

---

ROBUSTEZ BAYESIANA NO MODELO PARTIÇÃO PRODUTO

JACQUELINE ALVES FERREIRA

Orientadora: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Rosângela Helena Loschi

Area de Concentração: Estatística Bayesiana

Tese de Doutorado - Doutorado em Estatística  
Universidade Federal de Minas Gerais

Belo Horizonte, Março de 2015

---

Universidade Federal de Minas Gerais  
Tese de Doutorado - Doutorado em Estatística

---

(Data)

Nós recomendamos que a tese de doutorado de autoria de **Jacqueline Alves Ferreira** intitulada ROBUSTEZ BAYESIANA NO MODELO PARTIÇÃO PRODUTO seja aceita como cumprimento parcial dos requerimentos para o grau de Doutora em Estatística.

---

Coordenador do programa de Pós-Graduação em Estatística

Banca examinadora:

---

Dr<sup>a</sup>. Rosângela Helena Loschi

Orientadora

---

Prof. Dr. Flávio Bambirra Gonçalves (DEST/UFMG)

---

Prof. Dr. Marcelo Azevedo Costa (DEP/UFMG)

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Luciana Graziela de Godoi (DE-CCE/UFES)

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Márcia D'Elia Branco (IME/USP)

©Copyright by  
JACQUELINE ALVES FERREIRA  
2015  
Todos os direitos reservados  
Typeset by L<sup>A</sup>T<sub>E</sub>X

## AGRADECIMENTOS

Agradeço aos meus pais, em especial a minha mãe que sempre me incentivou nessa caminhada.

Agradeço a professora Rosangela pelos ensinamentos e pelos momentos de força e compreensão quando as coisas pareciam, *a priori*, conturbadas e sem solução.

Agradeço a paciência e as valiosas sugestões das professoras Luciana e Márcia que sem dúvida enriqueceram esse trabalho.

Agradeço também ao professor Marcelo pela parceria no artigo publicado e prontidão em partilhar seus conhecimentos.

Agradeço a todos colegas de pós-graduação pelos momentos de descontração e discussão de ideias. Agradeço em especial ao amigo Gustavo e a amiga Márcia que sempre estiveram presentes.

Agradeço a Rogéria e a Rosi pela excelência no desempenho de suas funções facilitando a nossa vida acadêmica.

Agradeço em especial, a Patrícia pelos momentos de companheirismo e pelas palavras positivas nos momentos difíceis.

Por fim, agradeço a CAPES e a FAPEMIG pelo apoio financeiro.

## RESUMO

O Modelo Partição Produto proposto por Barry e Hartigan (1992, 1993) tem sido explorado na literatura por diversos autores. Particularmente, este modelo tem sido utilizado para a identificação de múltiplos pontos de mudança em dados sequencialmente observados e permite inferir sobre o número de pontos de mudanças, sobre a partição induzida por estes pontos a qual indica a posição onde as mudanças ocorreram e sobre os parâmetros que indexam a distribuição amostral. Apesar de muito explorado na literatura, não existe um estudo de robustez no modelo partição produto que investigue os impactos nas inferências quando perturbações são feitas no modelo. Neste trabalho, o interesse está em identificar mudanças na média populacional quando a variância é fixa e conhecida, e serão consideradas apenas escolhas do modelo em classes de distribuições contaminadas de forma multiplicativa, especificamente, considerar-se-á a classe de distribuições normal assimétrica introduzida por Azzalini (1985). Diferentemente do que é feito na literatura aqui será assumindo dois tipos de contaminação, a contaminação na distribuição *a priori* para as médias dos blocos e a contaminação na função de verossimilhança dos dados. O objetivo é investigar o comportamento das estimativas *a posteriori* para os parâmetros de interesse no modelo partição produto, a saber, o número de pontos de mudanças e suas posições, a probabilidade de cada instante ser um ponto de mudança e a média populacional. Este estudo é feito através de uma análise de sensibilidade informal e da análise de robustez global. Neste último caso, algumas medidas para avaliar a robustez global são introduzidas. Outra contribuição deste trabalho é investigar a estrutura de correlação imposta pelo Modelo Partição Produto com correlação entre os “*clusters*” proposto por Ferreira (2010) e também realizar um estudo Monte Carlo comparando tal modelo com o modelo partição produto original em diversos cenários. Análise de dados reais são utilizadas para comparar o desempenho do modelo proposto com o modelo partição produto introduzidos por Barry e Hartigan (1993) e sua extensão proposta por Monteiro et al (2011).

## ABSTRACT

The Product Partition Model (PPM) proposed by Barry and Hartigan (1992-1993) has been explored in the literature by many authors. In particular, this model has been used to identify multiple change points in data sequentially observed allowing concluding about their number, the partition induced by these points that indicates the position in which the change has occurred, and about the parameters that index the sampling distribution. Although much explored in the literature, there is no robustness study in the partition product model that investigates the effects on inferences, when contaminations are made in the model (prior or likelihood). This paper focuses in identifying changes along the time in the population mean, whenever the model belongs to a contaminated class of distributions of multiplicative form. Specifically, the class of regular asymmetric distributions introduced by Azzalini (1985) will be considered. Unlike the literature, here, two types of model contaminations will be considered, contamination of the distribution for the mean of the blocks and contamination in the likelihood function. The objective is to investigate the effect of the contamination in the posterior estimates for the parameters of interest, namely, the number of change points and their positions, the probability of each moment to be a change point, and the population means. A sensitivity informal analysis as well as a global robustness are presented. In this latter case, some measures to assess the overall robustness are introduced. Another contribution of this study is to investigate the correlation structure imposed by Product Partition Model with correlation among the clusters proposed by Ferreira (2010) and also to perform a Monte Carlo study, comparing such model with the original product partition model under different scenarios. Case studies are considered in order to compare the performance of the proposed model with that of the product partition model introduced by Barry and Hartigan (1993) and its extension proposed by Monteiro et al. (2011).

<b>Lista de Tabelas</b>		<b>ix</b>
<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Modelo Partição Produto para Identificação de Pontos de Mudança em Modelos Normais Assimétricos</b>	<b>7</b>
2.1	Modelo Partição Produto (MPP)	8
2.2	Métodos Computacionais para MPP	12
2.3	MPP para dados normais com médias normais	14
2.4	MPP para dados normais com médias normais assimétricas	17
2.4.1	Métodos computacionais - Modelo I	20
2.5	MPP dados normais assimétricos com médias normais	21
2.5.1	Métodos computacionais - Modelo II	24
<b>3</b>	<b>Análise de Sensibilidade Informal no MPP: Uma ilustração do que pode acontecer em relação a especificações equivocadas da distribuição <i>a priori</i> ou da distribuição amostral</b>	<b>27</b>
3.1	Sensibilidade informal na especificação <i>a priori</i>	28
3.1.1	A influência de $\lambda$ e $n$ nas estimativas produto e na variância <i>a posteriori</i> das médias	30
3.1.2	A influência de $\lambda$ e $n$ nas inferências sobre $\rho$ e $B$	42
3.1.3	A influência de $\lambda$ e $n$ nas inferências sobre $P(A_l)$	54
3.2	Sensibilidade informal na especificação da distribuição amostral	61
3.2.1	A influência de $\lambda$ e $n$ nas estimativas produto e na variância <i>a posteriori</i> das médias	62
3.2.2	A influência de $\lambda$ e $n$ nas inferências sobre $\rho$ e $B$	92
3.2.3	A influência de $\lambda$ e $n$ nas inferências sobre $P(A_l)$	112
3.3	Análise Crítica do Estudo de Sensibilidade	129

3.3.1	Estudo de sensibilidade - Modelo I	129
3.3.2	Estudo de sensibilidade - Modelo II	131
<b>4</b>	<b>Robustez Global no Modelo Partição Produto</b>	<b>133</b>
4.1	Sensibilidade Global	134
4.1.1	Análise de Sensibilidade na probabilidade de mudança, na partição e no parâmetro $B$	137
4.2	Métodos computacionais para o estudo de robustez no MPP - Modelo I e Modelo II	140
<b>5</b>	<b>Resultados da robustez global no MPP</b>	<b>142</b>
5.1	Robustez no Modelo I	142
5.1.1	Robustez nas estimativas produto	145
5.1.2	Robustez na partição e no número de blocos	154
5.1.3	Robustez na probabilidade de mudança em cada instante	174
5.1.4	Considerações do estudo de robustez global no Modelo I	179
5.2	Robustez no Modelo II	180
5.2.1	Robustez nas estimativas produto e variância <i>a posteriori</i> de $\mu_i$	181
5.2.2	Robustez na partição e no número de blocos	190
5.2.3	Robustez na probabilidade de mudança em cada instante	210
5.2.4	Considerações do estudo de robustez global no Modelo II	215
5.3	Análise Crítica do Estudo de Robustez	218
5.3.1	Robustez no Modelo I	218
5.3.2	Robustez no Modelo II	220
<b>6</b>	<b>Aplicação</b>	<b>222</b>
6.1	Série do Preço Médio do Tomate(Kg)	222
6.1.1	Aplicação - Modelo I - <i>priori</i> normal assimétrica	224
6.1.2	Aplicação - Modelo II - dados normais assimétricos	229
6.2	Série de retorno mensal do Ibovespa	232
6.2.1	Aplicação - Modelo I - <i>priori</i> normal assimétrica	233
6.2.2	Aplicação - Modelo II - dados normais assimétricos	236
<b>7</b>	<b>Considerações Finais e Trabalhos Futuros</b>	<b>241</b>
<b>A</b>	<b>Artigo Publicado</b>	<b>244</b>
A.1	Artigo: Detecting changes in time series: A product partition model with across-cluster correlation	244
<b>B</b>	<b>Demonstrações da Proposição 1 - Modelo I</b>	<b>245</b>
B.1	Demonstração da Proposição 1	246
<b>C</b>	<b>Demonstração da Proposição 2 - Modelo II</b>	<b>250</b>
C.1	Demonstração da Proposição 2	250

<b>D Gráficos da sensibilidade informal na verossimilhança</b>	<b>255</b>
D.1 Sensibilidade informal nas Estimativas produto	255
D.1.1 A influência de $\lambda$ e $n$ nas estimativas produto e na variância <i>a posteriori</i> das médias	255
D.1.2 A influência de $\lambda$ e $n$ nas inferências sobre $\rho$ e $B$	272
D.1.3 A influência de $\lambda$ e $n$ nas inferências sobre $P(A_l)$	278
D.1.4 Gráficos do estudo de sensibilidade informal - Modelo II: robustez na média e no instante da mudança	286
<b>Referências Bibliográficas</b>	<b>323</b>

---

## Lista de Tabelas

---

3.1	Cenários avaliados	29
3.3	Cenários avaliados	62
5.1	Cenários avaliados	144
6.1	Estatísticas Descritivas - Série mensal do preço do tomate	224
6.2	Valores do DIC para a Série mensal do preço do tomate	228
6.3	Valores do DIC para a Série mensal do preço do tomate	232
6.4	Estatísticas Descritivas - Série dos retornos mensais do Ibovespa	232
6.5	Valores do DIC para a Série dos retornos mensais do Ibovespa	236
6.6	Valore do DIC para a Série dos retornos mensais do Ibovespa	240

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 do Modelo Partição Produto (MPP).

O modelo partição produto foi proposto por Hartigan (1990). Em sua versão geral, o MPP permite identificar “*clusters*” ou blocos não contíguos o que torna sua implementação um pouco mais complexa dado o grande número de partições distintas a serem visitadas numa implementação computacional. Na análise de pontos de mudança em séries temporais ou em dados sequencialmente observados faz sentido pensar em um modelo que considere apenas partições em blocos contíguos de observações. Nesta direção, Barry & Hartigan (1992) propõem um MPP em que apenas blocos contíguos são considerados diminuindo o número de partições possíveis. Barry & Hartigan (1993) aplicam o MPP proposto por Barry & Hartigan (1992) para identificação de múltiplos pontos de mudança na média da distribuição normal com variância constante e desconhecida e introduzem uma forma de implementá-lo computacionalmente, via amostrador de Gibbs.

No seu artigo original em 1990, Hartigan faz duas suposições para construir o MPP. A primeira

afirma que as observações dentro de um mesmo “*cluster*” (ou bloco) são identicamente distribuídas, e a segunda assume que os parâmetros em diferentes “*clusters*” são independentes, dada a partição.

As suposições feitas em Hartigan (1990) são mantidas em Barry & Hartigan (1992, 1993), esses autores também consideram uma classe do MPP em que a partição é entendida como os blocos contíguos formados pelos índices dos pontos de mudança, ou seja, se uma série possui 60 observações e dois pontos de mudança, nas posições, 21 e 41, então a partição  $\rho = \{0, 20, 40\}$  é dada por  $\{[1 - 20], [21 - 40], [41 - 60]\}$ , mas 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). Já 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 tais áreas de acordo com os seus respectivos riscos relativos.

O modelo proposto por Hartigan (1990) é bastante explorado na literatura, dentre alguns autores pode-se citar Quintana & Iglesias (2003) que provam, entre outras coisa, que o processo de Dirichlet é um caso particular do MPP proposto por Hartigan (1990) o que permite usar alguns algoritmos já implementados para os processo de Dirichlet na implementação do MPP. Quintana & Iglesias (2003) apresentam uma formulação teórica do MPP para identificação de *outliers*. Tarantola et al. (2008) utiliza o modelo partição produto em tabelas de contingências. Zantedeschi et al. (2011) propõem um modelo partição produto dinâmico para prever movimentos na estrutura da taxa de juros e utilizam o algoritmo Monte Carlo sequencial.

A popularidade do PPM é justificada pela flexibilidade proporcionada por esse modelo para analisar o problema de ponto de mudança. Outras extensões desse modelo aparecem, por exemplo, em 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 um filtro de partículas para encontrar as distribuições *a posteriori* relacionadas ao MPP. Demarqui

et al. (2008, 2012) utilizam o MPP para estimar a grade do modelo exponencial por partes que é muito utilizado em análise de sobrevivência. As suposições originais do MPP foram flexibilizadas, posteriormente, no modelo introduzido por Monteiro et al. (2011) que assumem que as observações dentro dos blocos são correlacionadas e não identicamente distribuídas.

Um tópico ainda não explorado no contexto de identificação de pontos de mudança utilizando o MPP é o estudo de robustez bayesiana no modelo partição produto, logo o estudo que pretende-se desenvolver nesse trabalho é inédito nesse contexto, pois não existe um estudo na literatura que aborde de forma detalhada a robustez bayesiana no MPP.

O termo robustez é amplamente empregado, por exemplo, no contexto onde as estimativas dos parâmetros não são afetadas por observações discrepantes. No entanto, neste trabalho, o foco é estudar robustez bayesiana avaliando o comportamento das estimativas *a posteriori* quando, por exemplo, as distribuições *a priori* ou a função de verossimilhança são escolhidas numa certa classe de distribuições de probabilidade. Segundo Berger et al. (2000), o início dos anos 90 foi a era dourada da robustez bayesiana, pois foi quando esta área de estudo ganhou visibilidade com workshops internacionais, conferências e a publicação do livro Bayesian Robustness de Berger et al. (1996), além da publicação de diversos artigos.

Berger et al. (2000) afirma que a necessidade de se estudar a robustez bayesiana nasceu em virtude dos questionamentos sobre a escolha da distribuição *a priori* e explica que há dois tipos principais de robustez bayesiana: a robustez global e robustez local. A robustez local analisa como mudanças infinitesimais na distribuição *a priori* podem influenciar na distribuição *a posteriori*. No contexto de robustez global, o foco está, por exemplo, em se considerar uma classe de distribuições *a priori* e o objetivo é avaliar o quanto a distribuição *a posteriori* dos parâmetros se modifica para escolhas de distribuições *a priori* em tal classe de distribuições. Exemplo de autores que discutem a robustez bayesiana são: Gustafson (1995) que busca quantificar o efeito de pequenas mudanças para a distribuição *a priori* ao longo de um espaço de parâmetros  $k$ -dimensional e Fortini & Ruggeri (1994) que propõem o uso da função de concentração para avaliar robustez bayesiana.

A década de 90 foi um marco para o estudo de robustez, mas na última década vários estudos continuam sendo realizados como, por exemplo, van der Linde (2004) que estuda a influencia local na distribuição *a posteriori* sob uma classe de modelos contaminados de forma multiplicativa. Godoi & Branco (2014) propõem medidas de robustez global quando a distribuição *a priori* pertence a classe de distribuições normais assimétricas. Ruggiere & Sivaganesan (2000) que propõem uma medida de sensibilidade global que possui uma interpretação comum independente do contexto do problema.

Voltando a década de 90, Carota e Ruggeri (1994), pela primeira vez, fazem um estudo de robustez Bayesiana em modelos que objetivam fornecer inferências sobre partições do espaço paramétrico. Em seu trabalho, avaliam como as contaminações nas distribuições *a priori* sobre as partições influenciam nas inferências *a posteriori* sobre os parâmetros de interesse como, por exemplo, a média da característica em estudo. No entanto, esses autores não levam em consideração a partição do espaço amostral ao longo do tempo.

Neste trabalho, o principal foco está no estudo de robustez no MPP. Restringir-se-á à classe de modelos partição produto definidas por Barry & Hartigan (1992) onde apenas blocos ou “clusters” contíguos são permitidos e assume-se independência entre os blocos. Um dos principais objetivos quando ajusta-se este tipo de modelo é inferir sobre o número de blocos ( $B$ ) ou, equivalentemente, o número de pontos de mudanças que ocorreram, a partição  $\rho$  do eixo do tempo induzida pela posição dos pontos de mudanças, além dos parâmetros  $\mu$  que indexam a distribuição amostral ajustada aos dados. O foco do trabalho será o estudo da robustez global e a sugestão de algumas medidas para avaliar a influência das escolhas do modelo e da distribuição *a priori* em todos estes parâmetros. Diferentemente do estudo feito em Carota e Ruggeri (1994), aqui será avaliado como as estimativas *a posteriori* para  $B$ ,  $\rho$ ,  $\mu$  e a probabilidade de mudança em cada instante são afetadas quando a distribuição *a priori* para  $\mu$  é escolhida em uma certa classe de distribuições contaminada. Em outra direção, também será realizado um estudo de robustez quando a distribuição amostral é contaminada. Este trabalho se restringe a estudos com classes de distribuições contaminadas multiplicativamente. Essa classe de distribuições também foi considerada em Godoi & Branco (2014), onde se argumenta que a família de distribuições normal assimétrica, introduzida por Azzalini (1985), pertence a tal classe.

