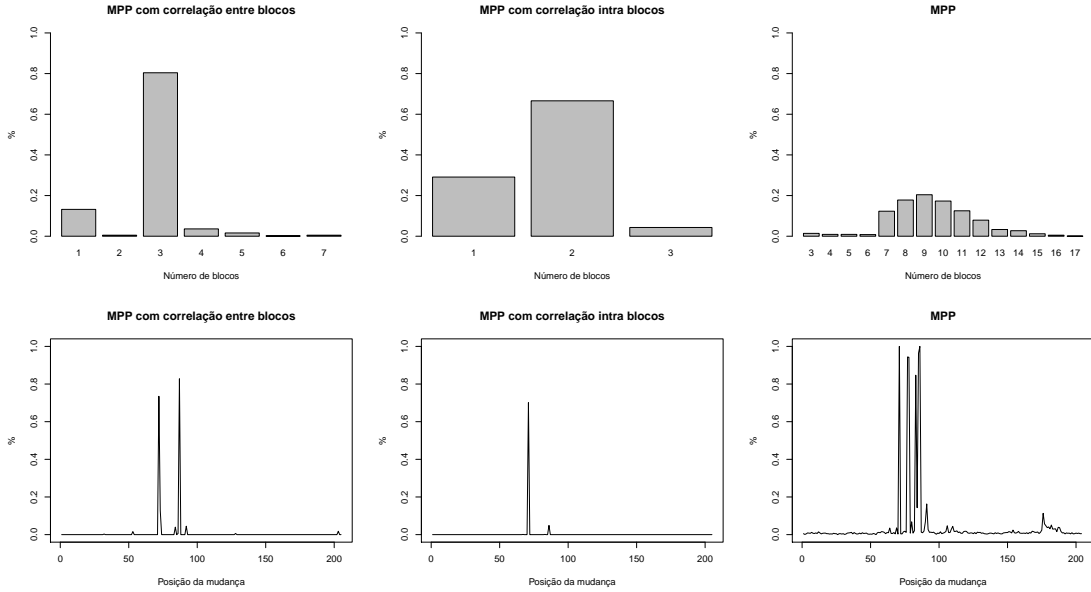


Figura 5.5: Distribuições *a posteriori* do número de blocos e probabilidade de mudança em cada ponto para o cromossomo 11 - coriell

Na Tabela 5.3 observa-se que o MPP com correlação intra blocos só identifica dois blocos. No entanto, visualmente, as observações aparentam estar divididas em pelo menos três segmentos. Analisando a Figura 5.6 e a Tabela 5.4 nota-se que as estimativas e distribuições *a posteriori* para o parâmetro p obtidas, via MPP com correlação entre blocos e MPP com correlação intra blocos, estão muito próximas. Porém, em relação ao parâmetro σ_y^2 não se pode afirmar o mesmo, uma vez que as distribuições *a posteriori* possuem formas bem diferentes.

O conjunto de dados cromossomo 11 - coriell também foi analisado sob uma distribuição *a priori* não informativa para τ_μ , ou seja, $\tau_\mu \sim Gama(0,001; 0,001)$. Nesse caso o MPP com correlação intra blocos identificou a mesma partição sugerida pelo MPP com correlação entre blocos, porém com uma probabilidade bem menor, $P(\rho = \{[1 - 71], [72 - 86], [87 - 205]\}) = 0,179$.

Tabela 5.3: Partições mais prováveis *a posteriori* - Cromossomo 11 - Coriell.

Modelo	Partição
MPP com correlação entre blocos	$P(\rho = \{[1 - 71], [72 - 86], [87 - 205]\}) = 0,722$ $P(\rho = \{[1 - 205]\}) = 0,132$
MPP com correlação intra blocos	$P(\rho = \{[1 - 71], [72 - 204]\}) = 0,658$ $P(\rho = \{[1 - 204]\}) = 0,291$
MPP	$P(\rho = \{[1 - 71], [72 - 77], [78], [79 - 83], [84 - 85], [86], [87 - 205]\}) = 0,107$ $P(\rho = \{[1 - 71], [72 - 77], [78], [79 - 84], [85], [86], [87 - 205]\}) = 0,020$

Tabela 5.4: Resumo das distribuições *a posteriori* - Cromossomo 11 - Coriell.

Modelo	Parâmetro	Média	Des. Padrão	Intervalo HPD - 95%
MPP com correlação entre blocos	σ_μ^2	3,675	2,711	[1,003; 8,347]
	σ_y^2	0,0120	0,010	[0,006; 0,039]
	p	0,991	0,008	[0,976; 1,000]
	B	2,825	0,837	[1; 4]
MPP com correlação intra blocos	σ_μ^2	0,027	0,004	[0,020; 0,035]
	σ_y^2	0,022	0,003	[0,017; 0,027]
	p	0,992	0,007	[0,979; 0,999]
	B	1,752	0,522	[1; 2]

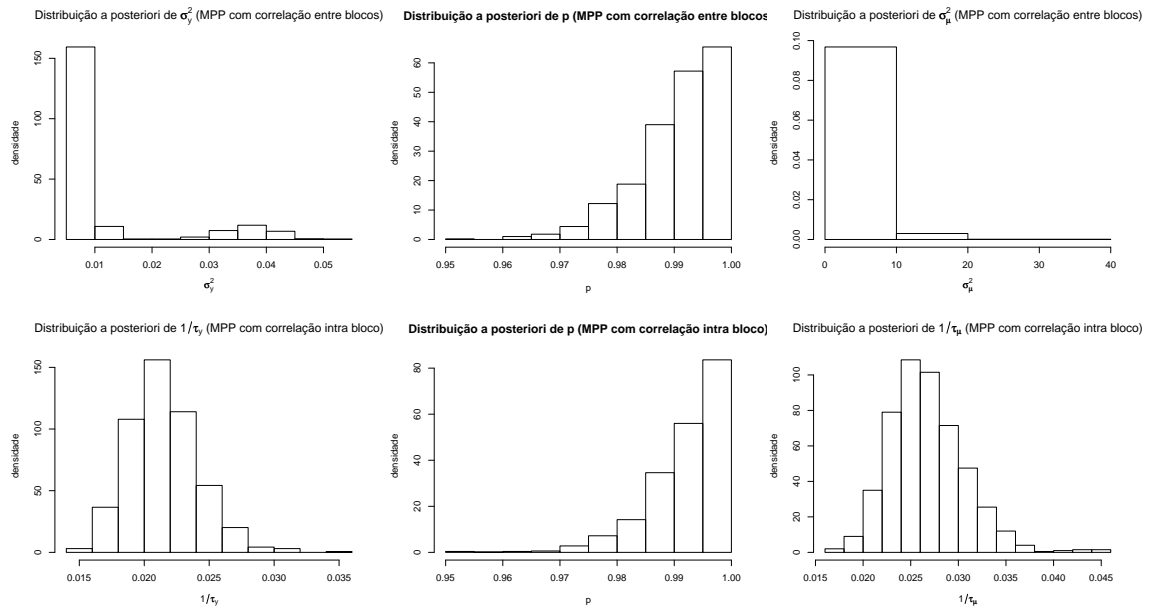


Figura 5.6: Distribuições *a posteriori* dos parâmetros σ_y^2 , p e σ_μ^2 para o cromossomo 11 - coriell.

Nesse trabalho introduziu-se um modelo tipo partição produto o qual é inovador por incluir em sua estrutura a correlação entre os blocos temporais. Este modelo foi considerado para identificar blocos temporais na média de observações normalmente distribuídas. Para amostrar da partição ρ foi proposto um algoritmo MCMCSR, uma vez que o espaço paramétrico não tem dimensão fixa. Vários cenários foram utilizados para avaliar o desempenho do modelo proposto o qual foi comparado a outros dois modelos presentes na literatura, a saber, o MPP proposto por Barry & Hartigan (1993) e a sua extensão proposta por Monteiro, Loschi & Assunção (2008).