O estudo será feito considerando o MPP para a identificação de pontos de mudança na média de dados que são sequencialmente observados com variância comum e conhecida. Primeiramente, considera-se que a distribuição amostral é a distribuição normal e é assumido que a distribuição *a priori* para a média pertence à classe de distribuições normal assimétrica (Azzalini, 1985). Em um segundo estudo, é assumido que a distribuição amostral é a distribuição normal assimétrica e, neste caso, assume-se que a distribuição normal descreve a incerteza sobre o parâmetro de locação. Em ambos os casos, deseja-se investigar a robustez global na locação da distribuição amostral e, para este fim, considera-se como operadores a média *a posteriori*. O interesse principal está em avaliar a influência dessas escolhas na inferência sobre o número de blocos ( $B$ ), sobre as estimativas produto  $\mu_i$ , sobre a probabilidade de cada instante ser um ponto de mudança,  $P(A_t)$ , e sobre a partição verdadeira  $\rho$ .

Um estudo para identificar um único ponto de mudança em dados com distribuição normal assimétrica foi feito por Arellano-Valle et al. (2013). Nesta tese também estende-se tal trabalho por considerar a identificação de múltiplos pontos de mudança nesse mesmo contexto. Adicionalmente, também assume-se que os dados são normalmente distribuídos e uma distribuição normal assimétrica descreve o comportamento da média em cada bloco. Nessa direção, foi realizado um estudo de simulação visando mostrar como mudanças na distribuição *a priori* influenciam as estimativas *a posteriori* no modelo partição produto.

A justificativa para se trabalhar com uma classe de distribuições assimétricas é que a suposição de simetria em torno da média pode ser muito forte em algumas situações reais, como por exemplo, a modelagem de dados de mercado financeiro, mais especificamente, os retornos de ações em mercados emergentes que são bastante influenciados por eventos internacionais.

Além disso, o primeiro ano do doutorado foi dedicado a estender os resultados obtidos no mestrado. Realizou-se um estudo Monte Carlo para o modelo partição produto com correlação entre blocos proposto em 2010 na dissertação de mestrado comparando-o com o MPP original proposto por Barry & Hatingan (1993). A estrutura de correlação entre os “*clusters*” imposta pelo modelo foi

estudada mais detalhadamente. Outro resultado importante, foi comparar através de estudo de casos o modelo proposto, o MPP e o modelo partição produto com correlação intra blocos proposto por Monteiro et al. (2011). Neste caso, um estudo de sensibilidade a respeito da escolha da distribuição *a priori* para as médias do bloco também é feito. Os resultados podem ser encontrados no artigo em anexo no final do texto.

O presente trabalho está organizado como descrito a seguir. O Capítulo 2 apresenta o modelo partição produto na sua forma original e contemplando a classe de distribuição normal assimétrica. O Capítulo 3 ilustra por meio de uma análise de sensibilidade informal o que pode ocorrer se o modelo é especificado erroneamente. No Capítulo 4 são apresentadas as medidas adotadas neste trabalho para o estudo de robustez global no modelo partição produto, na sequência, o Capítulo 5 apresenta os resultados obtidos por meio das medidas propostas no Capítulo 4. Após o estudo de robustez é apresentado no Capítulo B um estudo de sensibilidade aplicado aos dados de retorno mensal do Ibovespa e ao preço médio de venda do quilo do tomate. Por fim, são apresentadas propostas de trabalhos futuros e as considerações finais deste texto.

---

### Modelo Partição Produto para Identificação de Pontos de Mudança em Modelos Normais Assimétricos

---

O modelo partição produto foi proposto por Hartigan (1990) em sua forma geral, no qual permite a identificação de agrupamentos não contíguos num conjunto de dados quando o número de *clusters* não é previamente conhecido. Posteriormente, Barry & Hartigan (1992) propõem uma modificação no MPP a qual é útil na identificação de pontos de mudança em dados sequencialmente observados, onde apenas *clusters* contíguos são permitidos, obtendo as formas fechadas das distribuições *a posteriori* para uma partição aleatória ( $\rho$ ), para o número de blocos 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 em Barry & Hartigan (1993) é a introdução de uma forma de implementar computacionalmente, via amostrador de Gibbs, o MPP apesar da dimensão do espaço paramétrico não ser inicialmente conhecida.

Como o objetivo é fazer um estudo de robustez no modelo considerado por Barry & Hartigan (1992), neste capítulo é feita uma breve revisão da metodologia proposta por Barry & Hartigan (1992 ; 1993) para identificação de pontos de mudança utilizando o modelo partição produto.

## 2.1 Modelo Partição Produto (MPP)

Sejam  $Y_1, \dots, Y_n$  observações feitas ao longo do tempo. Admita que esta série experimenta mudanças em seu comportamento induzindo uma partição dessa série em blocos ou subsequências contíguas no tempo. Para estabelecer notação, chame de  $\rho$  a partição aleatória definida pelos pontos de mudança no conjunto  $I = \{0, 1, \dots, n\}$ . Assuma que um valor particular  $\{i_0, \dots, i_b\}$  de  $\rho$  é tal que  $0 = i_0 < i_1 < \dots < i_b = n$  a qual divide a série em  $B = b$  blocos (“clusters”) contíguos com pontos finais  $i_j$ ,  $j = 1, \dots, b - 1$ . Sendo assim, um bloco (ou *cluster*) temporal é definido 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,  $i_j + 1$  é o instante onde o  $j$ -ésimo ponto de mudança ocorreu. Uma partição em  $I$ , induz uma partição na série  $Y_1, \dots, Y_n$ . Dessa forma, ao bloco  $[i_j i_{j+1}]$  está associado o bloco de observações  $\mathbf{Y}_{[i_j i_{j+1}]} = [Y_{i_j+1}, \dots, Y_{i_{j+1}}]$ . Denote por  $B$  o número de blocos ou subsequências na série, a variável aleatória  $B$  está diretamente relacionada com  $\rho$  de tal forma que se  $\rho = \{i_0, \dots, i_b\}$ ,  $B$  assume o valor  $b$ .

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, traduz o grau de similaridade que se julga existir entre as observações dentro de um mesmo bloco. Portanto, coesões são escolhas pessoais e subjetivas. 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 pelos pontos de mudança. Como pode ser visto em Barry & Hartingan (1992).

Particularmente, neste trabalho, considerar-se-á as coesões propostas por Yao (1984), estas coesões *a priori* indicam que o ponto de mudança atual só depende do ponto de mudança imediatamente anterior e os tempos de ocorrência dos pontos de mudança são identicamente distribuídos com distribuição geométrica. Desta forma, 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:

$$C_{[i_j i_{j+1}]} = \begin{cases} p(1-p)^{i_{j+1}-i_j-1}, & i_{j+1} < n; \\ (1-p)^{i_{j+1}-i_j-1}, & i_{j+1} = n, \end{cases} \quad (2.1)$$

para todo  $i_j \in I$ . As coesões *a priori* em (2.1) 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 pontos de mudanças futuros, o que pode ser razoável em alguns contextos como, por exemplo, na análise de quebras estruturais produzidas por choques econômicos ou mudança de políticas governamentais.

Considere uma sequência de parâmetros desconhecidos  $\theta_1, \dots, \theta_n$ , condicionalmente nos quais, as variáveis aleatórias  $Y_1, \dots, Y_n$  são independentes e tem densidades condicionais marginais dadas por  $f(Y_1|\theta_1), \dots, f(Y_n|\theta_n)$ , respectivamente. Dada uma partição  $\rho = \{i_0, \dots, i_b\}$ , para  $b \in \{1, 2, \dots, n\}$ , assuma que existam parâmetros comuns  $\theta_{[i_0 i_1]}, \dots, \theta_{[i_{b-1} i_b]}$ , tais que  $\theta_i = \theta_{[i_j i_{j+1}]}$ , para todo  $i_j < i \leq i_{j+1}$ ,  $j = 0, \dots, b-1$ , que são considerados independentes entre si e, também, independentes 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,  $\theta_{[i_j i_{j+1}]} = \theta_{i_{j+1}} = \dots = \theta_{i_j+1}$ . Assuma as coesões *a priori* definidas em (2.1). Barry & Hartingan (1992) definem o modelo partição produto como sendo a distribuição conjunta de  $(Y_1, \dots, Y_n, \rho)$ , dado  $p$ ,  $p \in (0, 1)$  da seguinte forma:

i) dado  $p$ , a distribuição *a priori* de  $\rho$  tem a seguinte distribuição produto:

$$\pi(\rho = \{i_0, \dots, i_b\} | p) = p^{b-1} (1-p)^{n-b}, \quad \forall b = 1, 2, \dots, n; \quad (2.2)$$

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:

$$f(y_1, \dots, y_n | \rho = \{i_0, \dots, i_b\}, p) = \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) \quad (2.3)$$

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  $\mathbf{Y}_{[i_j i_{j+1}]}$ .

Para facilitar o cálculo da partição  $\rho = \{i_0, i_1, \dots, i_b\}$  *a posteriori*, Barry & Hartigan (1993) sugerem que partição seja escrita de forma equivalente ao vetor  $\mathbf{U} = (U_{i_0} = 1, \dots, U_{i_{b-1}} = 1, U_{i_b} = 0, U_{i_{b+1}} = 1, \dots, U_{i_2} = 0, \dots, U_b = 1)$ , onde  $U_i$  é uma variável aleatória definida da seguinte forma:

$$U_i = \begin{cases} 1, & \text{se } \theta_i = \theta_{i+1} \\ 0, & \text{se } \theta_i \neq \theta_{i+1}, \end{cases} \quad (2.4)$$

$i = 1, \dots, n-1$  e os componentes de  $U_i$ 's do vetor  $\mathbf{U}$  são independentes. Assim, se  $p$  denota a probabilidade de uma mudança em um instante qualquer, segue que

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

Também, segue da relação entre  $\rho$  e  $B$ , que a distribuição *a priori* para o número de blocos é

$$P(B = b | \rho) = \binom{n-1}{b-1} p^b (1-p)^{n-b}, \quad b = 1, \dots, n.$$

Assumindo o modelo definido em (2.2) e (2.3), Barry & Hartigan (1992) mostram que a distribuição e a esperança (estimativa produto) *a posteriori* para  $\theta_k, k = 1, \dots, n$  são dadas, respectivamente, por

$$f(\theta_k | y_1, \dots, y_n) = \sum_{i_j=0}^{k-1} \sum_{i_{j+1}=i}^n r_{[i_j i_{j+1}]}^* f(\theta_k | \mathbf{y}_{[i_j i_{j+1}]}), \quad (2.5)$$

$$E(\theta_k | y_1, \dots, y_n) = \sum_{i_j=0}^{k-1} \sum_{i_{j+1}=i}^n r_{[i_j i_{j+1}]}^* E(\theta_k | \mathbf{y}_{[i_j i_{j+1}]}), \quad (2.6)$$

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* associada ao bloco  $[i_j i_{j+1}]$ , ou seja, é a probabilidade *a posteriori* de que o bloco  $[i_j i_{j+1}]$  faça parte da partição  $\rho$ .

No caso considerado nesse trabalho, a relevância *a posteriori* para o bloco  $[i_j i_{j+1}]$  é

$$r_{[i_j i_{j+1}]}^* = \frac{\tau_{[0 i_j]} C_{[i_j i_{j+1}]}^* \tau_{[i_j i_n]}}{\tau_{[i_0 i_n]}}, \quad (2.7)$$

onde  $\tau_{[i_0 i_j]} = \sum \prod_{k=1}^b C_{[i_{k-1} i_k]}^*$  em que a soma é feita sobre todo conjunto de números inteiros satisfazendo  $i_j = i_0 < i_j < \dots < i_b = i_{j+1}$ , e  $C_{[i_j i_{j+1}]}^*$  denota a coesão *a posteriori* do bloco  $[i_j i_{j+1}]$  que é dada por

$$C_{[i_j i_{j+1}]}^* = \begin{cases} f_{[i_j i_{j+1}]}(y_{[i_j i_{j+1}]})p(1-p)^{i_{j+1}-i_j-1}, & j < n; \\ f_{[i_j i_{j+1}]}(y_{[i_j i_{j+1}]})p(1-p)^{i_{j+1}-i_j-1}, & j = n, \end{cases} \quad (2.8)$$

O modelo original de Barry & Hartigan (1992, 1993) foi posteriormente estendido por Loschi et al. (2003) e Loschi & Cruz (2005) para obter as distribuições *a posteriori* dos parâmetros  $p$ ,  $B$ ,  $\rho$  e também para obter a probabilidade *a posteriori* de cada instante "l" ser um ponto de mudança (denote por  $A_l$  esse evento). Essa extensão é mostrada nos dois parágrafos a seguir no caso particular em que, *a priori*,  $p \sim \text{Beta}(\alpha, \beta)$ . Como consequência dessa consideração, segue que as distribuições *a posteriori* de  $\rho$ ,  $B$ ,  $p$  e  $A_l$  são dadas, respectivamente, por:

$$f(\rho = \{i_0, \dots, i_b\} | y_1, \dots, y_n) = \frac{\Gamma(\alpha + b - 1)\Gamma(n + \beta - b) \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]})}{\sum_{b=1}^n \left\{ \Gamma(\alpha + b - 1)\Gamma(n + \beta - b) \sum_{C_b} \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) \right\}}, \quad (2.9)$$

$$f(B = b | y_1, \dots, y_n) = \frac{\Gamma(\alpha + b - 1)\Gamma(n - \beta - b) \sum_{C_b} \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]})}{\sum_{b=1}^n \left\{ \Gamma(\alpha + b - 1)\Gamma(n + \beta - b) \sum_{C_b} \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) \right\}}, \quad (2.10)$$

$$f(p | y_1, \dots, y_n) = \frac{\Gamma(\alpha + b - 1) \sum_{b=1}^n \left[ p^{\alpha+b-2} (1-p)^{\beta+n-b-1} \sum_{C_b} \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]}) \right]}{\sum_{b=1}^n \left[ \Gamma(\alpha + b - 1)\Gamma(n + \beta - b) \sum_{C_b} \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) \right]}, \quad (2.11)$$

$$f(A_l | y_1, \dots, y_n) = \frac{\sum_{C_l} \Gamma(\alpha + b - 1)\Gamma(n - \beta - b) \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(y_{[i_j i_{j+1}]})}{\sum_{b=1}^n \left\{ \Gamma(\alpha + b - 1)\Gamma(n + \beta - b) \sum_{C_b} \prod_{j=0}^{b-1} f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) \right\}}, \quad (2.12)$$

onde  $C_b$  é o conjunto de todas as partições com  $B = b$  blocos tais que  $0 = i_0 < i_1, \dots, < i_b = n$  e  $C_l$  é o conjunto de todas as partições que contém o instante  $l$  como um ponto de mudança, isto é, cada partição em  $C_l$  assume a forma  $\{i_0, \dots, i_{k-1}, i_k = l, i_{k+1}, \dots, i_b\}$  tal que  $0 = i_0 < i_{k-1} < l + 1 < i_{k+1} < \dots < i_b = n$ .

Na Seção 2.2 os métodos computacionais para o MPP discutidos em Barry & Hartigan (1993), Loschi et al. (2003) e Loschi & Cruz (2005) são apresentados.

## 2.2 Métodos Computacionais para MPP

Apesar das distribuições *a posteriori* terem formas fechadas no caso particular visto na Seção 2.1, suas determinações são caras do ponto de vista computacional. Para gerar das distribuições *a posteriori* para  $\rho$ ,  $B$ ,  $p$  e  $A_l$  e calcular as estimativas produto de  $\theta_k$ , alguns artigos na literatura sugerem utilizar o amostrador de Gibbs. As distribuições *a posteriori* de  $B$  e  $A_l$  são obtidas a partir da distribuição *a posteriori* de  $\rho$ . Desta forma, faz-se necessário encontrar-se apenas as distribuições condicionais completas dos parâmetros  $\theta_k$ ,  $\rho$  e  $p$ .

Sejam  $\mathbf{Y} = (Y_1, \dots, Y_n)$  e  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ . Denote por  $\boldsymbol{\theta}_{-k}$  o vetor  $(\theta_1, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_n)$ . Assuma que, dado uma partição  $\rho = \{i_0, \dots, i_b\}$ ,  $k \in [i_j + 1, i_{j+1}]$ , ou seja, tem-se que o  $k$ -ésimo componente de  $\theta$  é tal que  $\theta_k = \theta_{i_j+1} = \dots = \theta_{i_{j+1}}$ , para  $k = 1, \dots, n$  e  $i_j \in I$ . As distribuições condicionais completas *a posteriori* de  $p$ ,  $\rho$  e  $\theta_k$  para  $k = 1, \dots, n$  são dadas, respectivamente, por

$$f(p|\rho = \{i_0, \dots, i_b\}, \boldsymbol{\theta}, \mathbf{y}) \propto \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha+b-2} (1-p)^{\beta+n-b-1};$$

$$f(\rho = \{i_0, \dots, i_b\}|p, \boldsymbol{\theta}, B = b, \mathbf{y}) \propto p^{b-1} (1-p)^{n-b} \prod_{j=1}^b f_{[i_j i_{j+1}]}(y_{[i_j i_{j+1}]});$$

$$f(\theta_k|\rho = \{i_0, \dots, i_b\}, p, \boldsymbol{\theta}_{-k}, \mathbf{y}) \propto f_{[i_j i_{j+1}]}(\theta_k|y_{[i_j i_{j+1}]}).$$

Gerar amostras das distribuições condicionais completas de  $p$  e  $\theta_k$  é relativamente simples, ver Loschi & Cruz (2002), por exemplo. No entanto, não é fácil gerar diretamente da distribuição condicional completa de  $\rho$ , mas é possível definir uma quantidade aleatória auxiliar que torna mais fácil gerar da distribuição condicional completa de  $\rho$  e  $\theta_k$  como feito em Barry & Hartingan (1993). Para isto, será usada a equivalência entre a partição aleatória  $\rho$  e o vetor aleatório  $\mathbf{U} = (U_1, \dots, U_{n-1})$  definido na Seção 2.1. Assumindo esta equivalência entre  $\rho$  e  $\mathbf{U}$ , cada partição  $\rho^s = (U_1^{(s)}, \dots, U_{n-1}^{(s)})$ ,  $s \geq 1$ , 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  $U_r^{(s)}$  é gerado da distribuição condicional de  $U_r$ , dados  $U_1^{(s)}, \dots, U_{r-1}^{(s)}, U_{r+1}^{(s-1)}, \dots, U_{n-1}^{(s-1)}, p^{(s-1)}, \boldsymbol{\theta}^{(s-1)}$  e  $\mathbf{Y}$ . Como cada coordenada  $U_r$  de  $\mathbf{U}$  só assume valores no conjunto  $\{0, 1\}$  utiliza-se a seguinte razão para, no passo  $s$ , gerar de sua distribuição condicional completa *a posteriori*

$$\begin{aligned}
R_r &= \frac{P(U_r = 1 | U_1^{(s)}, \dots, U_{r-1}^{(s)}, U_{r+1}^{(s-1)}, \dots, U_{n-1}^{(s-1)}, p^{(s-1)}, \boldsymbol{\theta}^{(s-1)}; \mathbf{y})}{P(U_r = 0 | U_1^{(s)}, \dots, U_{r-1}^{(s)}, U_{r+1}^{(s-1)}, \dots, U_{n-1}^{(s-1)}, p^{(s-1)}, \boldsymbol{\theta}^{(s-1)}; \mathbf{y})} \\
&= \frac{f_{[xz]}(y_{[xz]}) \int_0^1 p^{b-2} (1-p)^{n-b+1} d\pi(p)}{f_{[xr]}(y_{[xr]}) f_{[rz]}(y_{[rz]}) \int_0^1 p^{b-1} (1-p)^{n-b} d\pi(p)},
\end{aligned} \tag{2.13}$$

onde  $x$  é o último ponto de mudança antes de  $r$ ,  $z$  é o próximo ponto de mudança depois de  $r$  e  $f_{[x,z]}(y_{[x,z]})$  denota a distribuição preditiva *a priori* relacionada às observações no bloco  $[x, z]$  dada em (2.3).

O critério de escolha dos valores para  $U_r^{(s)}$  é dado pela seguinte condição. Gere  $u$  de uma distribuição Uniforme em  $(0, 1)$ . Daí,

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

onde  $r = 1, \dots, n-1$ . O critério de escolha de  $U_r$  é baseado na forma de amostrar da distribuição de Bernoulli, discutido por exemplo, em Gamermam & Lopes (2006). Para maiores detalhes de como amostrar das distribuições *a posteriori* dos demais parâmetros veja Loschi & Cruz (2005).

Note que no MPP a dimensão do espaço paramétrico muda de partição para partição. A estratégia de amostragem sugerida por Barry & Hartigan (1993) apresenta uma solução para o problema de amostrar da distribuição *a posteriori* sem o uso do Monte Carlo via Cadeias de Markov com Saltos Reversíveis (MCMCSR) que só foi proposto posteriormente por Green (1995).

Utilizando a ideia sugerida por Barry & Hartigan (1993) para amostrar da partição  $\rho$ , 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  $B$  é o número de pontos de mudança acrescido de uma unidade.

Para calcular a probabilidade de mudança em cada instante Loschi & Cruz (2005) propõe um esquema também baseado no vetor de  $U'$ s o qual aproxima tal probabilidade por

$$\widehat{P(A_l)} = \frac{N}{T}, \quad l = 2, \dots, n, \tag{2.14}$$

onde  $N$  é a quantidade de vetores nos quais foi observada a mudança no instante  $l$  e  $T$  é o número de iterações de Gibbs, ou seja, esta probabilidade é estimada dividindo o número de zeros na  $l$ -ésima posição do vetor  $\mathbf{U}$  pelo tamanho da amostra final.

Nas próximas seções, o MPP será aplicado em alguns contextos particulares. Na Seção (2.3) o modelo proposto por Barry & Hartigan (1993) será apresentado e nas Seções (2.4) e (2.5) serão introduzidas algumas extensões desse modelo.

## 2.3 MPP para dados normais com médias normais

Assuma que, dados  $\mu_1, \dots, \mu_n$  e  $\sigma_y^2$ ,  $\mu_i \in \mathbb{R}$  e  $\sigma_y^2 \in \mathbb{R}^+$ , as variáveis aleatórias  $Y_1 \dots Y_n$  são independentes e tal que  $Y_i \mid \mu_i \sim \mathcal{N}(\mu_i, \sigma_y^2)$  com variância  $\sigma_y^2$  conhecida. Dada uma partição  $\rho = \{i_0, \dots, i_b\}$  de  $I = \{0, 1, \dots, n\}$ , assuma que existam  $\mu_{i_0 i_1}, \dots, \mu_{[i_{b-1} i_b]}, j = 0, \dots, b-1$  tais 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  $\rho = \{i_0, \dots, i_b\}$ , considere que  $Y_1 \dots Y_n$  são variáveis aleatórias tais que:

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

Denote por  $\boldsymbol{\mu}$  o vetor  $(\mu_1, \dots, \mu_n)$  e  $\mathbf{y} = (y_1, \dots, y_n)$ . Sob tais supostos, segue que a verossimilhança é dada por:

$$\begin{aligned}
 f(\mathbf{y} \mid \sigma_y^2, \boldsymbol{\mu}, \rho = \{i_0, \dots, i_b\}) &= \prod_{j=0}^{b-1} f(y_{i_{j+1}}, \dots, y_{i_j} \mid \mu_{[i_j i_{j+1}]}) \\
 &= \prod_{j=0}^{b-1} \prod_{k=i_j+1}^{i_{j+1}} f(y_k \mid \mu_{[i_j i_{j+1}]}) \\
 &= \prod_{j=0}^{b-1} \prod_{k=i_j+1}^{i_{j+1}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp \left\{ -\frac{1}{2\sigma_y^2} (y_k - \mu_{[i_j i_{j+1}]})^2 \right\} \right).
 \end{aligned} \tag{2.15}$$

Para a construção da distribuição *a priori* conjunta para  $\boldsymbol{\mu}$  assumamos que os parâmetros comuns  $\mu_{[i_0 i_1]}, \dots, \mu_{[i_{b-1} i_b]}$  são independentes e identicamente distribuídos com  $\mu_{[i_j i_{j+1}]} \sim \mathcal{N}(\mu_0, \sigma_\mu^2)$ ,  $j = 0, \dots, b-1$ . Note que o vetor  $\boldsymbol{\mu}$  pode ser escrito em função dos parâmetros comuns da seguinte forma

$$\boldsymbol{\mu} = \sum_{j=0}^{b-1} \left( \mu_{[i_j i_{j+1}]} \mathbf{1}\{i_j < 1 < i_{j+1}\}, \dots, \mu_{[i_j i_{j+1}]} \mathbf{1}\{i_j < n < i_{j+1}\} \right).$$

Como consequência, tem-se que a distribuição *a priori* conjugada dos  $\boldsymbol{\mu}$  é

$$\begin{aligned} \pi(\boldsymbol{\mu} \mid \rho = \{i_0, \dots, i_b\}, \sigma_\mu^2) &= \prod_{j=0}^{b-1} \pi(\mu_{[i_j i_{j+1}]}) \\ &\propto \left( \frac{1}{\sigma_\mu^2} \right)^{\frac{b}{2}} \prod_{j=0}^{b-1} \exp \left\{ -\frac{1}{2\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\}, \end{aligned} \quad (2.16)$$

onde  $\mu_0$  e  $\sigma_\mu^2$  são constantes. Um estudo de sensibilidade sobre a escolha de  $\mu_0$  foi realizado por Ferreira et al. (2014), esses autores propõem que  $\mu_0$  seja escolhido como sendo o valor mínimo da série, o valor máximo ou a média das observações, nesses casos as estimativas produto obtidas foram extremamente próximas.

A distribuição *a priori* para a partição aleatória  $\rho$ , será considerada a mesma distribuição produto proposta Barry e Hartigan (1992), a saber,

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

onde  $p \in (0, 1)$ .

Finalizando as especificações *a priori*, assumamos que,  $p \sim \text{Beta}(\alpha, \beta)$ ,  $\alpha > 0, \beta > 0$ . Como consequência das suposições anteriores segue que, a distribuição *a posteriori* conjunta de  $\boldsymbol{\mu}, \sigma_y^2, \rho, \sigma_\mu^2, p$  é dada por:

$$\begin{aligned}
f(\boldsymbol{\mu}, \sigma_y^2, \rho, \sigma_\mu^2, p \mid \mathbf{y}) &\propto f(\mathbf{y} \mid \boldsymbol{\mu}, \sigma_y^2, \rho) \pi(\boldsymbol{\mu} \mid \rho, \sigma_\mu^2) \pi(\rho \mid p) \pi(p) \pi(\sigma_\mu^2) \\
&\propto (1-p)^{n+\alpha-b-1} p^{b+\beta-2} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{j=0}^{b-1} \left[ \sum_{k=i_j+1}^{i_{j+1}} (y_k - \mu_{[i_j i_{j+1}]})^2 \right] \right\} \\
&\times \exp \left\{ -\frac{1}{2\sigma_\mu^2} \sum_{j=0}^{b-1} [(\mu_{[i_j i_{j+1}]} - \mu_0)^2] \right\}.
\end{aligned}$$

Como visto em (2.5) e de (2.9) a (2.12) para obter-se as distribuições *a posteriori* marginais de cada parâmetro faz-se necessário obter a distribuição preditiva *a priori* por bloco e a distribuição de cada parâmetro comum dadas as observações do bloco. Logo, a distribuição preditiva por bloco nesse modelo é dada por:

$$f_{[i_j i_{j+1}]}(y_{[i_j i_{j+1}]}) = 2\sqrt{\frac{\sigma_y^2}{2\pi n_j}} \times \phi_{n_j}(\mathbf{y}_{[i_j i_{j+1}]} \mid \bar{y}\mathbf{1}_{n_j}, \sigma_y^2) \phi\left(\bar{y} \mid \mu_0, \frac{n_j \sigma_\mu^2 + \sigma_y^2}{n_j}\right), \quad (2.17)$$

onde  $m^* = (\sigma_\mu^2 \sum_{l=i_j+1}^{i_{j+1}} y_l + \mu_0 \sigma_y^2) (\sigma_\mu^2 (i_{j+1} - i_j) + \sigma_y^2)^{-1}$ ,  $V^* = \sigma_y^2 \sigma_\mu^2 (n_j \sigma_\mu^2 + \sigma_y^2)^{-1}$  e  $n_j = i_{j+1} - i_j + 1$ . Ao longo desse texto,  $\phi(\cdot)$  e  $\Phi(\cdot)$  são, respectivamente, a função densidade de probabilidade (f.d.p) e a função de distribuição acumulada (F.D.A) da distribuição normal padrão.

A distribuição para  $\mu_{[i_j i_{j+1}]}$  dadas as observações  $y_{[i_j i_{j+1}]}$  do bloco  $[i_j i_{j+1}]$  é uma distribuição normal com média  $m^*$  e variância  $V^*$  dadas em (2.17), isto é,

$$f(\mu_{[i_j i_{j+1}]} \mid y_{[i_j i_{j+1}]}) = \phi\left(\mu_{[i_j i_{j+1}]} \mid m^*, V^*\right).$$

Também segue da equação (2.6) que a estimativa produto para  $\mu_i$  é

$$E(\mu_i \mid \mathbf{y}) = \sum_{i=0}^{i-1} \sum_{j=i}^n r_{ij}^* m_{ij}^*. \quad (2.18)$$

Considerando a identificação de pontos de mudanças em situações um pouco mais gerais do que a considerada aqui, nas duas próximas seções serão consideradas duas extensões do MPP. Na Seção 2.4 será assumido que os dados são normalmente distribuídos como em (2.15), mas a distribuição *a priori* para  $\mu_i$  possui distribuição normal assimétrica (Azzalini, 1985). Na Seção 2.5 será definido

o MPP para dados que tem distribuição normal assimétrica e a distribuição *a priori* para  $\mu_i$  será a distribuição normal.

## 2.4 MPP para dados normais com médias normais assimétricas

Sejam  $Y_1, \dots, Y_n$  variáveis aleatórias sequencialmente observadas e assumamos que, dados  $\mu_1, \dots, \mu_n$ , são independentes e tais que  $Y_i | \mu_i \sim \mathcal{N}(\mu_i, \sigma_y^2)$ ,  $i = 1, \dots, n$ , em que a variância  $\sigma_y^2$  é considerada comum para todo  $i$  e conhecida *a priori*,  $\sigma_y^2$  será considerada conhecida para minimizar o ruído de estimação do modelo. Assumamos a estrutura de agrupamento descrita na Seção 2.1 onde  $\rho = \{i_0, \dots, i_b\}$ ,  $0 = i_0 < i_1 < \dots < i_b = n$ , é uma partição aleatória do conjunto  $I = \{0, \dots, n\}$ . Para introduzir uma estrutura de agrupamento, dado  $\rho = \{i_0, \dots, i_b\}$ , assume-se que

i) Existem parâmetros comuns  $\mu_{[i_0 i_1]}, \dots, \mu_{[i_{b-1} i_b]}$  tais que

$$\mu_{[i_j i_{j+1}]} = \mu_{i_{j+1}} = \dots = \mu_{i_{j+1}}, \quad j = 1, \dots, b-1,$$

ii) As médias comuns dos blocos  $\mu_{[i_0 i_1]}, \dots, \mu_{[i_{b-1} i_b]}$  são independentes.

Esta partição em  $\mu$  induz uma partição na sequência  $\mathbf{Y} = (Y_1, \dots, Y_n)$  em  $B = b$  blocos contíguos de observações, denotados por  $Y_{[i_0 i_1]}, \dots, Y_{[i_{b-1} i_b]}$ . Seja  $\mu_{[i_j i_{j+1}]}$  a média comum que indexa a distribuição das observações no bloco ou “cluster”  $Y_{[i_j i_{j+1}]}$ . Sob tais suposições, a função de verossimilhança para o MPP torna-se:

$$f(\mathbf{y} | \mu_1, \dots, \mu_n, \rho = \{i_0, \dots, i_b\}, p) = \prod_{j=0}^{b-1} \prod_{k=i_j+1}^{i_{j+1}} \left( \frac{1}{2\pi\sigma_y^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(y_k - \mu_{[i_j i_{j+1}]})^2}{2\sigma_y^2} \right\}. \quad (2.19)$$

A especificação do modelo fica concluída quando as distribuições *a priori* para os parâmetros são especificadas. A distribuição *a priori* para  $\rho$  possui forma produto e, nesse trabalho, assume-se as coesões de Yao (1984) dadas em (2.1). Desta forma, dado  $p \in (0, 1)$ , a distribuição *a priori* para  $\rho$  é a distribuição produto definida em (2.2). Para  $p$  considere que

$$p \sim \text{Beta}(\alpha, \beta), \quad (2.20)$$

onde  $\alpha$  e  $\beta$  são números reais positivos.

Para realizar o estudo de sensibilidade que será mostrado no próximo capítulo é assumido que a distribuição *a priori* para  $\mu_{[i_j i_{j+1}]}$  pertence à classe de distribuição de probabilidade normal assimétrica introduzida por Azzalini (1985) a qual é dada por

$$\mathcal{P}_\lambda = \left\{ \pi_\lambda(\mu) \in \mathcal{P} : \pi_\lambda(\mu) = \frac{2}{\sigma_\mu} \phi \left( \frac{(\mu - \mu_0)}{\sigma_\mu} \right) \Phi \left( \lambda \frac{(\mu - \mu_0)}{\sigma_\mu} \right), \lambda \in \mathbb{R} \right\}, \quad (2.21)$$

onde  $\mathcal{P}$  é o conjunto de medidas de probabilidade definidas em  $\mathbb{R}$ . Vale ressaltar que como o foco é o estudo de sensibilidade no modelo partição produto, faz-se necessário que seja fixada uma grade de valores para  $\lambda$  e posteriormente avaliar o comportamento das estimativas a medida que  $\lambda$  se afasta de zero. A família em (2.21) inclui, como caso particular, a distribuição *a priori* normal para  $\mu_{[i_j i_{j+1}]}$  que foi assumida no MPP introduzido em Barry & Hartigan (1993).

A distribuição preditiva *a priori* por bloco e a distribuição da média *a posteriori* por bloco dos parâmetros comum  $\mu_{[i_j i_{j+1}]}$  necessárias para se obter as distribuições *a posteriori* de  $\rho$ ,  $B$  e  $\mu_i$  são dadas na proposição a seguir.

**Proposição 1.** *Assuma que, dados  $\mu_1, \dots, \mu_n$ , as variáveis aleatórias  $Y_1, \dots, Y_n$  são independentes e tais que  $Y_i | \mu_i \sim \mathcal{N}(\mu_i, \sigma_y^2)$ ,  $i = 1, \dots, n$ . Suponha que, a priori,  $\mu_{[i_j i_{j+1}]}$  tenha distribuição na família  $\mathcal{P}_\lambda$  definida em (2.21),  $\rho$  tenha a distribuição produto definida em (2.2) e  $p$  tenha a distribuição a priori dada em (2.20), então*

*i) a distribuição preditiva a priori para o bloco  $Y_{[i_j i_{j+1}]}$  é*

$$\begin{aligned} f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) &= 2 \sqrt{\frac{\sigma_y^2}{2\pi n_j}} \phi_{n_j} \left( \mathbf{y}_{[i_j i_{j+1}]} | \bar{y} \mathbf{1}_{n_j}, \sigma_y^2 \mathbf{I}_{n_j} \right) \\ &\times \phi \left( \bar{y} | \mu_0, \frac{n_j \sigma_\mu^2 + \sigma_y^2}{n_j} \right) \Phi \left( \lambda m^* | \lambda \mu_0, \sigma_\mu^2 + \lambda^2 V^* \right), \end{aligned} \quad (2.22)$$

*ii) a distribuição a posteriori, por bloco, para  $\mu_{[i_j i_{j+1}]}$  é*

$$f(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) = \phi \left( \mu_{[i_j i_{j+1}]} | m^*, V^* \right) \frac{\Phi \left( \lambda \mu_{[i_j i_{j+1}]} | \lambda \mu_0, \sigma_\mu^2 \right)}{\Phi \left( \lambda m^* | \lambda \mu_0, \sigma_\mu^2 + \lambda^2 V^* \right)}, \quad (2.23)$$

iii) o valor esperado de  $\mu_{[i_j i_{j+1}]}$  dado  $\mathbf{y}_{[i_j i_{j+1}]}$  é

$$E(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) = m^* + \frac{\lambda V^* \phi(\lambda m^* | \lambda \mu_0, \sigma_\mu + \lambda^2 V^*)}{\Phi(\lambda m^* | \lambda \mu_0, \sigma_\mu + \lambda^2 V^*)}, \quad (2.24)$$

onde  $m^* = (\sigma_\mu^2 \sum_{l=i_j+1}^{i_{j+1}} y_l + \mu_0 \sigma_y^2) (\sigma_\mu^2 (i_{j+1} - i_j) + \sigma_y^2)^{-1}$ ,  $V^* = \sigma_y^2 \sigma_\mu^2 (n_j \sigma_\mu^2 + \sigma_y^2)^{-1}$ ,  $n_j = i_{j+1} - i_j + 1$  e  $\mu_{[i_j i_{j+1}]}$  é a média comum do bloco  $[i_j i_{j+1}]$  e não um vetor de médias.