Considerando os resultados obtidos para os cenários gerados o modelo com correlação entre blocos mostrou ser mais eficiente, na maioria dos casos, no que tange à estimação das médias e variância da distribuição normal, bem como para o número de mudanças ocorridas ao longo do tempo e suas posições. Constatou-se também que todos os modelos estudados não apresentaram bom desempenho quando a série é decrescente ou apresenta uma maior variabilidade.

Em linhas gerais, a luz dos resultados obtidos, o modelo proposto apresentou, na maioria dos casos abordados, resultados aparentemente um pouco mais satisfatórios que o MPP e o MPP com cor-

relação intra blocos em termos de estimativas da partição e do número de blocos. Outra vantagem da abordagem proposta é que quando comparada ao MPP com correlação intra blocos, na maioria dos casos, apresentou estimativas pontuais um pouco mais precisas e menores intervalos HPD para a variância das observações. Observou-se, também, que mesmo em um tamanho de amostra relativamente pequeno, 60 observações, o modelo proposto obteve estimativas satisfatórias para os parâmetros de interesse. Além disso, verificou-se que o modelo é bastante sensível quando a série se comporta de maneira alternada como descrito no Cenário 5. No entanto, cabe salientar que todos os cenários considerados na análise favoreceram o MPP original e o modelo proposto, uma vez que sempre se assumiu médias comuns dentro dos blocos. Além disso, apenas uma amostra foi considerada em cada cenário. Portanto, um estudo Monte Carlo seria útil para avaliar o quão distante estariam as estimativas obtidas por pesquisadores que possuem a mesma informação *a priori*, porém não possuem a mesma informação amostral.

Como proposta para trabalhos futuros sugere-se avaliar o desempenho do modelo proposto para dados de contagem. Outra proposta interessante seria tentar incorporar a estrutura de correlação dentro dos blocos utilizada por Monteiro, Loschi e Assunção (2008) ao modelo proposto nessa dissertação, além de estender o modelo proposto para o caso onde a variância não seja considerada constante, visando melhorar as estimativas quando a série apresenta uma maior variabilidade.

Referências Bibliográficas

- Barry, D. & Hartigan, J. A. (1992), ‘Product partition models for change point problems’, *The Annals of Statistics* **20**(1), 260–279.
- Barry, D. & Hartigan, J. A. (1993), ‘A bayesian analysis for change point problem’, *Journal of the American Statistical Association* **88**(421), 309–319.
- Crowley, E. M. (1997), ‘Product partition models for normal means’, *Journal of the American Statistical Association* **92**(437), 192–198.
- Doornik, J. & Ooms, M. (2005), An introduction to Ox, Technical report, <http://www.doornik.com/ox/OxIntro.pdf>.
- Ehlers, R. S. (2007), *Introdução à Inferência Bayesiana*, 5 edn, Disponível em <http://leg.ufpr.br/~ehlers/bayes>. Acesso em: 14/12/2009.
- Erdman, C. & Emerson, J. W. (2007), ‘bcp: An r package for performing a bayesian analysis of change point problems’, *Journal of Statistical Software* **23**(3), 1–13.
- Fearnhead, P. (2006), ‘Exact and efficient bayesian inference for multiple changepoint problems’, *Statistics and Computing* **16**(2), 203–213.
- Gamerman, D. & Lopes, H. F. (2006), *Markov chain Monte Carlo: stochastic simulation for*

- Bayesian inference*, Texts in Statistical Science, 2 edn, Chapman & Hall/CRC, Boca Raton, U.S.A.
- Green, P. J. (1995), ‘Reversible jump markov chain monte carlo computation and bayesian model determination’, *Biometrika* **82**(4), 711–732.
- Hartigan, J. A. (1990), ‘Partition models’, *Communication in Statistics - Theory & Method* **19**(8), 2745–2756.
- Hegarty, A. & Barry, D. (2008), ‘Bayesian disease mapping using product partition models’, *Statistics in Medicine* **27**, 3868–3893.
- Loschi, R. H. & 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., Cruz, F. R. B., Iglesias, P. L. & Arellano-Valle, R. B. (2003), ‘A gibbs sampling scheme to product partition model: An application to change point problems’, *Computers & Operations Research* **30**(3), 463–482.
- Monteiro, J. V. D., Loschi, R. H. & Assunção, R. (2008), ‘Product partition models with correlated parameters’. unpublished.
- Perron, P. (1990), ‘Testing for a unit root in a time series with a changing mean’, *Journal of Business & Economic Statistics* **8**(2), 153–162.
- Quintana, F. A. & Iglesias, P. L. (2003), ‘Bayesian clustering and product partition models’, *Journal of the Royal Statistical Society B* **65**(2), 557–574.
- R Development Core Team. (2009), R: A language and environment for statistical computing. R Foundation for Statistical Computing. (ISBN 3-900051-07-0), Vienna, Austria. URL <http://www.R-project.org>.*
- Tarantola, C., Consoni, G. & Dallaportas, P. (2008), ‘Bayesian clustering for row effects models’, *Journal of Statistical Planning and Inference* **138**, 2223–2235.
- Yao, Y. (1984), ‘Estimation of a noisy discrete-time step function: Bayes and empirical bayes approaches’, *The Annals of Statistics* **12**(4), 1434–1447.

Gráficos das distribuições *a posteriori*

A seguir são apresentados os gráficos das distribuições *a posteriori* obtidas via MPP com correlação entre blocos e MPP com correlação intra blocos. As distribuições *a posteriori* são referentes a análise dos cenários gerados e correspondem aos parâmetros σ_y^2 , p e σ_μ^2 .

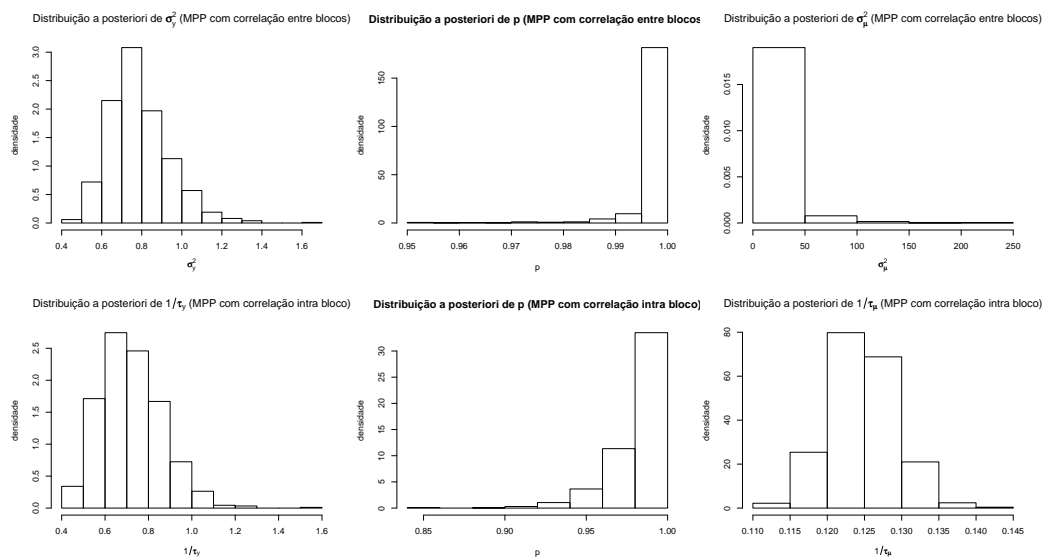


Figura A.1: Distribuições *a posteriori* do Cenário 1.