Note que o modelo apresentado em Barry & Hartigan (1993) surge como um caso particular do modelo introduzido aqui. Observe que dos resultados apresentados na **Proposição 1**, se  $\lambda = 0$  obtém-se os mesmos resultados apresentados na Seção 2.3. A prova da **Proposição 1** pode ser encontrada no Apêndice B.

A distribuição em (2.23) pertence à família normal assimétrica multivariada generalizada ( $\mathcal{GMSN}$ ) introduzida por Gupta et al. (2004), ver Apêndice B para definição, a qual é denotada por

$$\mu_{[i_j i_{j+1}]} | y_{[i_j i_{j+1}]} \sim \mathcal{GMSN}_{1,1}(\xi_{[i_j i_{j+1}]}, \Sigma, D, v, \Delta),$$

em que  $\xi_{[i_j i_{j+1}]} = m^*$ ,  $\Sigma = V^*$ ,  $D = \lambda$ ,  $v = \lambda \mu_0$  e  $\Delta = \sigma_\mu$ . Segue da expressão (2.6) e da **Proposição 1** que a média *a posteriori* para cada  $\mu_i$ ,  $i = 1, \dots, n$ , é dada por

$$E(\mu_i | \mathbf{y}) = \sum_{i=0}^{i-1} \sum_{j=i}^n \left[ r_{[ij]}^* \left( m_{[ij]}^* + \frac{\lambda V_{[ij]}^* \phi(\lambda m_{[ij]}^* | \lambda \mu_0, \sigma_\mu^2 + \lambda^2 V_{[ij]}^*)}{\Phi(\lambda m_{[ij]}^* | \lambda \mu_0, \sigma_\mu^2 + \lambda^2 V_{[ij]}^*)} \right) \right], \quad (2.25)$$

onde  $r^*$  denota a relevância *a posteriori* para o bloco  $[i_j i_{j+1}]$  definida em (2.7),  $m_{[ij]}^*$  e  $V_{[ij]}^*$  são calculados como em (2.24) utilizando as informações do bloco  $[i, j]$ .

Godoi (2011) mostra que, no contexto em que não é levado em consideração a partição do espaço em blocos contíguos, a distribuição *a posteriori* para a média é fechada sob conjugação. Isso é observado nesse contexto do MPP para a distribuição de  $\mu$  por blocos dada em (2.24).

As distribuições *a posteriori* para  $\rho$ ,  $B$ ,  $p$  e  $\mu_i$ ,  $i = 1, \dots, n$  são obtidas substituindo os resultados dados na **Proposição 1** nas expressões (2.9), (2.10), (2.11) e (2.5). No decorrer deste texto o modelo apresentado nessa seção será denominado Modelo I.

### 2.4.1 Métodos computacionais - Modelo I

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 é uma opção natural quando as distribuições condicionais completas têm formas fechadas. Para maiores detalhes sobre esse algoritmo ver, por exemplo, Gamermam & Lopes (2006).

No caso em que distribuições normais assimétricas são eliciadas, é também comum utilizar-se a representação de Henze (1986) para tornar mais simples as expressões das distribuições condicionais completas. A representação de Henze para a família normal assimétrica é dada no seguinte lema

**Lema 1.** *Sejam  $X$  e  $T$  variáveis aleatórias independentes e identicamente distribuídas, tais que  $X \stackrel{d}{=} T \sim \mathcal{N}(0, 1)$ . Se  $Z \stackrel{d}{=} \delta|T| + \sqrt{1 - \delta^2}X$  então segue que  $Z \sim \mathcal{SN}(0, 1, \lambda)$  e  $\delta = \lambda(\sqrt{1 + \lambda^2})^{-1}$ .*

Logo, utilizando a representação estocástica de Henze dada no Lema 1 e considerando algumas propriedades de cálculo de probabilidade para incluir locação e escala no modelo obtem-se a seguinte representação hierárquica para a distribuição *a priori*, esta é, se  $\mu_i \sim \mathcal{SN}(\mu_0, \sigma_\mu^2, \lambda)$  e  $T \sim \mathcal{N}(0, 1)$ , então

$$\begin{cases} \mu | T = t \sim \mathcal{N}(\mu_0 + \delta\sigma_\mu|t|, \sigma_\mu^2(1 - \delta^2)), \\ T \sim \mathcal{N}(0, 1). \end{cases} \quad (2.26)$$

onde  $\delta = \lambda(\sqrt{1 + \lambda^2})^{-1}$ . Na prática, ao utilizar esta representação, utiliza-se da técnica de aumento de dados o que facilita a geração de amostras *a posteriori*, no entanto, essa técnica pode aumentar a correlação da cadeia, principalmente pelo fato de  $t$  ser amostrado separadamente. Utilizando-se dessa representação, novas expressões são obtidas para as distribuições condicionais completas.

Suponha que  $i \in [i_j i_{j+1}]$ ,  $j = 0, \dots, n$  e considere a representação hierárquica em (2.4.1). Então, no modelo proposto para identificação de pontos de mudança, a distribuição condicional completa para  $\mu_i | T_i = t$  é uma distribuição normal:

$$\mu_i | T_i = t_i, \mathbf{y}_{[i_j i_{j+1}]}, \rho, p \sim \mathcal{N} \left( \frac{\sigma_\mu^2(1 - \delta^2) \sum_{l=i_j+1}^{i_{j+1}} y_l + \sigma_y^2(\mu_0 + \delta\sigma_\mu|t_i|)}{(1 - \delta^2)\sigma_\mu^2(i_{j+1} + 1 - i_j) + \sigma_y^2}; \frac{\sigma_y^2(1 - \delta^2)\sigma_\mu^2}{(1 - \delta^2)\sigma_\mu^2(i_{j+1} + 1 - i_j) + \sigma_y^2} \right).$$

A distribuição condicional completa para a variável latente  $t_i$ ,  $i = 1, \dots, n$  e para o parâmetro  $p$  são, respectivamente, dadas por

$$t_i | \mu_i, \rho, p, Y_i \sim N \left( \frac{\delta \sigma_\mu (\mu_i - \mu_0)}{\sigma_\mu^2}, 1 - \delta^2 \right),$$

e

$$p \sim \text{Beta}(b + \alpha + 1; n + \beta - b).$$

Já a razão  $R_r$  definida em (2.13) é dada por:

$$\begin{aligned} R_r &= \frac{\left( \frac{2\pi\sigma_y^2}{\sigma_\mu^2(z+1-x)+\sigma_y^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(\sigma_\mu^2 \sum_{l=x}^z y_l^2 + \sigma_y^2 \mu_0)}{2\sigma_y^2 \sigma_\mu^2} + \frac{(\sigma_\mu^2 \sum_{l=x}^z y_l + \mu_0 \sigma_y^2)^2}{\sigma_\mu^2 \sigma_y^2 [\sigma_\mu^2(z+1-x) + \sigma_y^2]} \right\}}{\left( \frac{2\pi\sigma_y^2}{\sigma_\mu^2(r+1-x)+\sigma_y^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(\sigma_\mu^2 \sum_{l=x}^r y_l^2 + \sigma_y^2 \mu_0)}{2\sigma_y^2 \sigma_\mu^2} + \frac{(\sigma_\mu^2 \sum_{l=x}^r y_l + \mu_0 \sigma_y^2)^2}{\sigma_\mu^2 \sigma_y^2 [\sigma_\mu^2(r+1-x) + \sigma_y^2]} \right\}} \\ &\times \frac{1}{\left( \frac{2\pi\sigma_y^2}{\sigma_\mu^2(z+1-r)+\sigma_y^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(\sigma_\mu^2 \sum_{l=r}^z y_l^2 + \sigma_y^2 \mu_0)}{2\sigma_y^2 \sigma_\mu^2} + \frac{(\sigma_\mu^2 \sum_{l=r}^z y_l + \mu_0 \sigma_y^2)^2}{\sigma_\mu^2 \sigma_y^2 [\sigma_\mu^2(z+1-r) + \sigma_y^2]} \right\}} \\ &\times \frac{\Phi(\lambda d; \lambda \mu_0; \sigma_\mu^2 + \lambda D \lambda)}{\Phi(\lambda d_1; \lambda \mu_0; \sigma_\mu^2 + \lambda D \lambda)} \frac{\Gamma(n-b+a) \Gamma(b+c-1)}{\Gamma(n-b-1+a) \Gamma(b+c)}, \end{aligned}$$

onde  $D = \frac{\sigma_y^2 \sigma_\mu^2}{\sigma_\mu^2(b+1-a) + \sigma_y^2}$ ,  $d = \frac{\sigma_\mu^2 \sum_{l=x}^z y_l - \mu_0 \sigma_y^2}{\sigma_\mu^2(z+1-x) + \sigma_y^2}$ ,  $d_1 = \frac{\sigma_\mu^2 \sum_{l=x}^r y_l - \mu_0 \sigma_y^2}{\sigma_\mu^2(r+1-x) + \sigma_y^2}$ ,  $d_2 = \frac{\sigma_\mu^2 \sum_{l=r}^z y_l - \mu_0 \sigma_y^2}{\sigma_\mu^2(z+1-r) + \sigma_y^2}$ ,  $x$  é o último ponto de mudança antes de  $r$  e  $z$  é o próximo ponto de mudança depois de  $r$ .

Após o cálculo da razão  $R_r$  é possível obter facilmente o vetor de  $U$ 's e, conseqüentemente, obter as estimativas para o número de blocos,  $B$ , e para a probabilidade de mudança em cada instante,  $P(A_l)$ , como descrito na Seção 2.2. Note que na razão  $R_r$  não foi utilizada a representação estocástica, pois como a distribuição acumulada utilizada é univariada o custo computacional não é alto. A seguir será discutido o MPP para dados normais assimétricos.

## 2.5 MPP dados normais assimétricos com médias normais

Sejam  $Y_1, \dots, Y_n$  variáveis aleatórias sequencialmente observadas. Para realizar o estudo de robustez no modelo é assumido que a distribuição de cada  $Y_i$ , dado  $\mu_i$ ,  $\sigma_y^2$  e  $\lambda$  pertence à seguinte classe de

distribuição de probabilidade contaminada por multiplicação

$$\mathcal{P}_\lambda = \left\{ f_\lambda(Y_i) \in \mathcal{P} : f_\lambda(Y_i) = \frac{2}{\sigma_y} \phi \left( \frac{(y_i - \mu_i)}{\sigma_y} \right) \Phi \left( \lambda \frac{(y_i - \mu_i)}{\sigma_y} \right), \lambda \in \mathbb{R} \right\}, \quad (2.27)$$

onde  $\mathcal{P}$  é o conjunto de medidas de probabilidade definidas em  $\mathbb{R}$ . Como no modelo anterior, assume-se que  $\sigma_y^2$  e  $\lambda$  são fixos e conhecidos. Decorre daí que a função de verossimilhança neste modelo é a distribuição normal assimétrica definida por Azzalini (1985) que inclui, como caso particular, a função de verossimilhança normal que foi assumida no MPP introduzido em Barry & Hartigan (1993). Uma motivação para se trabalhar com uma classe de distribuições assimétricas é que a suposição de simetria em torno da média pode ser muito forte em algumas situações reais.

Como anteriormente, seja  $\rho = \{i_0, \dots, i_b\}$ ,  $0 = i_0 < i_1 < \dots < i_b = n$ , uma partição aleatória do conjunto  $I = \{0, \dots, n\}$  e para introduzir a estrutura de agrupamento, dado  $\rho = \{i_0, \dots, i_b\}$ , assumamos que existam parâmetros comuns  $\mu_{[i_0 i_1]}, \dots, \mu_{[i_{b-1} i_b]}$  tais que  $\mu_{[i_j i_{j+1}]} = \mu_{i_{j+1}} = \dots = \mu_{i_{j+1}}$ ,  $j = 1, \dots, b-1$ , e que as médias comuns dos blocos  $\mu_{[i_0 i_1]}, \dots, \mu_{[i_{b-1} i_b]}$  são independentes. Fazendo-se isso induz-se uma partição na sequência  $\mathbf{Y} = (Y_1, \dots, Y_n)$  em  $B = b$  blocos contíguos de observações, a qual é denotada por  $Y_{[i_0 i_1]}, \dots, Y_{[i_{b-1} i_b]}$ . Seja  $\mu_{[i_j i_{j+1}]}$  a média comum que indexa a distribuição das observações  $Y_{[i_j i_{j+1}]}$  no bloco ou “cluster”  $[i_j i_{j+1}]$ . Sob tais suposições, a função de verossimilhança para o MPP torna-se:

$$\begin{aligned} f(\mathbf{y} | \mu_{[i_0 i_1]}, \dots, \mu_{[i_{b-1} i_b]}, \rho = \{i_0, \dots, i_b\}, p) &= \prod_{j=0}^{b-1} \left[ \prod_{k=i_j+1}^{i_{j+1}} \left( \frac{1}{2\pi\sigma_y^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(y_k - \mu_{[i_j i_{j+1}]})^2}{2\sigma_y^2} \right\} \right. \\ &\quad \left. \times \Phi \left( \lambda \left( \frac{Y_k - \mu_{[i_j i_{j+1}]}}{\sigma_y} \right) \right) \right]. \end{aligned} \quad (2.28)$$

O modelo fica completamente definido quando as distribuições *a priori* para os parâmetros são especificadas. As distribuições *a priori* para  $\rho$  e  $p$  são as mesmas consideradas na Seção (2.4). Assume-se que as médias comuns dos blocos são independentes e identicamente distribuídas com

$$\mu_{[i_j i_{j+1}]} \sim \mathcal{N}(\mu_0, \sigma_\mu^2),$$

onde  $\mu_0 \in \mathbb{R}$  e  $\sigma_\mu^2 \in \mathbb{R}^+$  são conhecidos.

As distribuições *a posteriori* de todos os parâmetros ficam definidas se a distribuição preditiva *a priori* por bloco e a distribuição *a posteriori* por bloco dos parâmetros comum  $\mu_{[i_j i_{j+1}]}$  forem conhecidas. Na **Proposição 2**, a seguir, tais distribuições são obtidas.

**Proposição 2.** *Assuma que, dados  $\mu_1, \dots, \mu_n$ , as variáveis aleatórias  $Y_1, \dots, Y_n$  são independentes e tais que  $Y_i | \mu_i \sim \mathcal{SN}(\mu_i, \sigma_y^2, \lambda)$   $i = 1, \dots, n$ . Suponha que, a priori,  $\mu_{[i_j i_{j+1}]}$  segue uma distribuição normal  $\mathcal{N}(\mu_0, \sigma_\mu^2)$ ,  $j = 0, \dots, b-1$ ,  $\rho$  tenha a distribuição produto definida em (2.2) e  $p$  tenha a distribuição a priori dada em (2.20), então*

*i) a distribuição preditiva a priori para o bloco  $Y_{[i_j i_{j+1}]}$  é*

$$f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) = 2^{n_j} \sqrt{\frac{2\pi\sigma_y^2}{n_j}} \phi\left(\bar{y} \mid \mu_0, \frac{\sigma_\mu^2 n_j + \sigma_y^2}{n_j}\right) \phi_{n_j}\left(\mathbf{y}_{[i_j i_{j+1}]} \mid \bar{y} \mathbf{1}_{n_j}, \sigma_y^2 \mathbf{I}_{n_j}\right) \times \Phi_{n_j}\left(-\lambda m^* \mathbf{1}_{n_j} \mid -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j}\right), \quad (2.29)$$

*ii) a distribuição a posteriori, por bloco, para  $\mu_{[i_j i_{j+1}]}$  é*

$$f(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) = \frac{\phi\left(\mu_{[i_j i_{j+1}]} \mid m^*, V^*\right) \Phi_{n_j}\left(-\lambda \mu_{[i_j i_{j+1}]} \mid -\lambda y_k, \sigma_y^2 \mathbf{I}_{n_j}\right)}{\Phi_{n_j}\left(-\lambda m^* \mathbf{1}_{n_j} \mid -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j}\right)}, \quad (2.30)$$

*iii) o valor esperado de  $\mu_{[i_j i_{j+1}]}$  dado  $Y_{[i_j i_{j+1}]}$  é*

$$E(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) = m^* + \frac{-\lambda V^* \phi_{n_j}(-\lambda m^* \mathbf{1}_{n_j} \mid -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j})}{\Phi_{n_j}(-\lambda m^* \mathbf{1}_{n_j} \mid -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j})}, \quad (2.31)$$

onde  $n_j = i_{j+1} - i_j + 1$ ,  $m^* = \frac{(\sigma_y^2 \mu_0 + n_j \bar{y} \sigma_\mu^2)}{\sigma_y^2 + \sigma_\mu^2 n_j}$  e  $V^* = \frac{\sigma_y^2 \sigma_\mu^2}{n_j \sigma_\mu^2 + \sigma_y^2}$ .