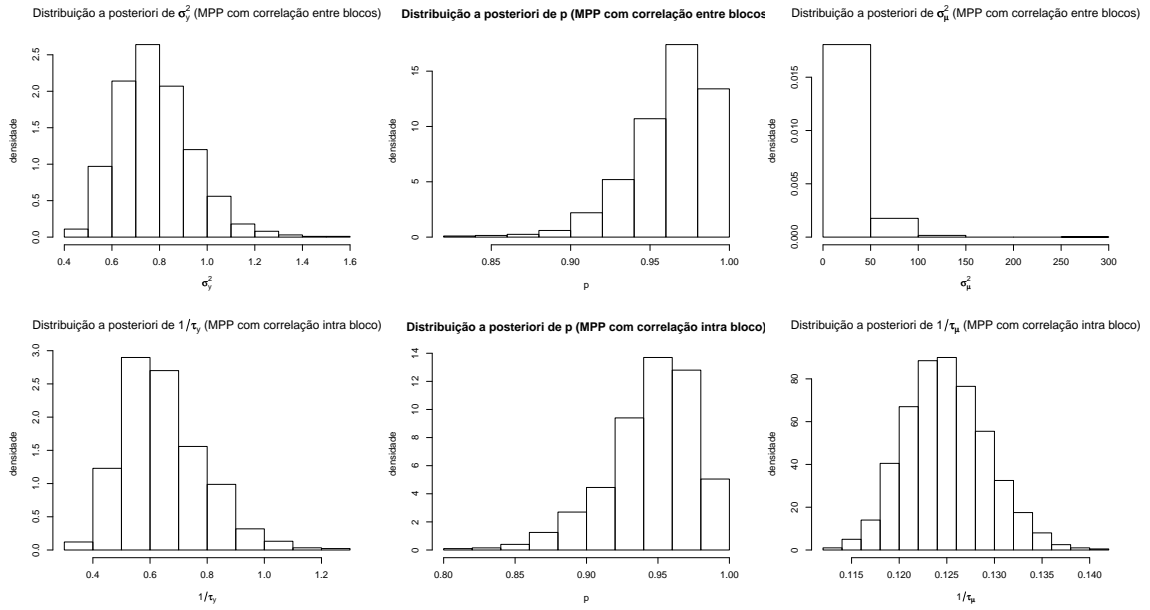


Figura A.2: Distribuições *a posteriori* do Cenário 2(a).

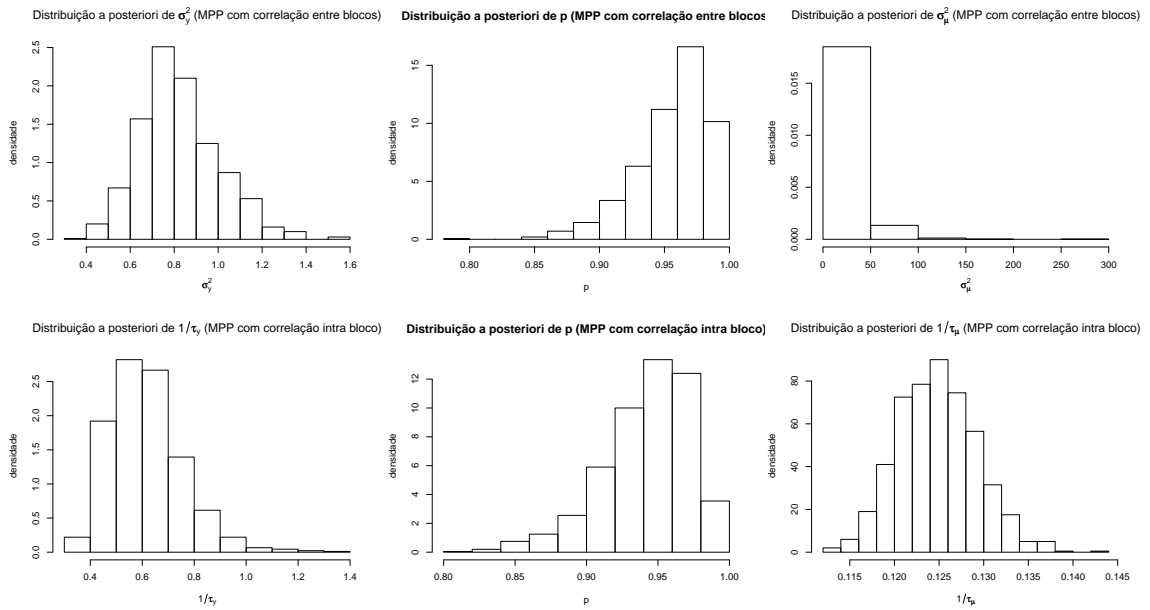


Figura A.3: Distribuições *a posteriori* do Cenário 2(b).

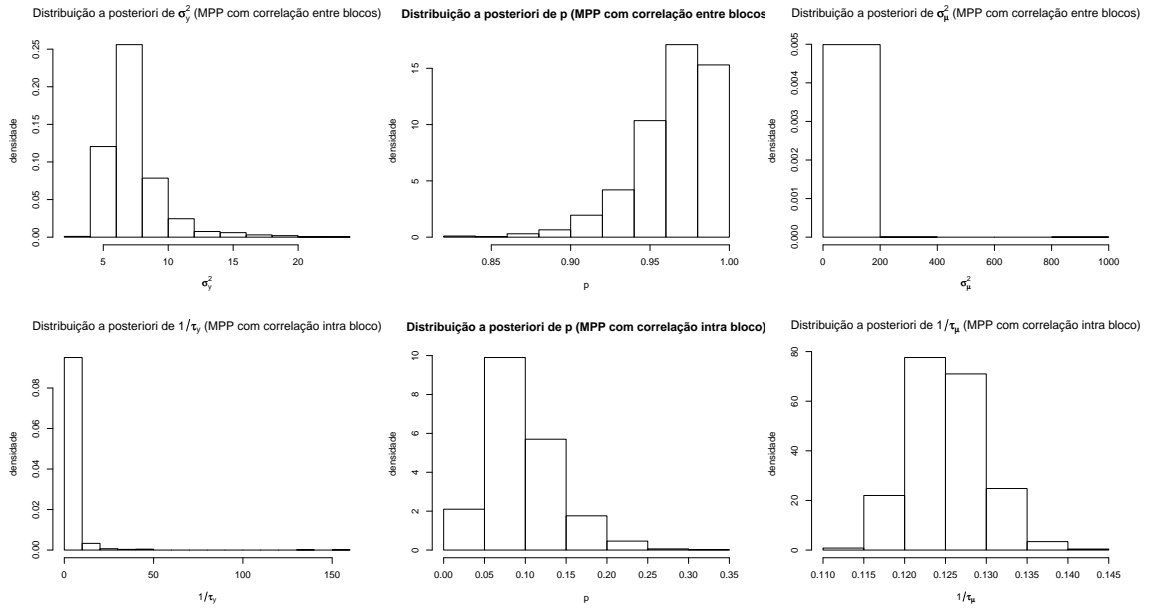


Figura A.4: Distribuições *a posteriori* do Cenário 3(a).

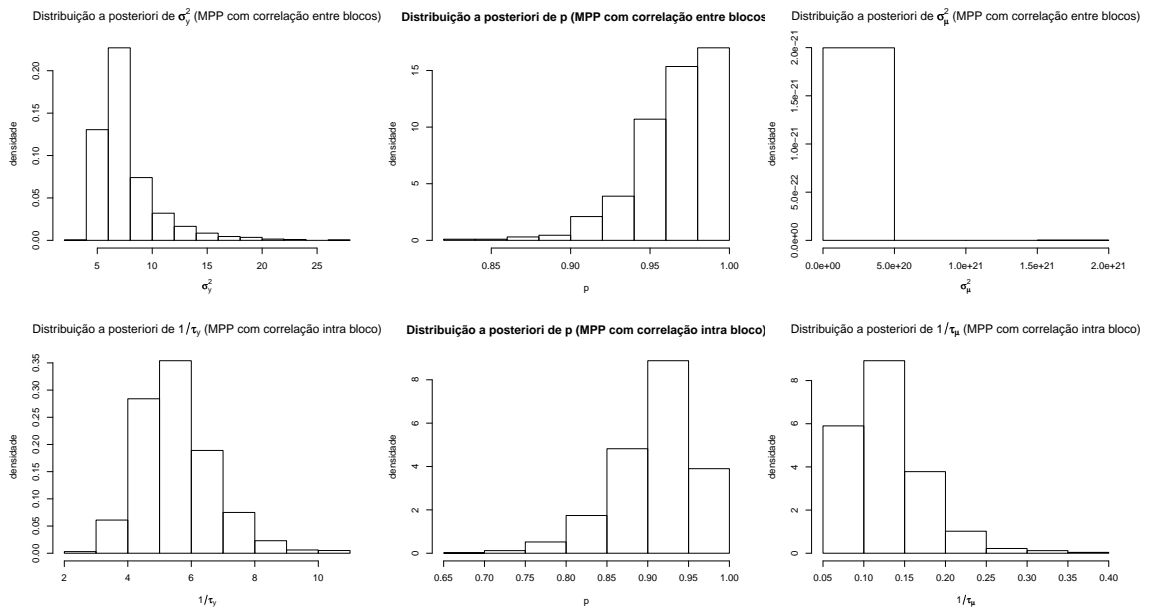


Figura A.5: Distribuições *a posteriori* do Cenário 3(b).

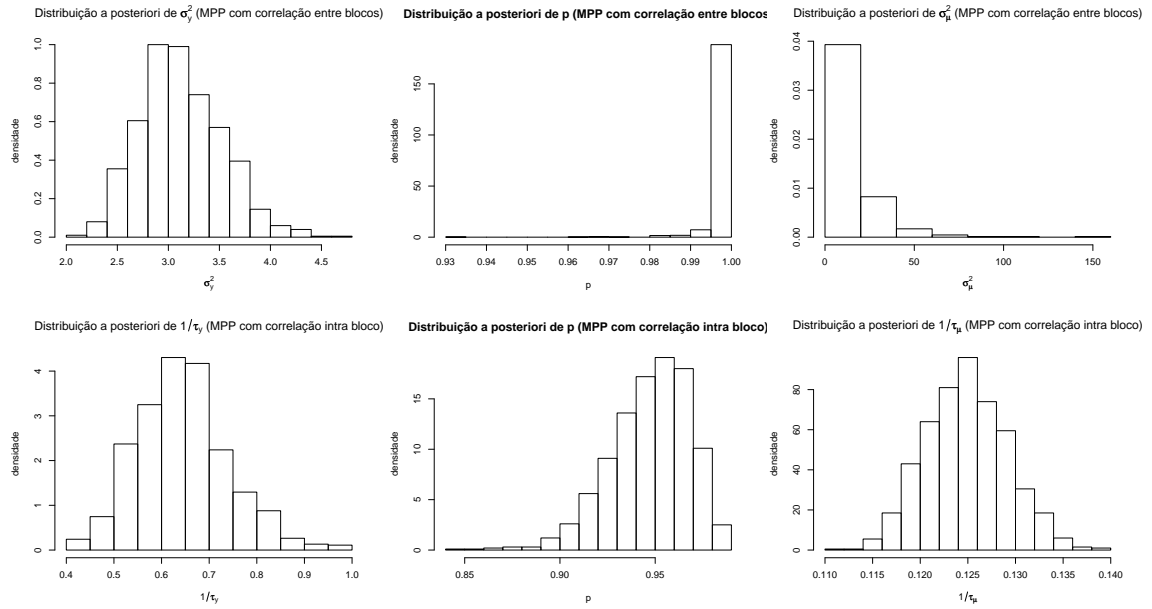


Figura A.6: Distribuições *a posteriori* do Cenário 5(a).

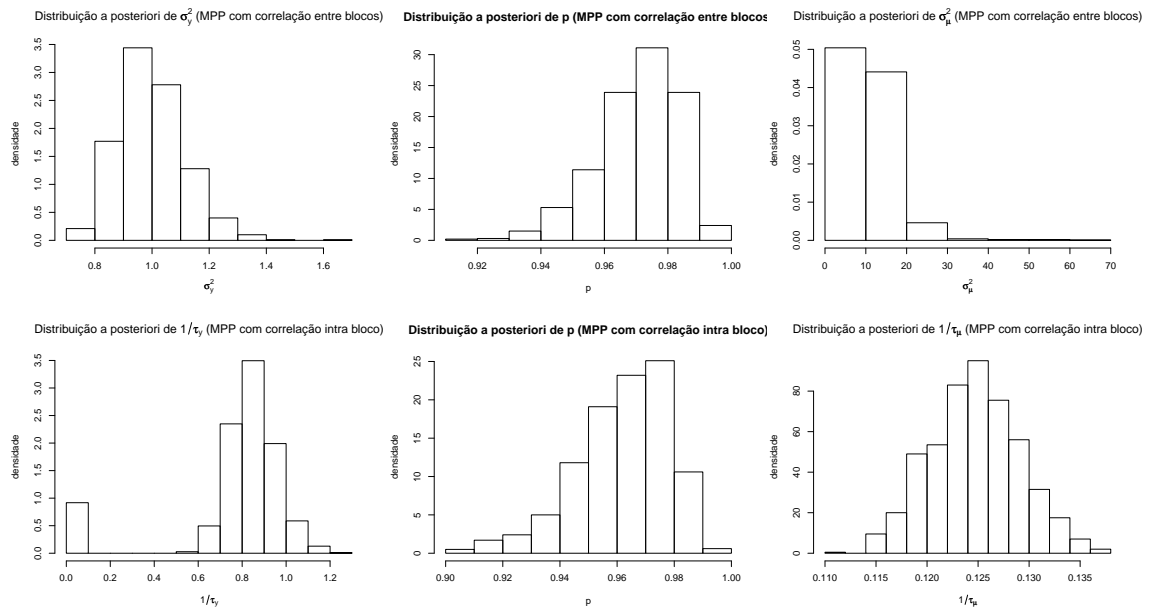


Figura A.7: Distribuições *a posteriori* do Cenário 5(b).

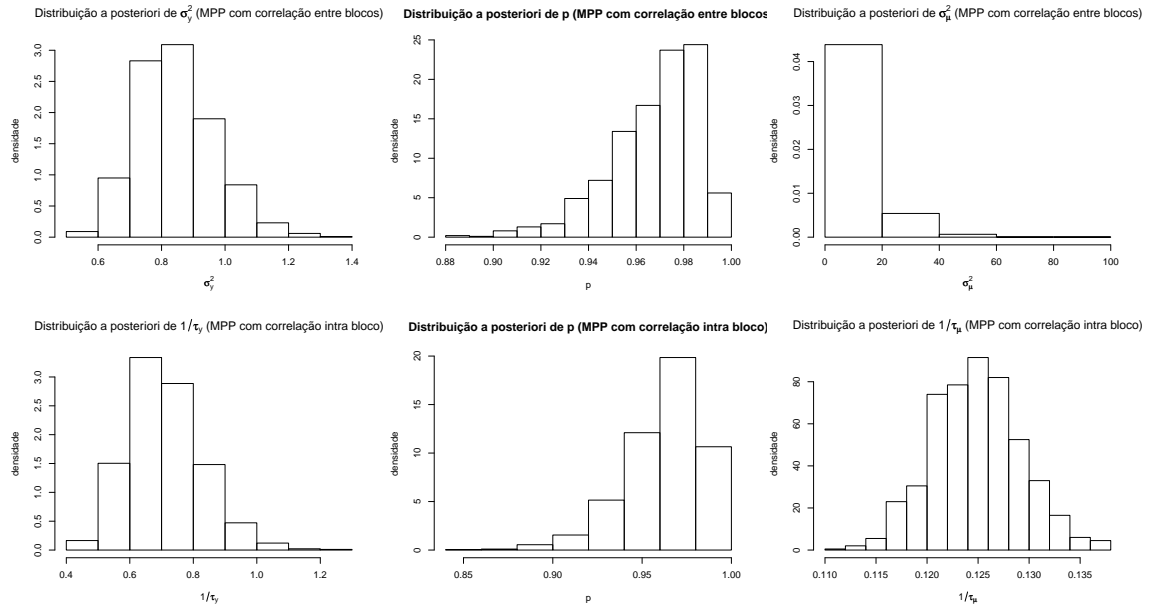


Figura A.8: Distribuições *a posteriori* do Cenário 6(a).