A prova da **Proposição 2** pode ser encontrada no Apêndice B. Note que o modelo apresentado em Barry & Hartigan (1993) surge como um caso particular do modelo introduzido aqui. Se for considerado  $\lambda = 0$  nos resultados apresentados na **Proposição 2** volta-se aos resultados obtidos na Seção 2.3.

A distribuição em (2.30) também pertencente à família de distribuição normal assimétrica multivariada generalizada  $\mathcal{GMSN}$  introduzida por Gupta et al. (2004) (Apêndice B), e é denotada por:

$$\mu_{[i_j i_{j+1}] | Y_{[i_j i_{j+1}]} \sim \mathcal{GMSN}_{1, n_j}(\xi_{[i_j i_{j+1}]}, \Sigma, D, v, \Delta),$$

assumindo que  $\xi_{[i_j i_{j+1}]} = m^*$ ,  $\Sigma = V^*$ ,  $D = -\lambda \mathbf{1}_{n_j}$ ,  $v = -\lambda \mathbf{y}_{[i_j i_{j+1}]}$  e  $\Delta = \sigma_\mu \mathbf{I}_{n_j}$ , tem-se a distribuição dada na equação (2.30).

Como consequência, segue de (2.6) que as estimativas produto para cada  $\mu_i$ ,  $i = 1, \dots, n$ , são dadas por

$$E(\mu_i | \mathbf{y}) = \sum_{i=0}^{i-1} \sum_{j=i}^n \left[ r_{[ij]}^* \left( m_{[ij]}^* + \frac{\lambda V_{[ij]}^* \phi_{n_j}(\lambda m_{[ij]}^* \mathbf{1}_{n_j} | \lambda \mathbf{y}_{i_j i_{j+1}}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V_{[ij]}^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j})}{\Phi_{n_j}(\lambda m_{[ij]}^* \mathbf{1}_{n_j} | \lambda \mathbf{y}_{i_j i_{j+1}}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V_{[ij]}^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j})} \right) \right], \quad (2.32)$$

onde  $n_j$  é o número de observações no bloco  $[i_j i_{j+1}]$ ,  $m_{[i_j i_{j+1}]}^*$  e  $V_{[ij]}^*$  são calculadas como em (2.31) usando as informações no bloco  $[ij]$ .

As distribuições *a posteriori* para  $\rho$  e  $B$  são obtidas substituindo a distribuição preditiva dada na **Proposição 2** nas expressões (2.9), (2.10) e (2.11).

O modelo apresentado nessa seção, assumindo que a função de verossimilhança segue uma distribuição normal assimétrica, será denotado por Modelo II ao longo desse texto.

### 2.5.1 Métodos computacionais - Modelo II

Como já discutido anteriormente, o custo computacional para encontrar as distribuições *a posteriori* no MPP é alto, por isso as condicionais completas dos parâmetros de interesse são úteis. Note dos resultados da Proposição 2 que, especificamente no Modelo II, existe um custo computacional extra na obtenção dos resultados *a posteriori*, uma vez que as distribuições de  $Y_{[i_j i_{j+1}]}$  e  $\mu_{[i_j i_{j+1}] | \mathbf{y}_{[i_j i_{j+1}]}}$

dependem da função de distribuição acumulada da normal multivariada. Como já foi mencionado na Seção 2.4 o uso da representação estocástica de Henze (1986) para a distribuição normal assimétrica facilita os cálculos computacionais.

Logo, utilizando a representação estocástica de Henze dada no Lema 1 e estendida para o caso de locação e escala, é possível reescrever o modelo hierarquicamente da seguinte forma. Como  $Y_i | \mu_i, \sigma_y^2, \lambda \sim \mathcal{SN}(\mu_i, \sigma_y^2, \lambda)$ , então  $\forall i = 1, \dots, n$ , tem-se que

$$\begin{cases} Y_i | T_i = t_i, \mu_i, \sigma_y^2, \lambda \sim \mathcal{N}(\mu_i + \delta\sigma_y|t_i, \sigma_y^2(1 - \delta^2)), \\ T_i \sim \mathcal{N}(0, 1), \end{cases} \quad (2.33)$$

onde  $\delta = \lambda(\sqrt{1 + \lambda^2})^{-1}$ .

Consequentemente, a função de verossimilhança estendida para  $(y_1, \dots, y_n, T_1, \dots, T_n)$  é

$$\begin{aligned} f(\mathbf{y} | \mu_{[i_0 i_1]}, \dots, \mu_{[i_{b-1} i_b]}, \rho = \{i_0, \dots, i_b\}, p) &= \prod_{j=0}^{b-1} \left[ \prod_{k=i_j+1}^{i_{j+1}} \left( \frac{1}{2\pi\sigma_y^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(Y_k - \mu_{[i_j i_{j+1}]})^2}{2\sigma_y^2} \right\} \right] \\ &\times \Phi \left( \lambda \left( \frac{Y_k - \mu_{[i_j i_{j+1}]}}{\sigma_y} \right) \right) \left( \frac{2}{\pi} \right)^{\frac{1}{2}} \exp -\frac{y_k^2}{2} \right]. \end{aligned} \quad (2.34)$$

No caso do Modelo II, assumindo que  $i \in [i_j i_{j+1}]$ ,  $j = 0, \dots, b$  e utilizando a representação estocástica dada no lema 1, a distribuição condicional completa de  $\mu_i$  tem forma fechada conhecida e é dada por

$$\mu_i | \mathbf{y}, \mathbf{T} = \mathbf{t}, \rho, p \sim \mathcal{N} \left( \frac{\sigma_\mu^2 \sum_{l=i_j+1}^{i_{j+1}} (y_l - \delta\sigma_y|t_l) + \sigma_y^2 \mu_0 (1 - \delta^2)}{\sigma_\mu^2 (i_{j+1} + 1 - i_j) + \sigma_y^2 (1 - \delta^2)}; \frac{\sigma_y^2 \sigma_\mu^2 (1 - \delta^2)}{\sigma_\mu^2 (i_{j+1} + 1 - i_j) + \sigma_y^2 (1 - \delta^2)} \right).$$

A distribuição condicional completa para  $t_i$ ,  $i = 1, \dots, n$  e  $p$  são, respectivamente,

$$t_i | \mu_{[i_j i_{j+1}]}, \rho, p, y_i \sim N \left( \frac{(y_i - \delta\sigma_y \mu_{[i_j i_{j+1}]})}{\sigma_y^2}, \frac{\sigma_y^2 (1 - \delta^2)}{\sigma_y^2} \right),$$

$$p | \mu, \rho, \mathbf{T} = \mathbf{t}, \mathbf{y} \sim \text{beta}(\alpha + b - 1, n - b + \beta).$$

Note que a distribuição condicional completa é definida para  $\mu$  que neste caso é o parâmetro de locação e não o valor esperado da variável aleatória  $Y_i$ . Logo, é necessário corrigir a média dos dados no processo de simulação. Desta forma,

$$E(Y_i|\mu_i, \sigma_y^2, \lambda) = \mu_i + \sigma_y \delta \sqrt{\frac{2}{\pi}}.$$

Para gerar da distribuição *a posteriori* de  $\rho$  é necessário construir a razão  $R_r$  a partir de (2.13) resultando em:

$$\begin{aligned} R_r &= \frac{\left(\frac{2\pi\sigma_\mu^2(1-\delta^2)\sigma_y^2}{\sigma_\mu^2(z+1-x)+(1-\delta^2)\sigma_y^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{(\sigma_\mu^2 \sum_{l=x}^z (y_l - \delta\sigma_y |t_l|)^2 + (1-\delta^2)\sigma_y^2 \mu_0)}{2(1-\delta^2)\sigma_y^2 \sigma_\mu^2} + \frac{(\sigma_\mu^2 \sum_{l=x}^z (y_l - \delta\sigma_y |t_l|) + \mu_0(1-\delta^2)\sigma_y^2)^2}{\sigma_\mu^2(1-\delta^2)\sigma_y^2[\sigma_\mu^2(z+1-x)+(1-\delta^2)\sigma_y^2]}\right\}}{\left(\frac{2\pi(1-\delta^2)\sigma_y^2}{\sigma_\mu^2(r+1-x)+(1-\delta^2)\sigma_y^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{(\sigma_\mu^2 \sum_{l=x}^r (y_l - \delta\sigma_y |t_l|)^2 + (1-\delta^2)\sigma_y^2 \mu_0)}{2(1-\delta^2)\sigma_y^2 \sigma_\mu^2} + \frac{(\sigma_\mu^2 \sum_{l=x}^r (y_l - \delta\sigma_y |t_l|) + \mu_0(1-\delta^2)\sigma_y^2)^2}{\sigma_\mu^2(1-\delta^2)\sigma_y^2[\sigma_\mu^2(r+1-x)+(1-\delta^2)\sigma_y^2]}\right\}} \\ &\times \frac{1}{\left(\frac{2\pi(1-\delta^2)\sigma_y^2}{\sigma_\mu^2(z+1-r)+(1-\delta^2)\sigma_y^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{(\sigma_\mu^2 \sum_{l=r}^z (y_l - \delta\sigma_y |t_l|)^2 + (1-\delta^2)\sigma_y^2 \mu_0)}{2(1-\delta^2)\sigma_y^2 \sigma_\mu^2} + \frac{(\sigma_\mu^2 \sum_{l=r}^z (y_l - \delta\sigma_y |t_l|) + \mu_0(1-\delta^2)\sigma_y^2)^2}{\sigma_\mu^2(1-\delta^2)\sigma_y^2[\sigma_\mu^2(z+1-r)+(1-\delta^2)\sigma_y^2]}\right\}} \\ &\times \frac{\Gamma(n-b+a)\Gamma(b+c-1)}{\Gamma(n-b-1+a)\Gamma(b+c)}, \end{aligned}$$

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

Para o cálculo da razão  $R_r$  no caso do Modelo II, foi necessário utilizar a representação estocástica de Henze (1986), pois o custo computacional para o cálculo das distribuições acumuladas nesse modelo é muito alto. As estimativas para  $B$  e  $P(A_l)$  são facilmente encontradas após calculado o vetor de  $U$ 's definido na Seção 2.2.

No próximo capítulo será apresentado um estudo de sensibilidade informal. O objetivo principal deste tipo de estudo é avaliar como a especificação de alguns valores para  $\lambda$ , parâmetro responsável pela assimetização da curva, afetam as inferências *a posteriori* para as médias  $\mu_i$ 's e também para  $\rho$ ,  $B$  e a probabilidade de cada instante ser ponto de mudança,  $P(A_l)$ .

---

### Análise de Sensibilidade Informal no MPP: Uma ilustração do que pode acontecer em relação a especificações equivocadas da distribuição *a priori* ou da distribuição amostral

---

Neste capítulo, a título de motivação para o estudo da robustez global realizado no próximo capítulo, será apresentado um estudo de sensibilidade informal no Modelo Partição Produto. Entende-se por análise de sensibilidade informal, um estudo feito sobre o comportamento das estimativas *a posteriori* mediante a escolha de algumas distribuições *a priori* ou distribuições amostrais na classe de interesse.

Nesse tipo de estudo de sensibilidade informal, uma análise visual do comportamento da distribuição *a posteriori* é feito, mas não é realizado um estudo mais formal para avaliar as diferenças nas estimativas. Aqui, o objetivo é mostrar como as distribuições *a posteriori* para os  $\mu'_i$ s, número de blocos ( $B$ ), probabilidade de mudança em cada instante ( $P(A_l)$ ) e a partição de interesse ( $\rho'$ ) são afetadas quando são considerados valores distintos de  $\lambda$  nos modelos apresentados nas Seções 2.4 e 2.5, Modelo I e Modelo II, respectivamente.

Vale ressaltar que em um estudo de sensibilidade o parâmetro ao qual deseja-se investigar é con-

siderado fixo e uma grade de valores desse parâmetro é especificada. Nesse trabalho deseja-se investigar o comportamento do parâmetro de assimetria, logo uma grade de valores para  $\lambda$  será especificada.

### 3.1 Sensibilidade informal na especificação *a priori*

Os resultados apresentados nesta seção são baseados em uma única amostra, pois as réplicas poderiam suavizar as estimativas e suprimir a influencia de  $\lambda$  nas estimativas dos parâmetros do Modelo Partição Produto. Os dados foram gerados da distribuição normal com variância fixa,  $\sigma_y^2 = 1$ . A distribuição *a priori* para o parâmetro  $\mu_i$  é uma distribuição normal assimétrica  $\mathcal{SN}(\mu_0, \sigma_\mu^2, \lambda)$ , onde  $\lambda$  é fixado, e os valores de  $\lambda$  avaliados foram  $\lambda = (-100, -2, -0.7, 0, 0.7, 2, 100)$ . Os hiperparâmetros  $\mu_0$  e  $\sigma_\mu^2$  foram considerados fixos e iguais a média dos dados e 1000, respectivamente. Assume-se *a priori*, que  $p \sim \text{Beta}(1.1, 1.1)$ , foram escolhidos  $\alpha = 1.1$  e  $\beta = 1.1$  de modo que a distribuição fosse pouco informativa *a priori* e para garantir a existência da moda *a priori*. A Figura, 3.1 ilustra o comportamento da assimetria para os valores de  $\lambda$  avaliados nesse capítulo.

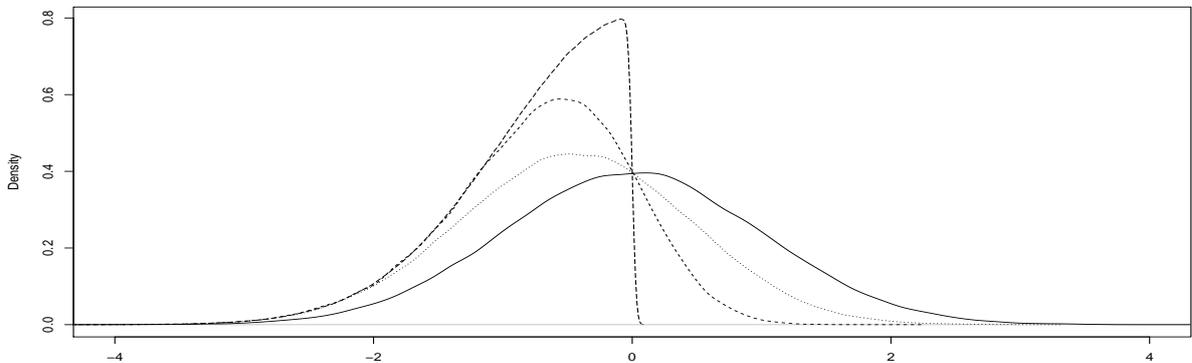


Figura 3.1: Distribuição normal assimétrica  $\mathcal{SN}(0,1,\lambda)$ .  $\lambda = -100$  (linha tracejada longa),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha pontilhada),  $\lambda = 0$  (linha contínua).

Para o MCMC, considerou-se uma cadeia com tamanho de 60000 iterações e o período de *burn-in* foi de 50000 iterações com um lag de 10 observações, logo o tamanho da amostra *a posteriori*

final foi de 1000 observações. O *burn-in* escolhido foi de 50000, pois com esse valor garantiu-se a convergência individual das cadeias dos parâmetros de interesse e o *lag* foi de tamanho 10, pois o custo computacional é alto. Para avaliar a sensibilidade foram considerados 2 tamanhos amostrais diferentes ( $n = 30, 90$ ) para os cenários com 3 blocos e ( $n = 40, 120$ ) para os cenários de 4 blocos, e não foram feitas réplicas Monte Carlo para garantir que o efeito do parâmetro de assimetria não fosse suavizado pela média das estimativas.

Todos os resultados foram obtidos por meio do software R (R Development Core Team, 2012). O estudo de sensibilidade apresentado nesta seção contempla o modelo apresentado na Seção 2.4, ou seja, quando a distribuição *a priori* para  $\mu$  pertence à classe de distribuição normal assimétrica. Lembrando que tal modelo é denominado nesse texto como Modelo I.

Nesse estudo de sensibilidade informal foram considerados os cenários descritos na Tabela 3.1 a seguir

Tabela 3.1: Cenários avaliados

Cenário	Dist. do 1º bloco.	Dist. do 2º bloco	Dist. do 3º bloco	Dist. do 4º bloco.
Sem mudança	$\mathcal{N}(0, 1)$	$\mathcal{N}(0, 1)$	$\mathcal{N}(0, 1)$	–
Degrau simétrico positivo	$\mathcal{N}(5, 1)$	$\mathcal{N}(8, 1)$	$\mathcal{N}(5, 1)$	–
Degrau simétrico negativo	$\mathcal{N}(-5, 1)$	$\mathcal{N}(-8, 1)$	$\mathcal{N}(-5, 1)$	–
Degrau não simétrico	$\mathcal{N}(-3, 1)$	$\mathcal{N}(7, 1)$	$\mathcal{N}(0, 1)$	–
Escada crescente	$\mathcal{N}(7, 1)$	$\mathcal{N}(10, 1)$	$\mathcal{N}(13, 1)$	–
Escada decrescente	$\mathcal{N}(13, 1)$	$\mathcal{N}(10, 1)$	$\mathcal{N}(7, 1)$	–
Epsilon	$\mathcal{N}(1.75, 1)$	$\mathcal{N}(2.5, 1)$	$\mathcal{N}(3.25, 1)$	–
Tendência	$\mathcal{N}(0, 0.01)$	$\mathcal{N}(0.3 + 0.008 * i, 0.1)$	$\mathcal{N}(0.6 + 0.018 * i, 0.2)$	$\mathcal{N}(1.1, 0.2)$

A seguir serão apresentados os resultados das estimativas *a posteriori* considerando os cenários apresentados na Tabela 3.1. Para facilitar a leitura do texto, como os resultados são apresentados por meio de vários gráficos. A apresentação dos resultados foi dividida em três etapas. As estimativas

produto e variância de  $\mu_i$ , serão apresentadas na Seção 3.1.1. As estimativas *a posteriori* para a partição e número de blocos são discutidas na Seção 3.1.2 e por fim, na Seção 3.1.3 são discutidas as estimativas *a posteriori* para a probabilidade de mudança em cada instante.