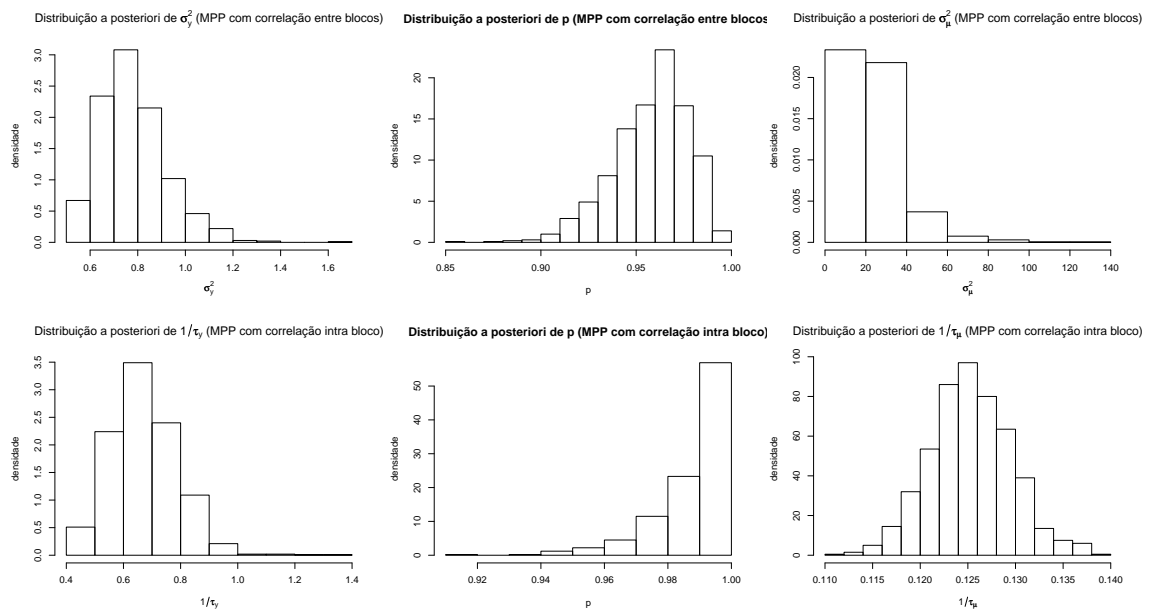


Figura A.9: Distribuições *a posteriori* do Cenário 6(b).

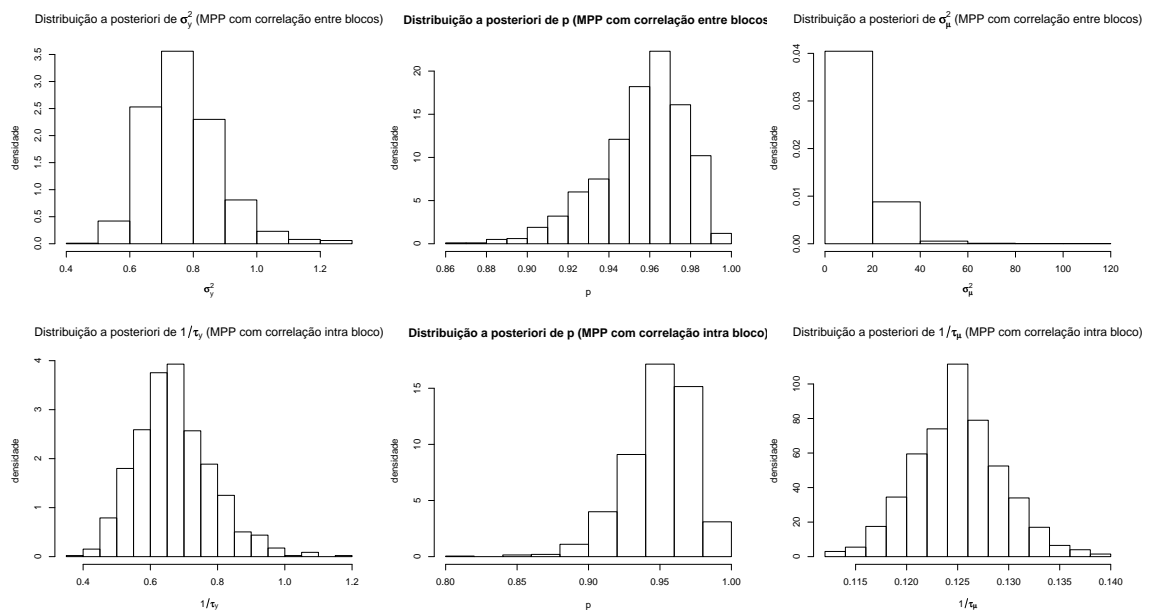


Figura A.10: Distribuições *a posteriori* do Cenário 4.

Gráficos de Convergência do Procedimento MCMC Implementado

A seguir são apresentados alguns gráficos referentes a convergência do procedimento MCMC utilizado nos cenários gerados.

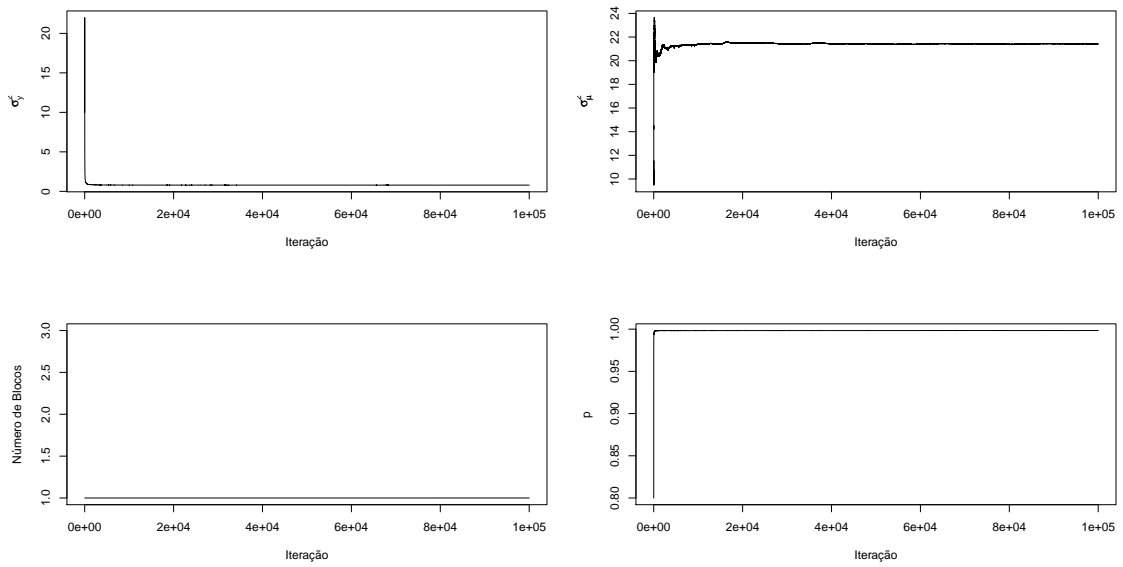


Figura B.1: Média ergódica dos parâmetros estimados via MPP com correlação entre blocos - Cenário 1.