### 3.1.1 A influência de $\lambda$ e $n$ nas estimativas produto e na variância *a posteriori* das médias

As Figuras de 3.2 a 3.13, apresentam as estimativas produto e a variância *a posteriori* para cada  $\mu_i$ . Os gráficos (a) e (b) mostram as estimativas produto para  $\lambda \leq 0$  e  $\lambda \geq 0$ , respectivamente. Já os gráficos (c) e (d) dessa seção, correspondem as estimativas da variância *a posteriori* para cada  $\mu_i$ , sendo o gráfico (c) os valores de  $\lambda \leq 0$  e o gráfico (d) os valores de  $\lambda \geq 0$ . O primeiro cenário analisado é o cenário sem mudança, ilustrado na Figura 3.2,  $n = 30$ , e Figura 3.3,  $n = 90$ .

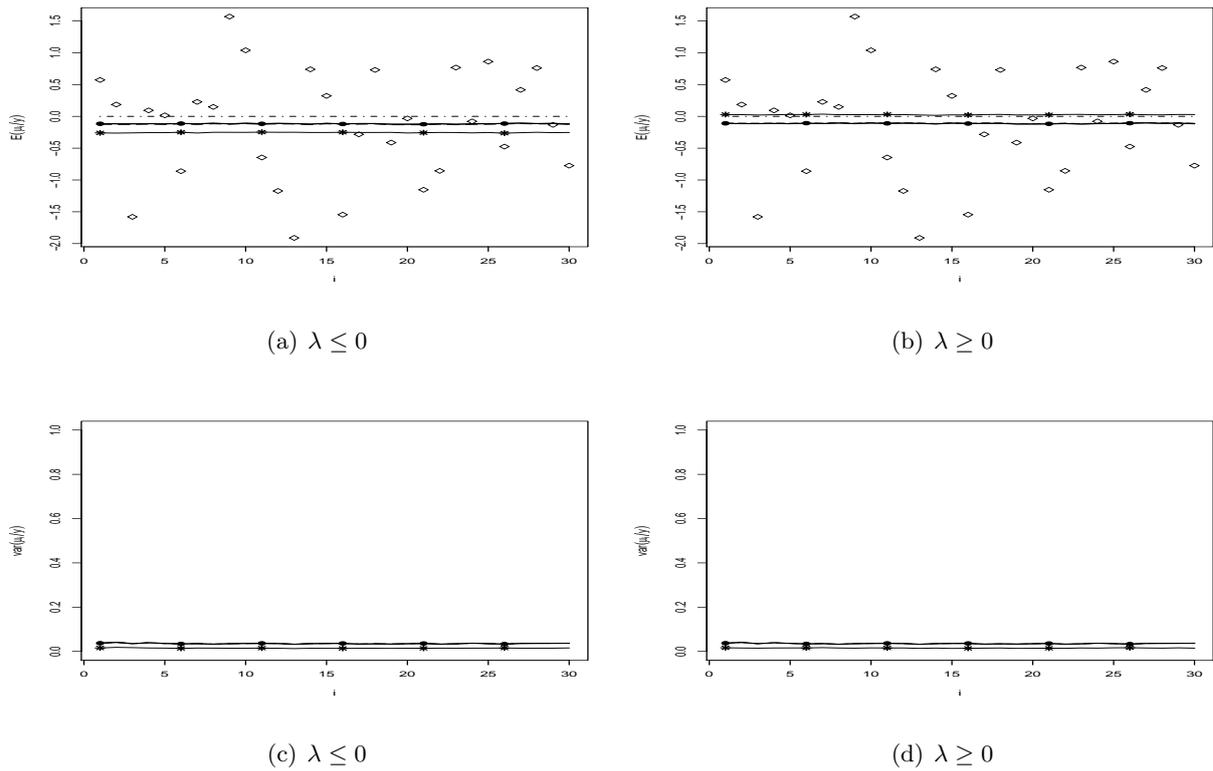


Figura 3.2: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 30$ .

Observa-se, no cenário sem mudança, que não houve grandes variações nas estimativas produto para os diferentes valores de  $\lambda$  e tamanho de amostra como pode-se observar nas Figuras 3.2 e 3.3. No cenário sem mudança para o tamanho de amostra 30, um valor de  $\lambda \leq 0$  e  $\lambda = 0.7$  e 2 resultou em uma leve subestimação das médias. Mas as estimativas produto se aproximam do verdadeiro valor da média quando a amostra é de tamanho 90. As estimativas das variâncias apresentam valores próximos a zero independente se os valores de  $\lambda$  são negativos ou positivos e quando o tamanho da amostra é de 90 esses valores são ainda mais próximos de zero.

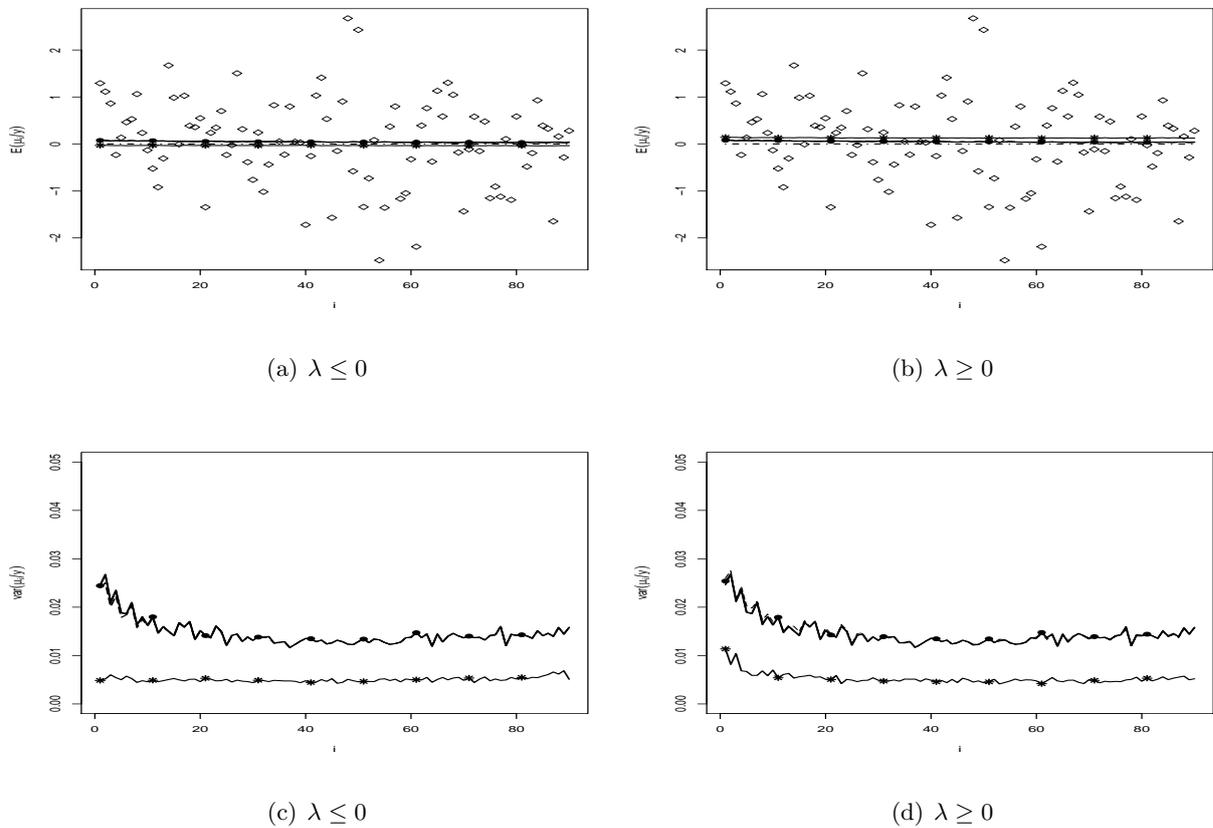


Figura 3.3: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 90$ .

Os próximos cenários avaliados são os cenários degrau positivo e degrau negativo, Figuras 3.4, 3.5, 3.6 e 3.7, respectivamente.

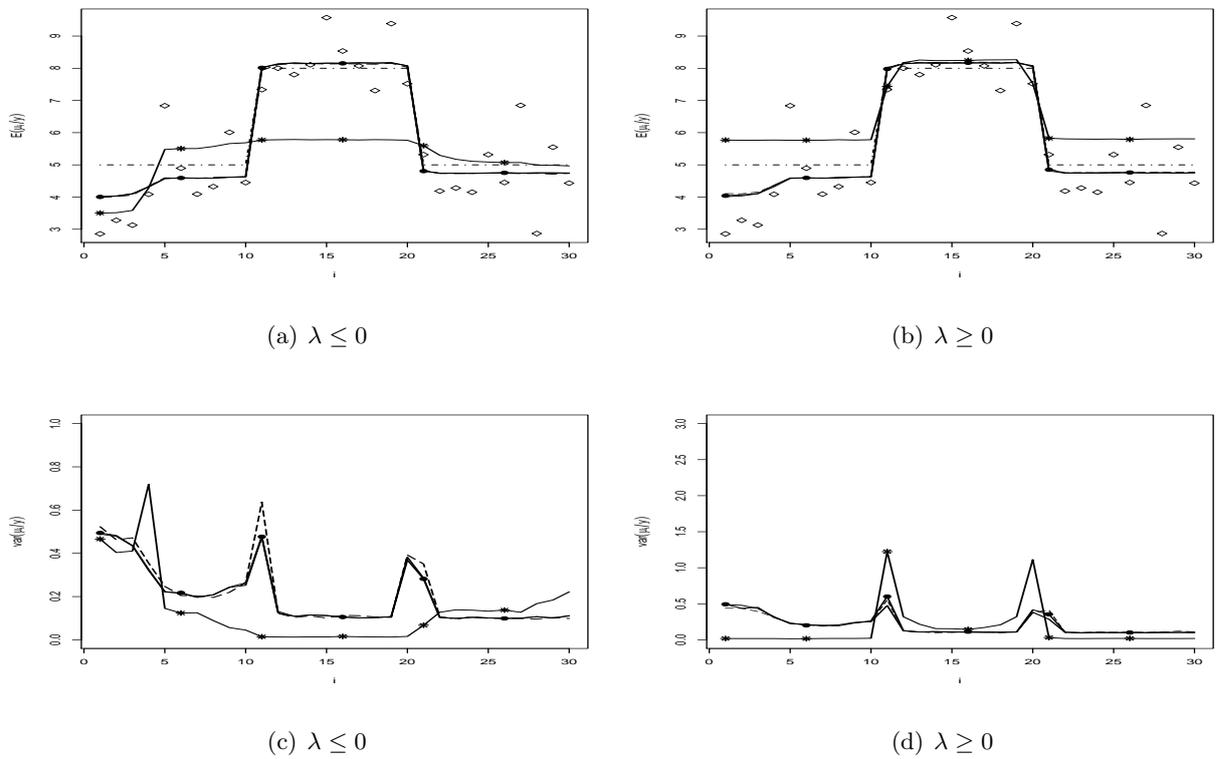


Figura 3.4: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 30$ .

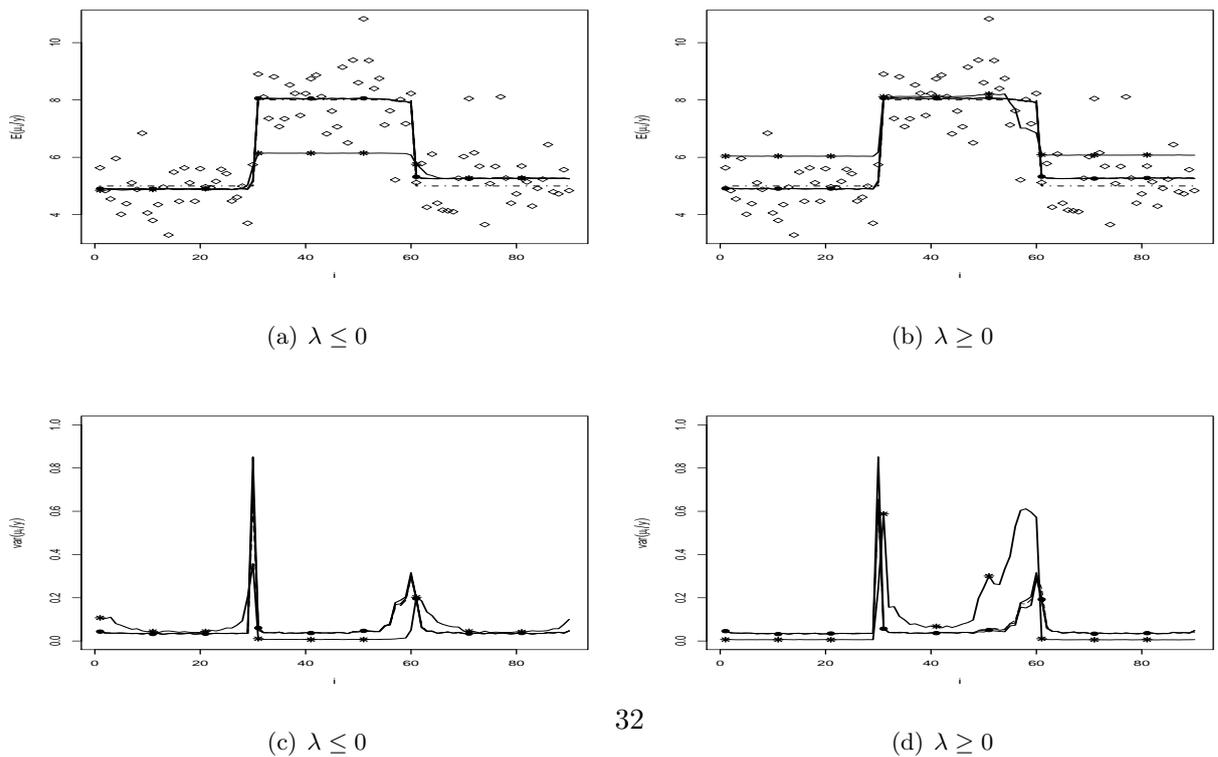


Figura 3.5: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 90$ .

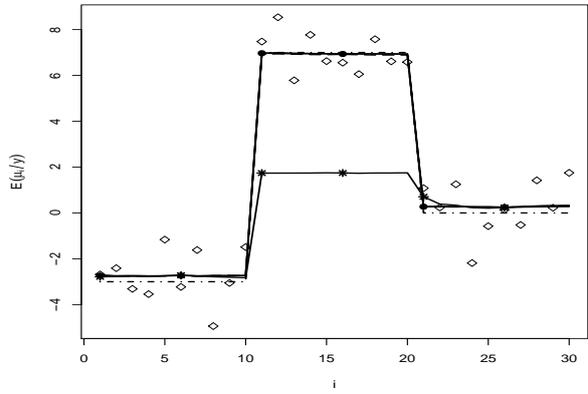


No cenário degrau positivo, cujas estimativas são dadas nas Figuras 3.4 e 3.5 para  $n = 30$  e  $90$ , respectivamente, como todos os dados são positivos, uma especificação negativa do parâmetro de assimetria poderia gerar estimativas equivocadas para as médias *a posteriori*, observe na Figuras 3.4 e 3.5 que para  $\lambda = \pm 100$  as estimativas das médias ficam comprometidas e, conseqüentemente, nota-se na Figura 3.4 que para  $\lambda = -100$  as estimativas são piores quando o tamanho de amostra é  $30$ .

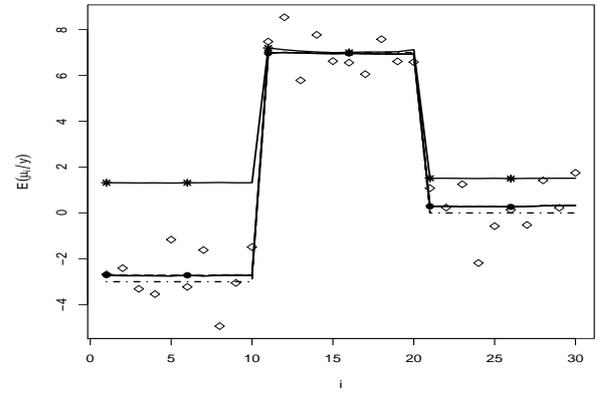
Note, na Figura 3.5, que com um tamanho de amostra maior,  $n = 90$ , as estimativas para a média obtidas utilizando  $\lambda = \pm 0.7$  e  $\pm 2$  *a priori* ficaram mais próximas da média real. Um comportamento similar é observado para o cenário degrau negativo, como mostram as Figuras 3.6 e 3.7 para  $n = 30$  e  $90$ , respectivamente, e observa-se que as piores estimativas para a média ocorreram para  $\lambda = 100$ , como era esperado uma vez que existe conflito entre a especificação de distribuição *a priori* e a função de verossimilhança. No entanto, nota-se que  $\lambda = -100$  gera estimativas melhores que  $\lambda = 100$ , mas são piores que as estimativas utilizando  $\lambda = \pm 0.7$  e  $\lambda = \pm 2$  que apresentam comportamento similar independente do sinal do parâmetro de assimetria.

O cenário degrau não simétrico é constituído por blocos negativos e positivos de tal modo que as diferenças entre as médias dos blocos não são constantes, ou seja, tal diferença é de 10 unidades do primeiro para o segundo bloco e de 7 unidades do segundo para o terceiro bloco. Os resultados desse cenário encontram-se nas Figuras 3.8 ( $n = 30$ ) e 3.9 ( $n = 90$ ).

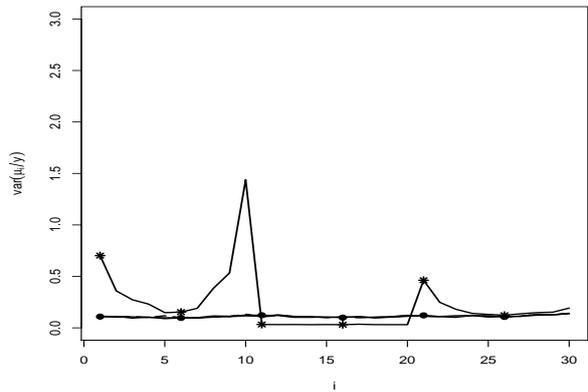
Nota-se, no cenário degrau não simétrico, que apesar do conjunto de dados dispor de observações positivas e negativas e o tamanho do salto ser maior que no degrau simétrico, as estimativas produto para a média quando o modelo assume o valor de  $\lambda = \pm 100$  estão distantes das verdadeiras médias  $(-3, 7, 0)$ . Mas, note que  $\lambda = -100$  fornece melhores estimativas, para o primeiro e terceiro blocos, onde a média é  $-3$  e  $0$ , respectivamente. Por outro lado, quando  $\lambda = 100$  a média (7) do segundo bloco é melhor estimada, uma vez que  $\lambda$  positivo neste caso não está em conflito com os dados do segundo bloco que também são positivos.