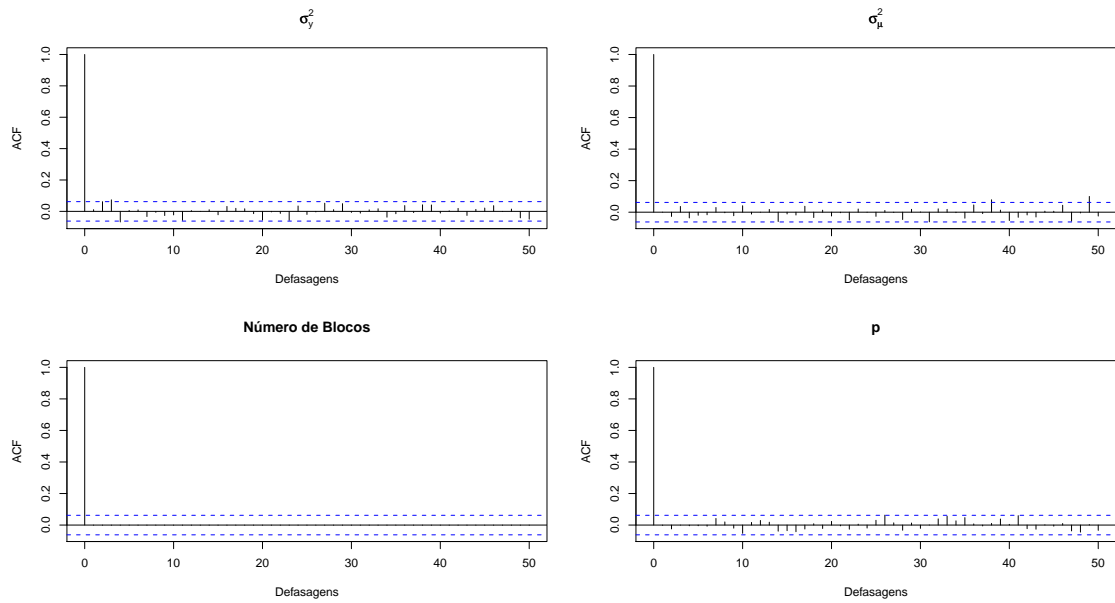


Figura B.2: Função de autocorrelação dos parâmetros estimados via MPP com correlação entre blocos - Cenário 1.

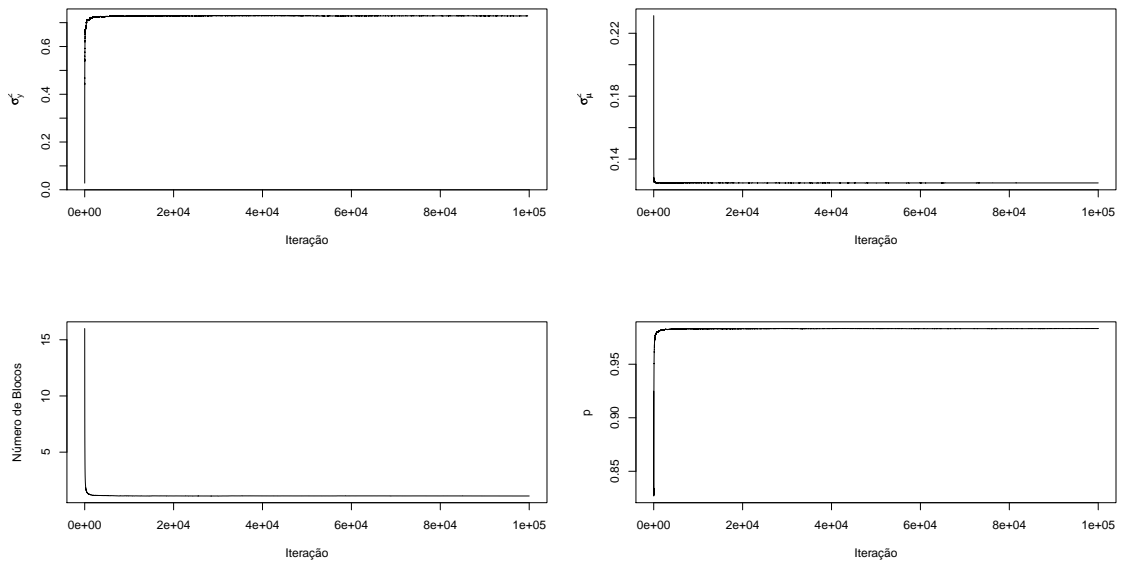


Figura B.3: Média ergódica dos parâmetros estimados via MPP com correlação intra blocos - Cenário 1.

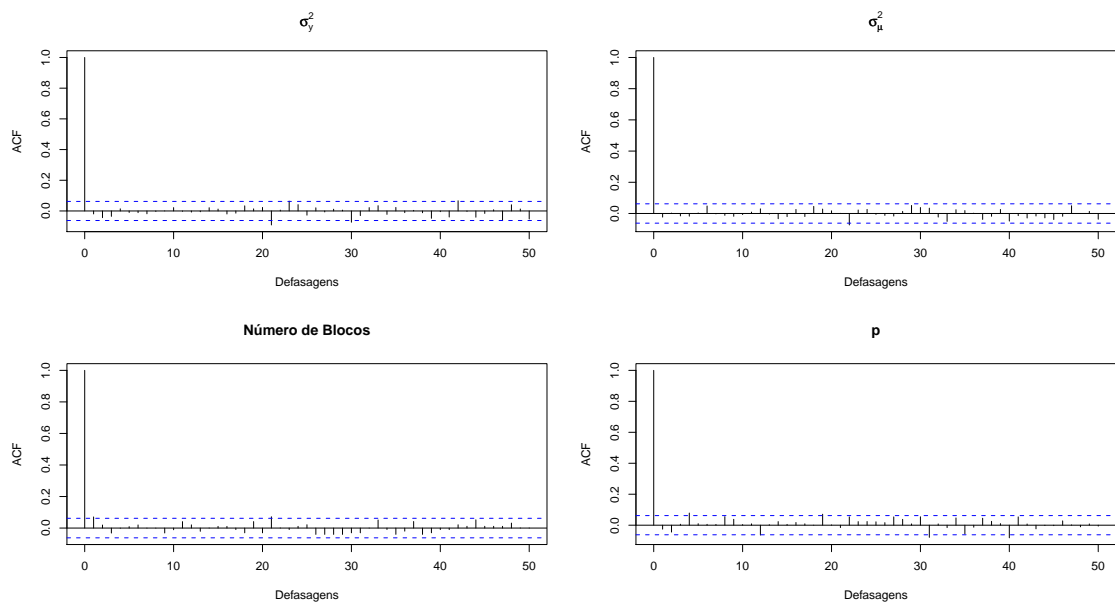


Figura B.4: Função de autocorrelação dos parâmetros estimados via MPP com correlação intra blocos - Cenário 1.

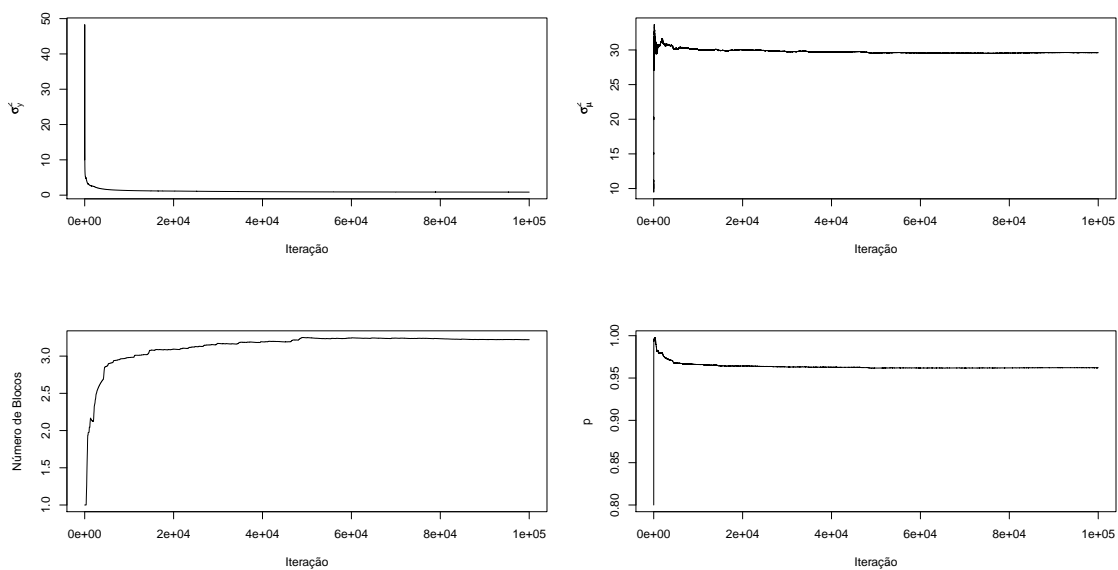


Figura B.5: Média ergódicas dos parâmetros estimados via MPP com correlação entre blocos - Cenário 2(a).