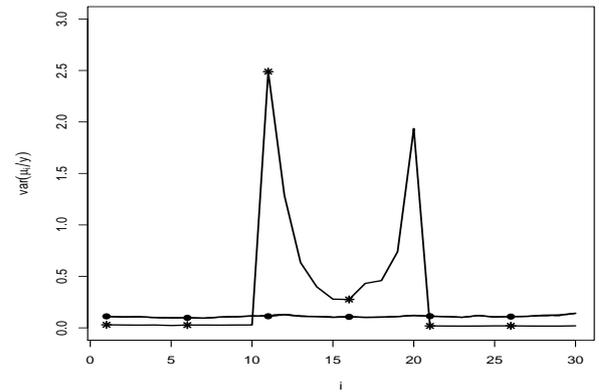
(a)  $\lambda \leq 0$



(b)  $\lambda \geq 0$



(c)  $\lambda \leq 0$



(d)  $\lambda \geq 0$

Figura 3.8: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 30$ .

Observa-se, também, que mesmo com o tamanho de amostra 90, as estimativas não melhoram para  $\lambda = \pm 100$ . Para valores de  $\lambda = 0, \pm 0.7$  e  $\pm 2$  as estimativas produto são praticamente as mesmas. Em relação a variância *a posteriori* de  $\mu_i$ , pode-se observar um pico para as estimativas nas posições 11 e 21, quando  $n = 30$ , e 31 e 61 quando  $n = 90$ , além de um pico na posição 1 quando  $\lambda = -100$  o que resulta maior incerteza sobre  $\mu_i$  nesses instantes.

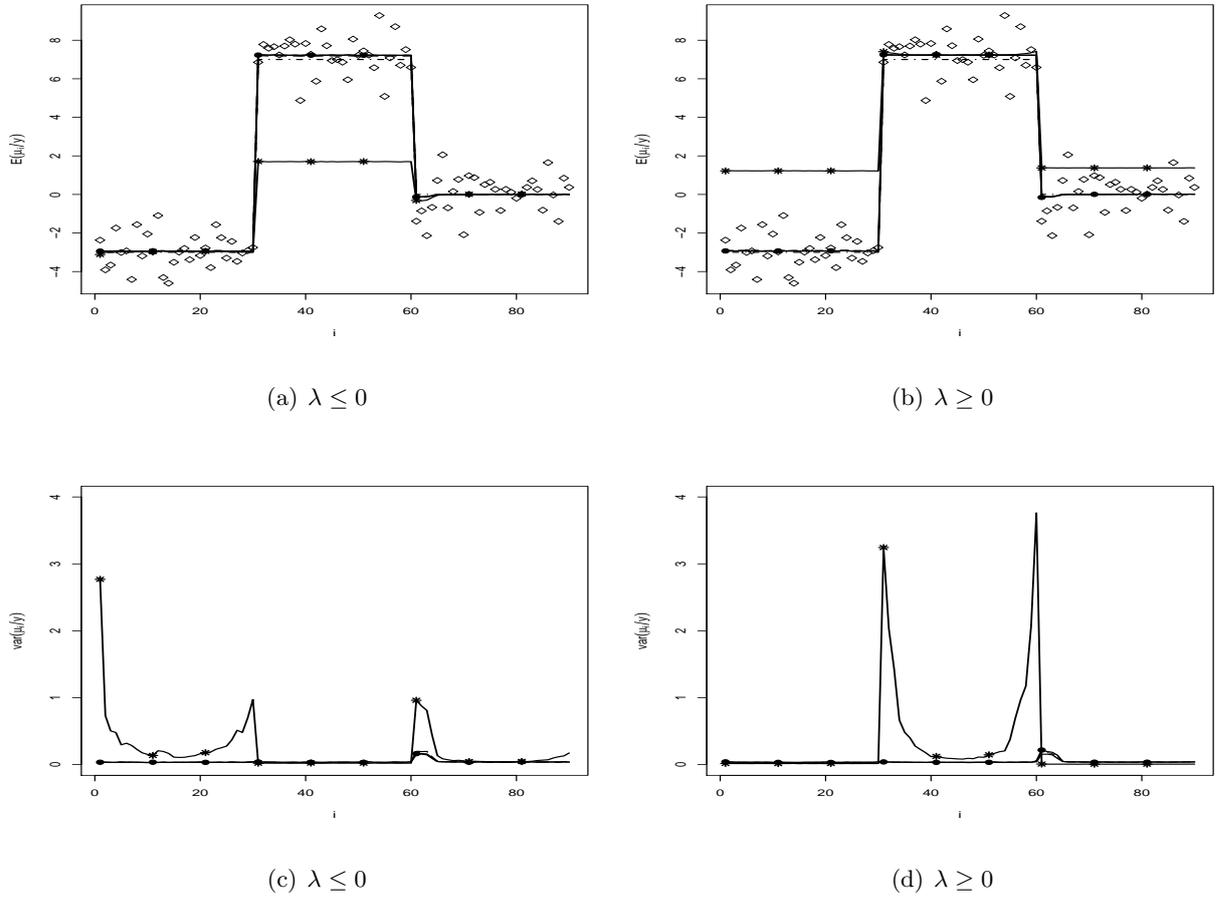


Figura 3.9: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 90$ .

No cenário escada, foram considerados dois pontos de mudança e as médias dos blocos foram consideradas sempre positivas e crescentes, cenário escada crescente, e positivas e decrescentes no cenário escada decrescente. A diferença entre os blocos vizinhos é de 3 unidades e esses cenários são ilustrados nas Figuras 3.10, 3.11, 3.12 e 3.13. Nota-se, nos cenários escada crescente e decrescente, que mesmo quando o tamanho da amostra aumenta as estimativas para a média quando o modelo assume valores de  $\lambda = \pm 100$  ficam comprometidas quando comparadas as estimativas obtidas assumindo  $\lambda = 0, \pm 0.7$  e  $\pm 2$ .

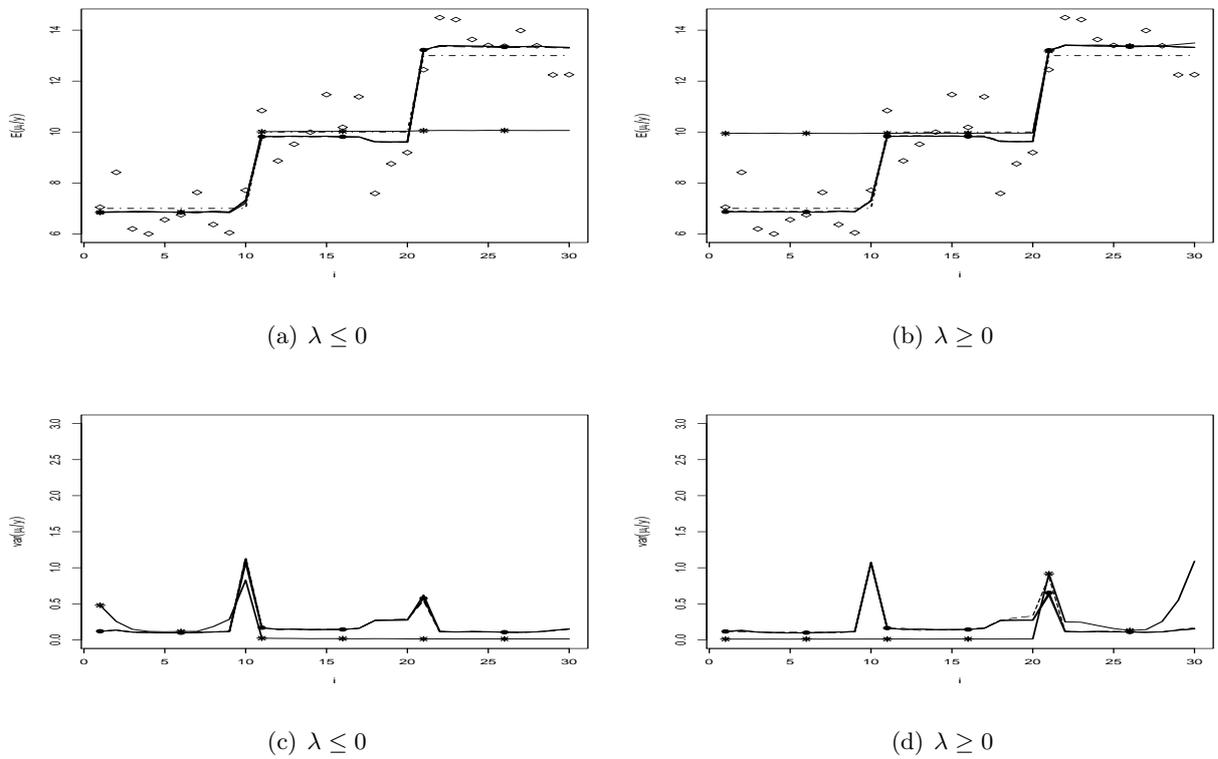


Figura 3.10: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 30$ .

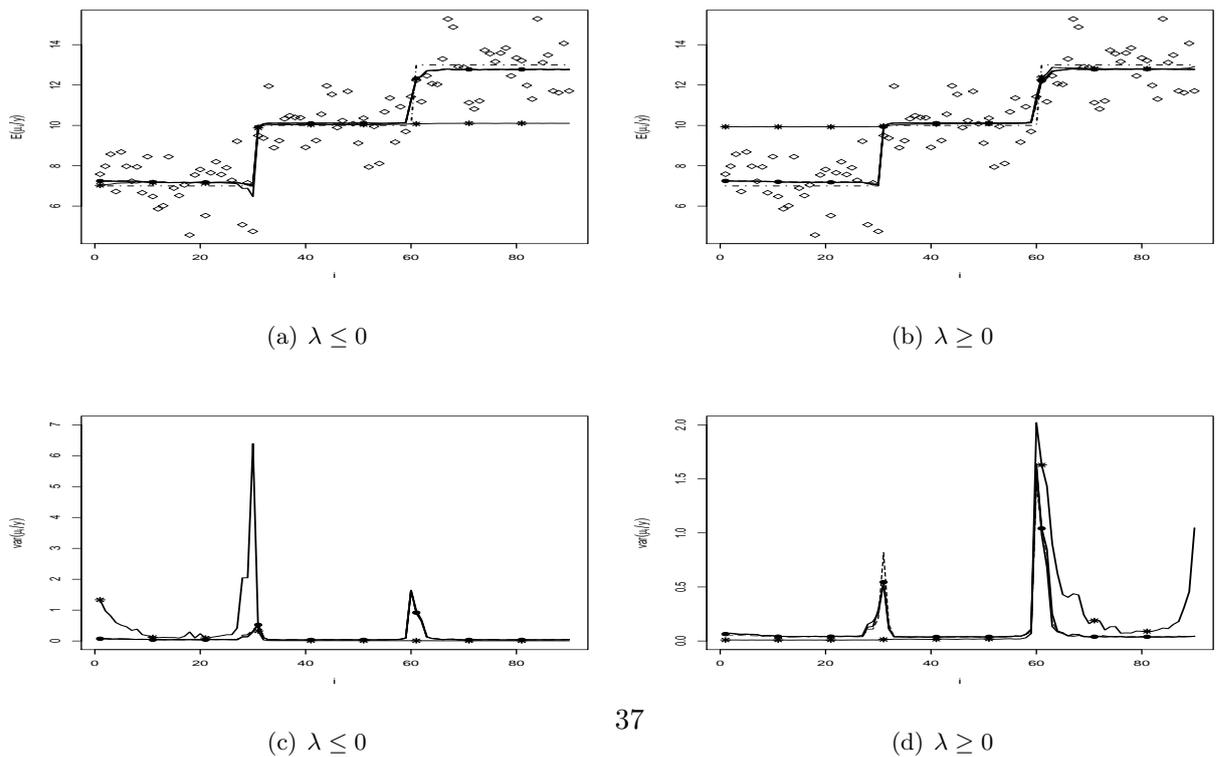


Figura 3.11: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 90$ .

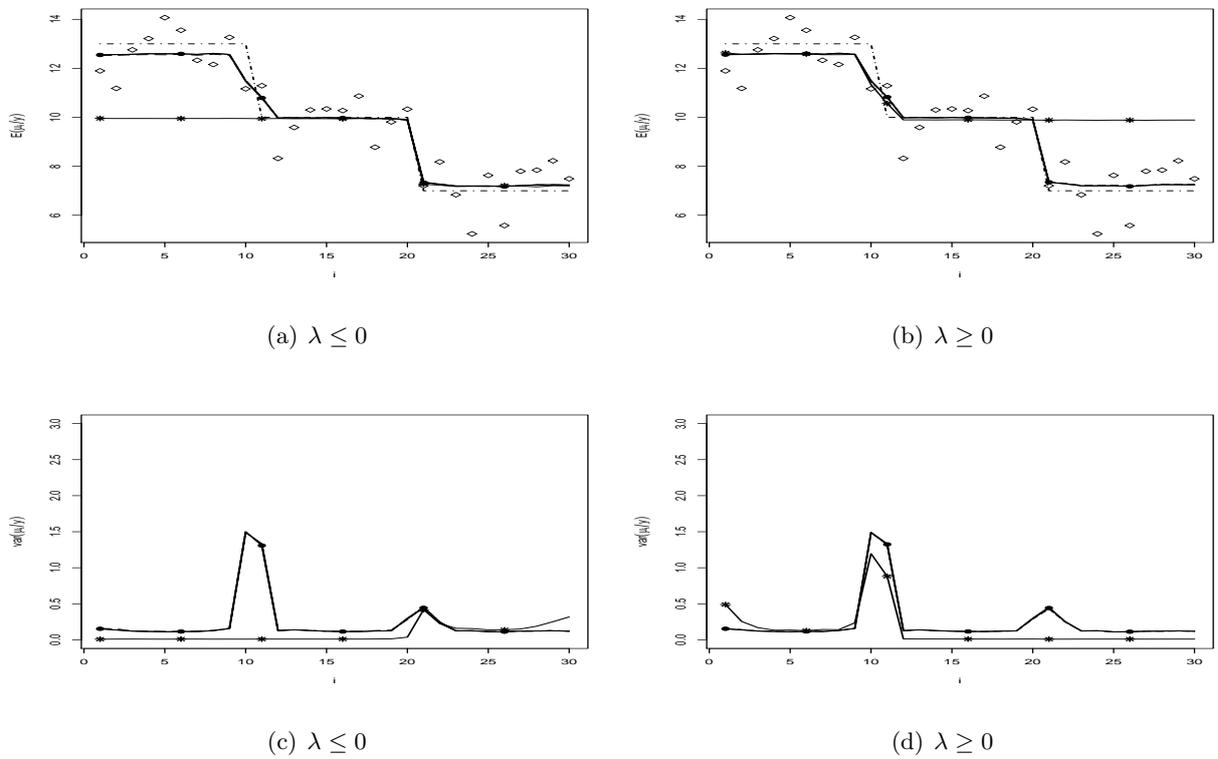


Figura 3.12: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 30$ .

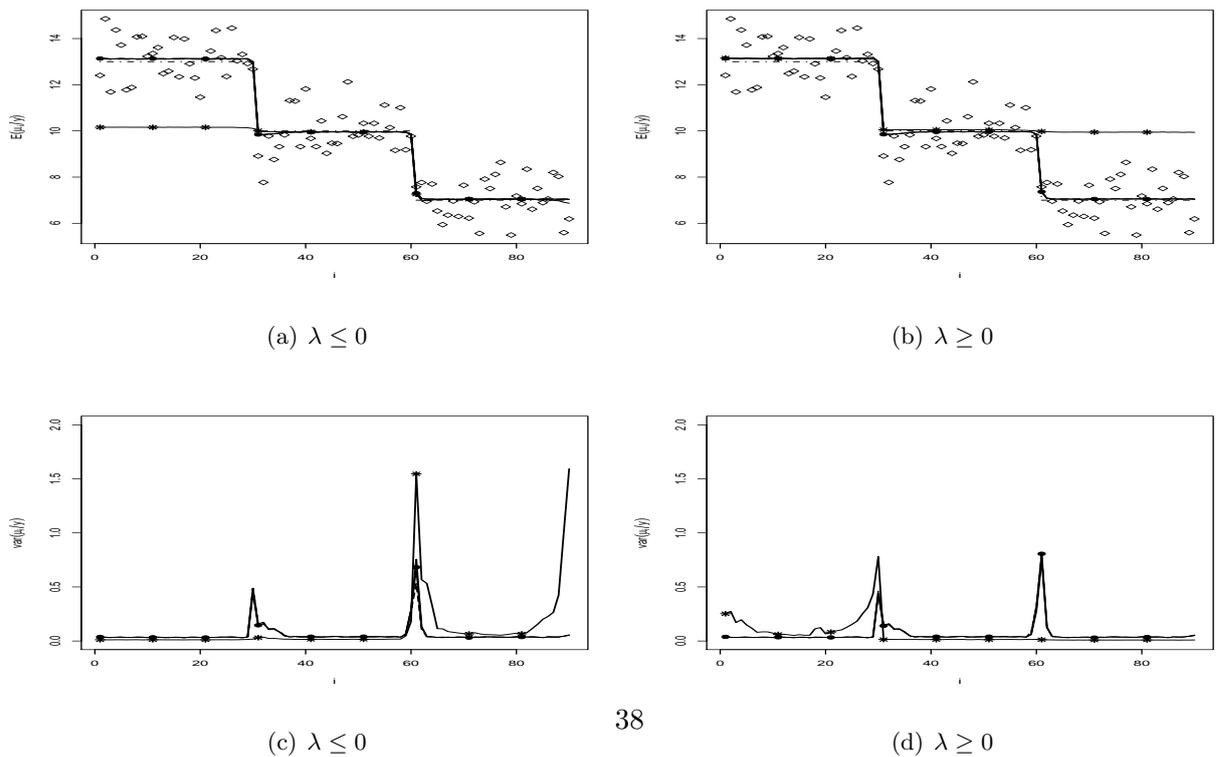


Figura 3.13: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 90$ .