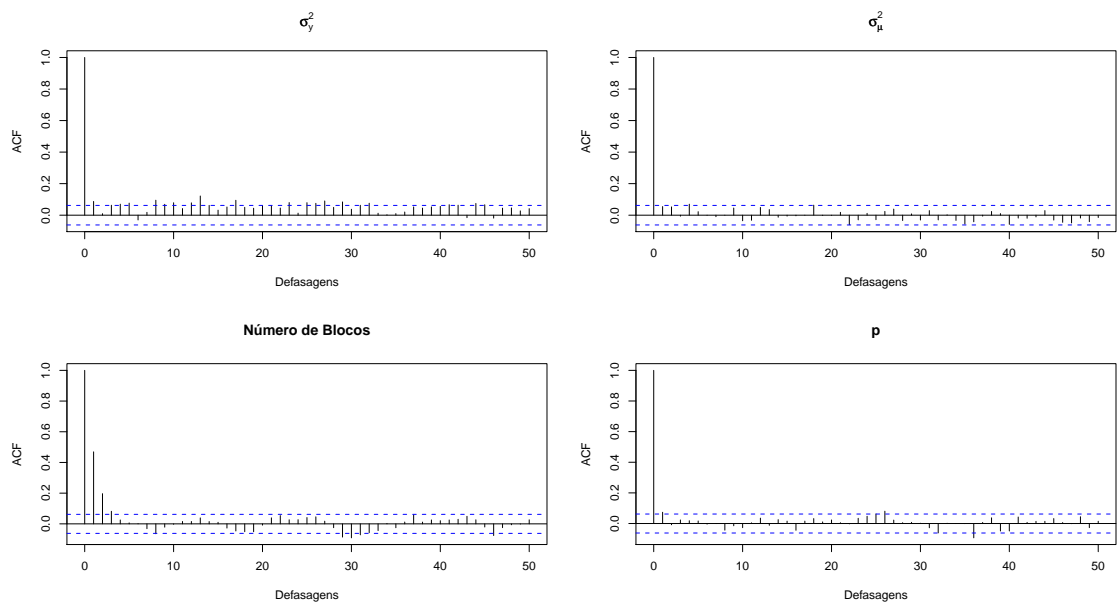


Figura B.6: Função de autocorrelação dos parâmetros estimados via MPP com correlação entre blocos - Cenário 2(a).

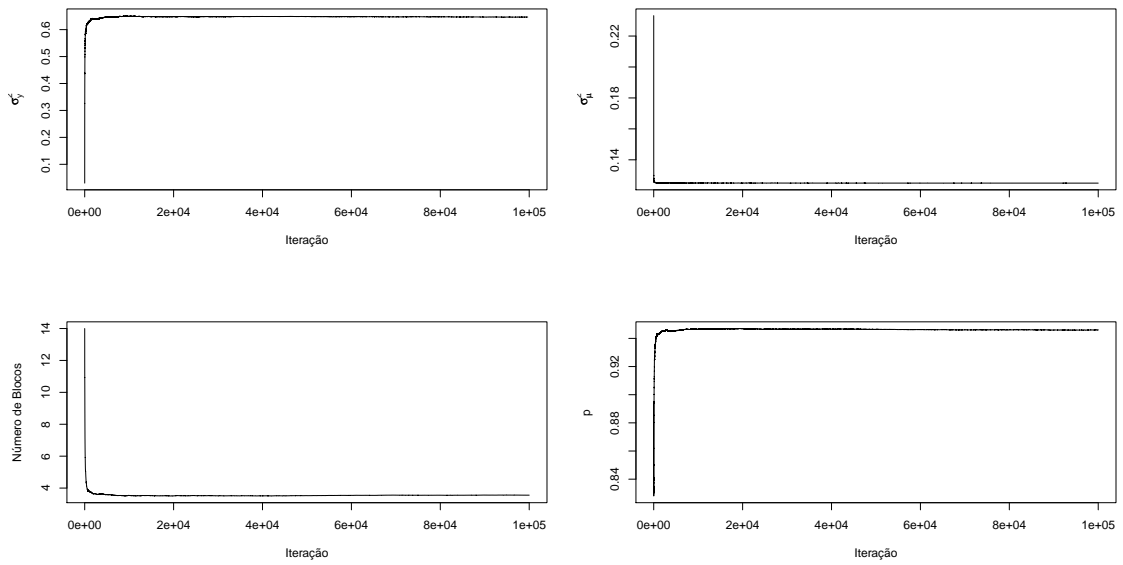


Figura B.7: Média ergódicas dos parâmetros estimados via MPP com correlação intra blocos - Cenário 2(a).

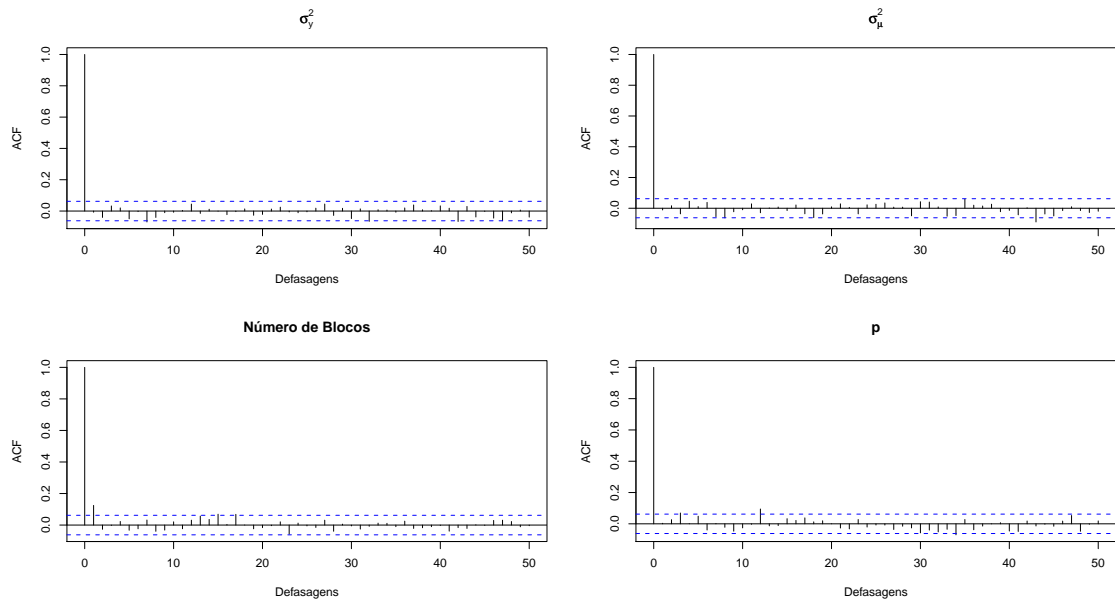


Figura B.8: Função de autocorrelação dos parâmetros estimados via MPP com correlação intra blocos - Cenário 2(a).

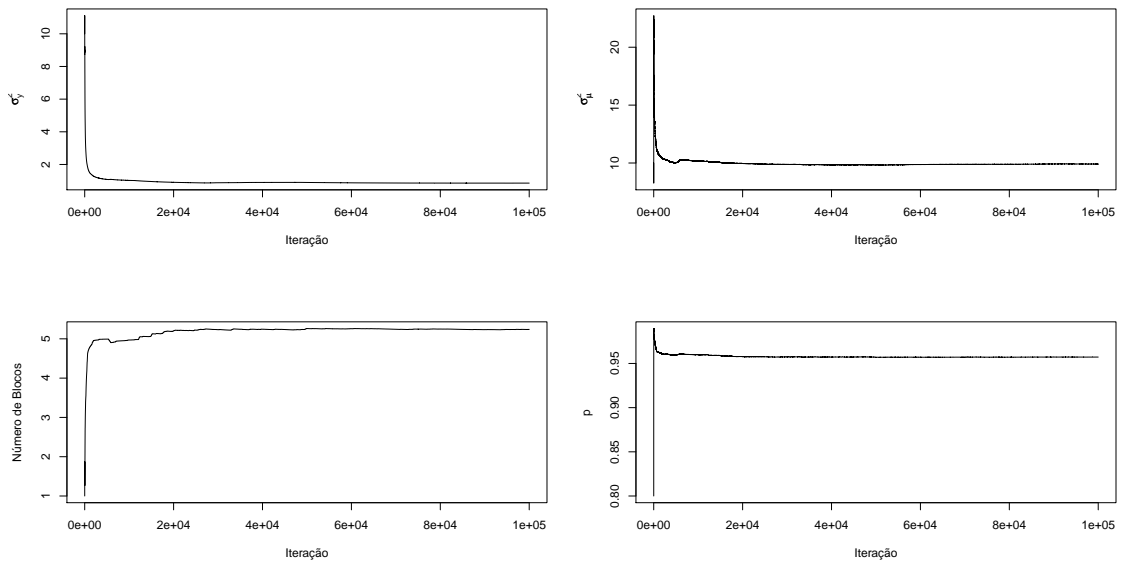


Figura B.9: Média ergódicas dos parâmetros estimados via MPP com correlação entre blocos - Cenário 6(b).

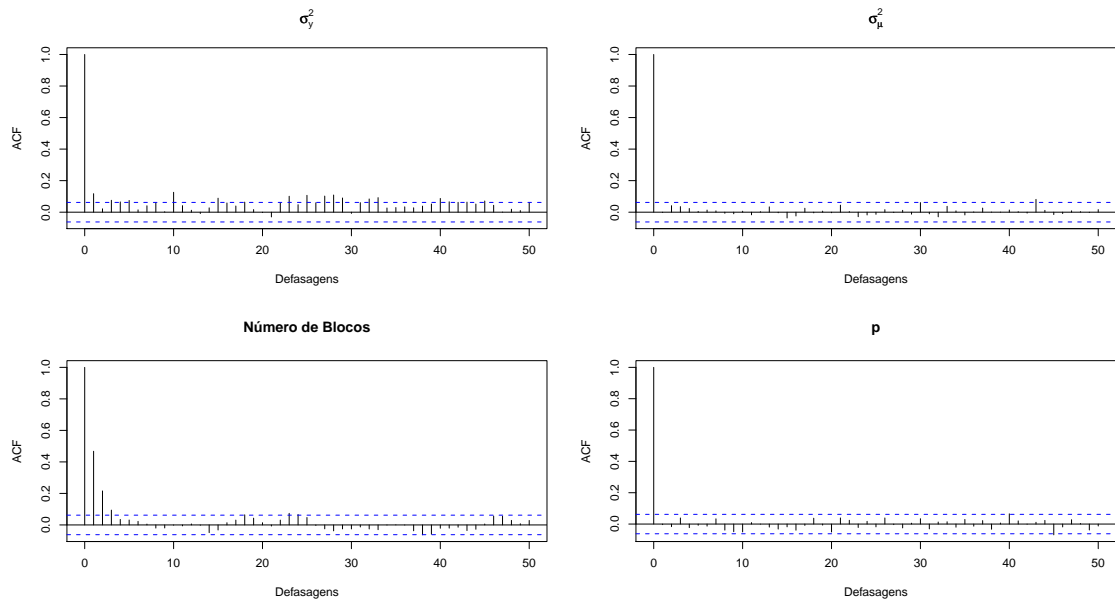


Figura B.10: Função de autocorrelação dos parâmetros estimados via MPP com correlação entre blocos - Cenário 6(b).

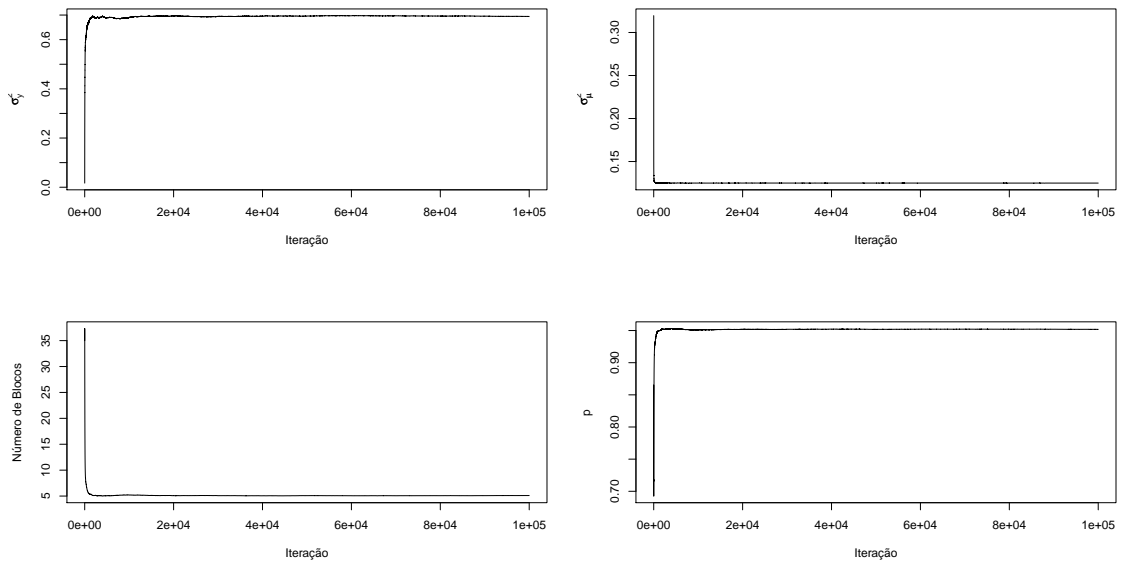


Figura B.11: Média ergódica dos parâmetros estimados via MPP com correlação intra blocos - Cenário 6(b).

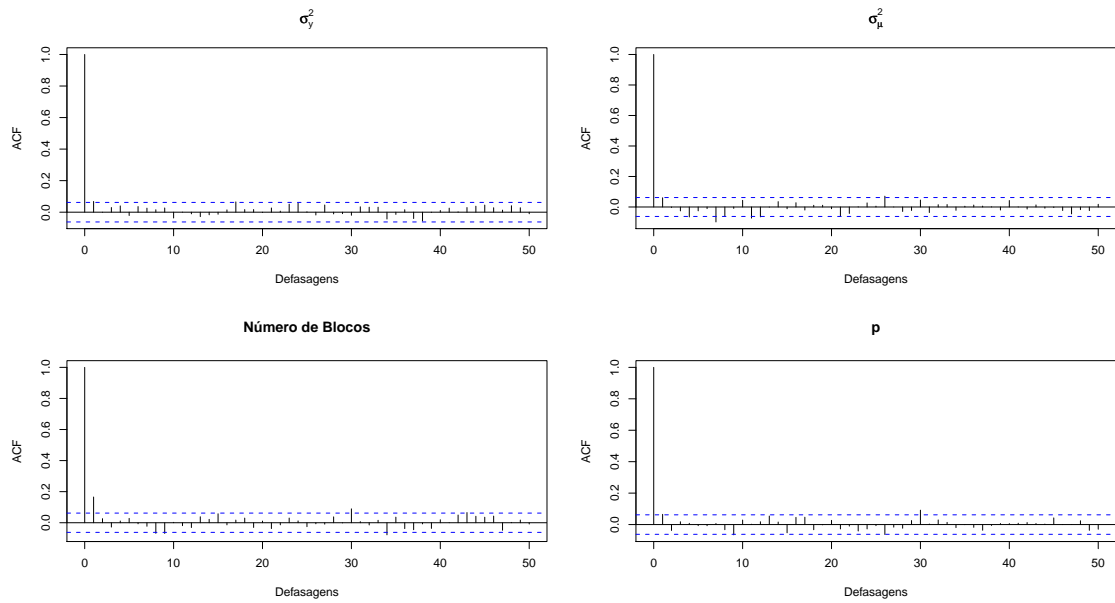


Figura B.12: Função de autocorrelação dos parâmetros estimados via MPP com correlação intra blocos - Cenário 6(b).