O cenário epsilon está ilustrado nas Figuras 3.14 e 3.15, este cenário também possui o formato escada, no entanto, a diferença entre os blocos é de apenas 0.75.

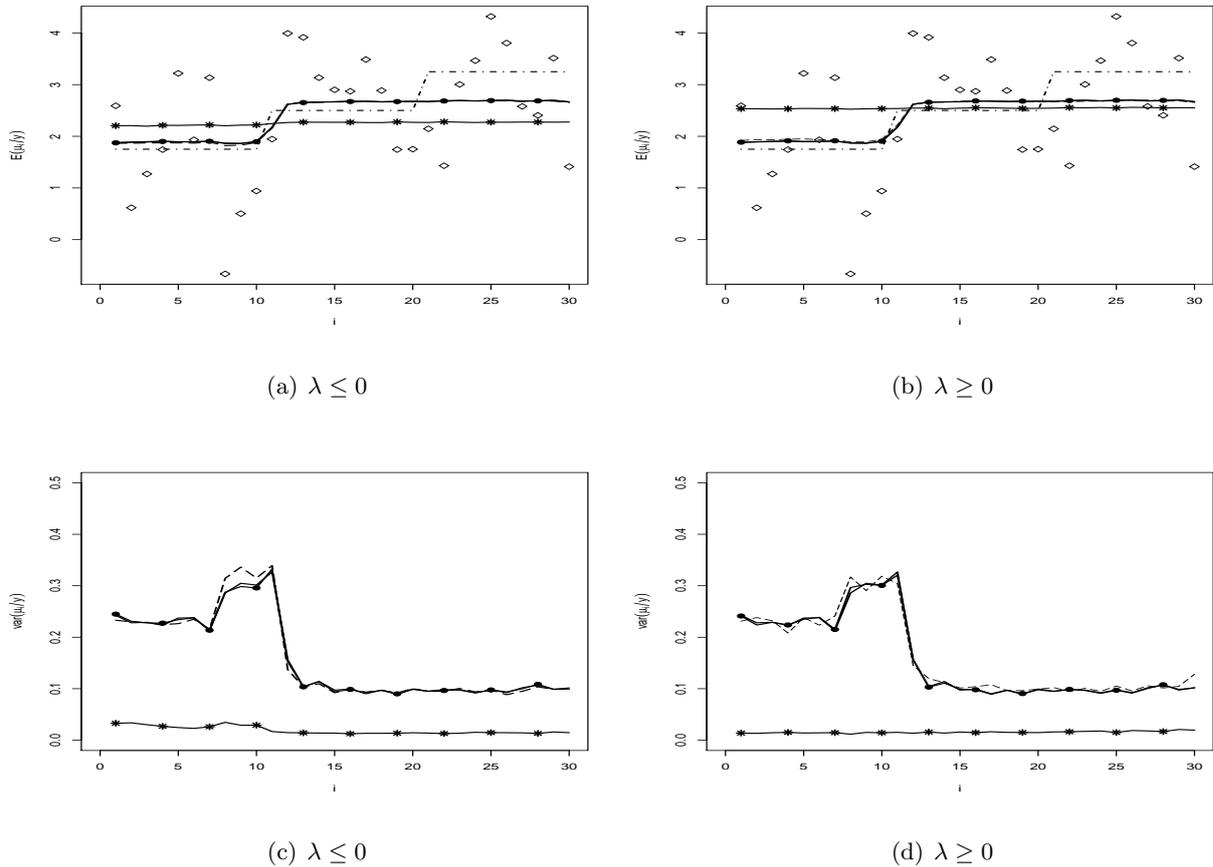


Figura 3.14: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 30$ .

Nota-se nesse cenário que para  $\lambda = \pm 100$  e  $n = 30$  o modelo tende a estimar as médias em torno de um único valor, os demais valores de  $\lambda$  estimam a média em torno de dois valores, e a média do último bloco é afetada. No cenário escada crescente com um salto maior entre os blocos isso não ocorre.

Observa-se no cenário epsilon que quando o tamanho da amostra aumenta as estimativas tendem

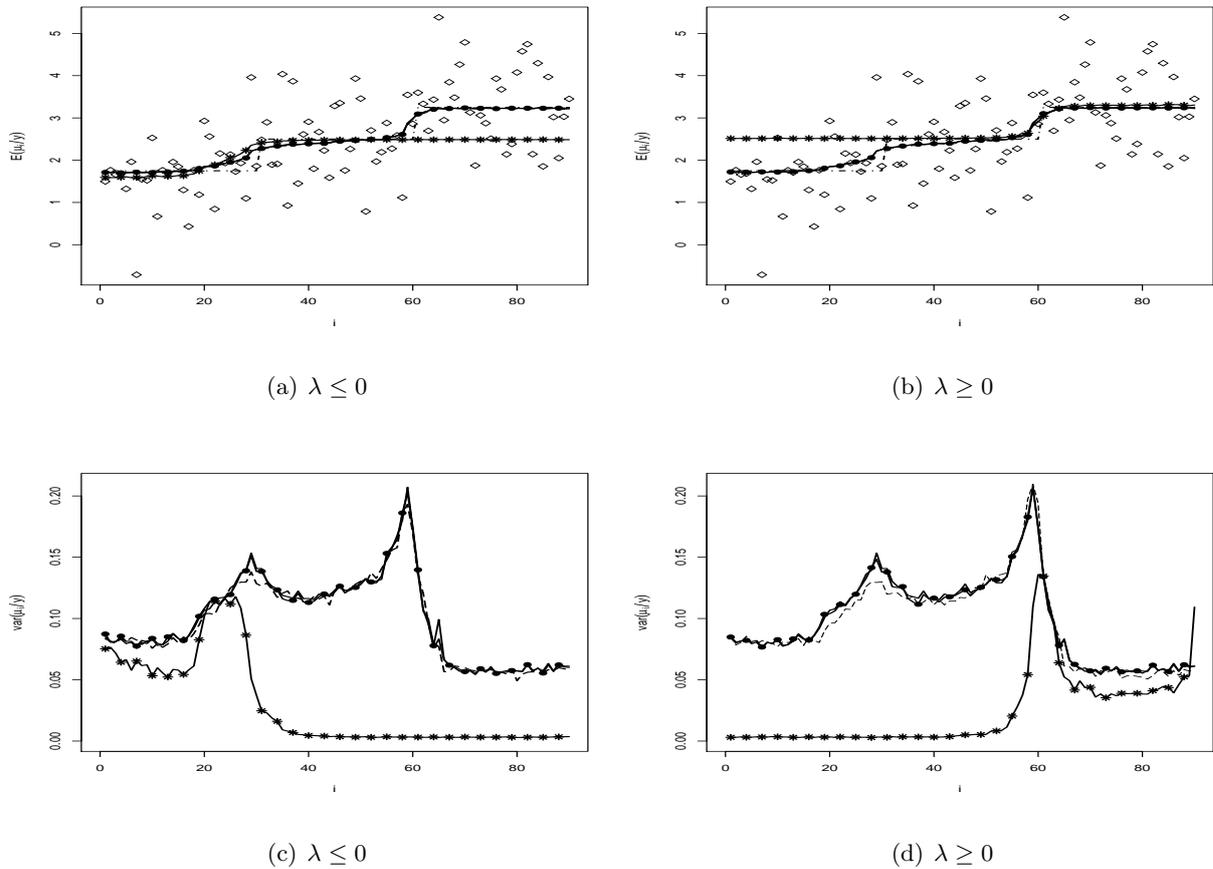


Figura 3.15: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 90$ .

a acompanhar melhor as médias verdadeiras, mas se  $\lambda = 100$  as médias do primeiro bloco são mal estimadas e se  $\lambda = -100$  a média do terceiro bloco tende a ser mal estimada. Os valores de  $\lambda = \pm 100$  suavizam demais as estimativas, conseqüentemente,  $\lambda = \pm 100$  produzem uma menor variação para as estimativas de  $\mu_i$ , como observa-se nas Figuras 3.14 e 3.15 (c) e (d).

As Figuras 3.16 e 3.17 mostram os resultados para o cenário tendência, nesse cenário a série possui 4 blocos, sendo que os dois blocos centrais possuem tendência determinística, mas a tendência do segundo bloco é mais suave que a tendência do terceiro bloco.

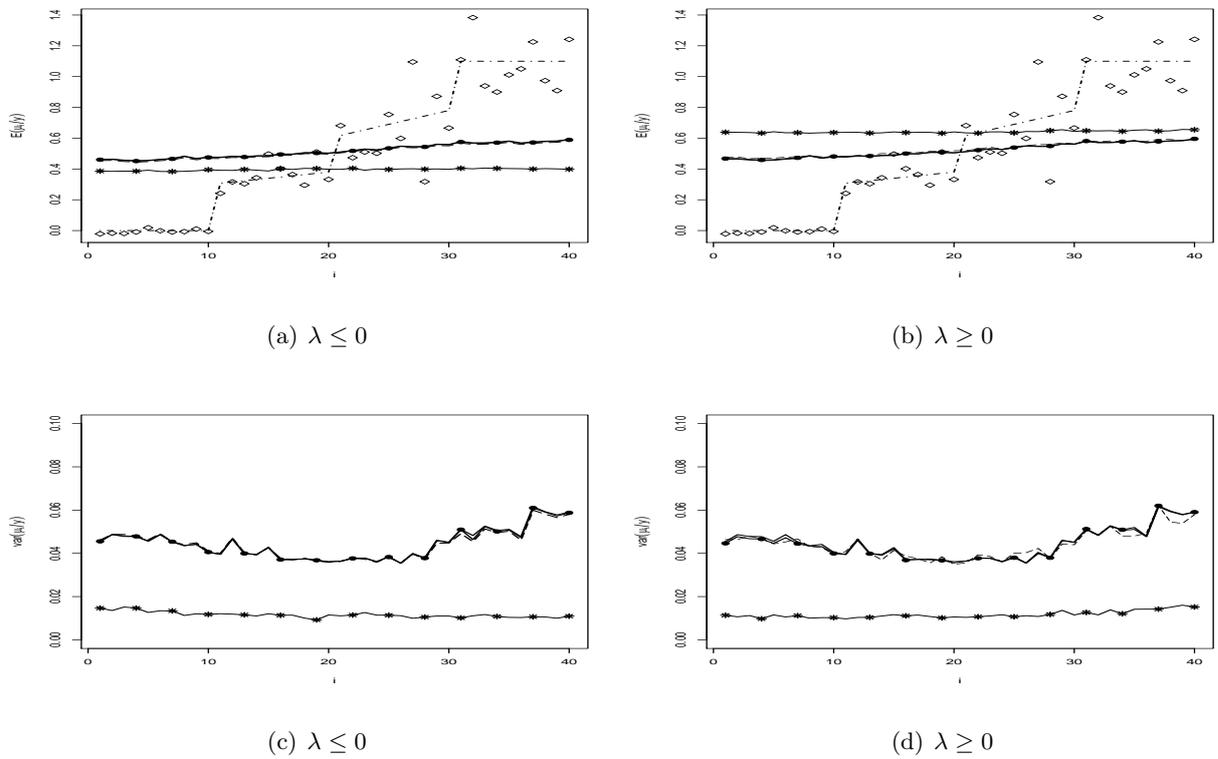


Figura 3.16: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 30$ .

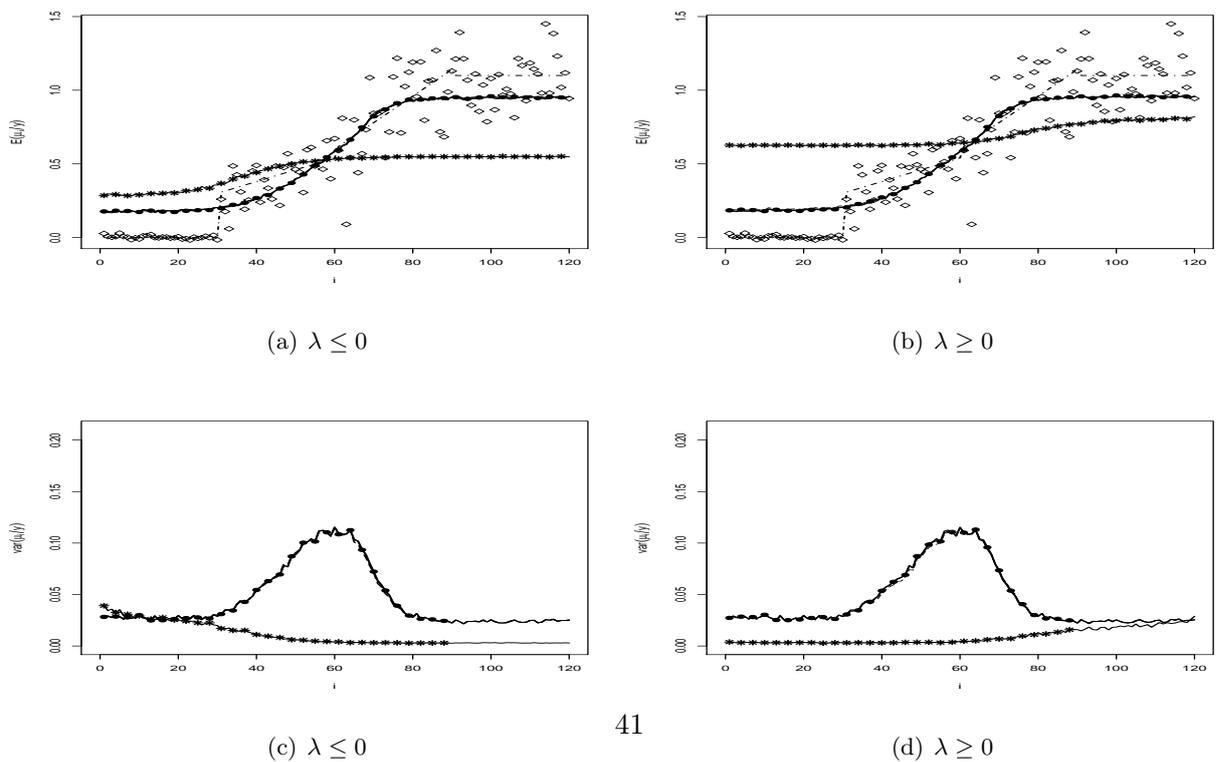


Figura 3.17: Estimativas produto (a) e (b) e variância (c) e (d) *a posteriori* para cada  $\mu_i$ . Dados simulados  $\diamond$ , média verdadeira (linha tracejada e pontilhada), estimativas para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 90$ .

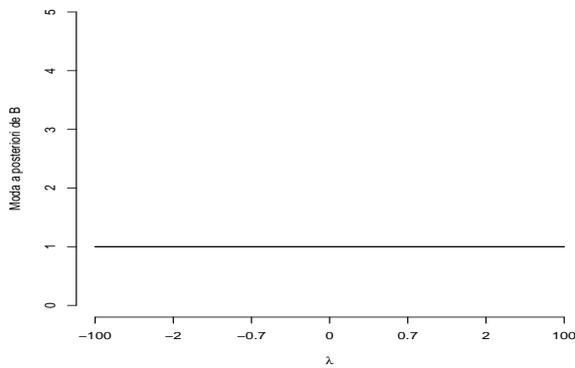
Nota-se no cenário tendência um comportamento similar ao cenário epsilon em que  $\lambda = \pm 100$  tende a fornecer as piores estimativas em relação aos demais valores de  $\lambda$  avaliados, mas se o tamanho da amostra aumenta as estimativas tendem a melhorar, mas não o suficiente para estimar bem as médias do primeiro e do último bloco. A variação para as estimativas de  $\mu_i$  e  $\lambda = \pm 100$  são maiores na presença de tendência determinística, ou seja, no segundo e terceiro bloco, com exceção de  $\lambda = \pm 100$  que suaviza demais as estimativas e conseqüentemente produz menor variação nas estimativas.

De forma geral, observa-se que as estimativas produto são piores para valores de  $\lambda = \pm 100$ , mesmo quando o tamanho da amostra foi maior, com exceção do cenário sem mudança. A medida que o tamanho da amostra aumenta as estimativas melhoram quando  $\lambda = 0, \pm 0.7$  e  $\pm 2$ . As variâncias de  $\mu_i$  para os valores de  $\lambda = \pm 100$  apresentam as maiores discrepâncias no cenário degrau não simétrico e escada. Nota-se que a variabilidade *a posteriori* da média é maior próximo aos pontos de mudança, como também é observado em Loschi et al. (2010) indicando que essa é uma característica do modelo partição produto.

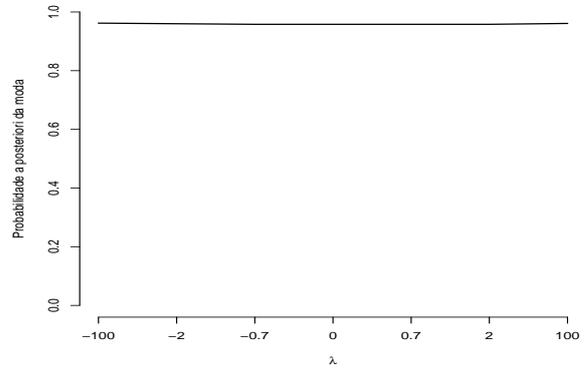
### 3.1.2 A influência de $\lambda$ e $n$ nas inferências sobre $\rho$ e $B$

Esta seção destina-se a mostrar através da sensibilidade informal como as estimativas *a posteriori* para o número de blocos e a partição mais provável são afetadas pelo parâmetro de assimetria da distribuição *a priori* das médias. As figuras desta seção são divididas em (a), (b), (c) e (d). Os gráficos (a) e (b) destinam-se, respectivamente, a ilustrar a moda *a posteriori* do número de blocos para cada  $\lambda$  avaliado e a probabilidade *a posteriori* atribuída a tal medida para cada valor de  $\lambda$ . Já os gráficos (c) e (d) representam, nessa ordem, a partição mais provável indicada para cada valor de  $\lambda$  e a probabilidade com que a partição mais provável foi indicada.

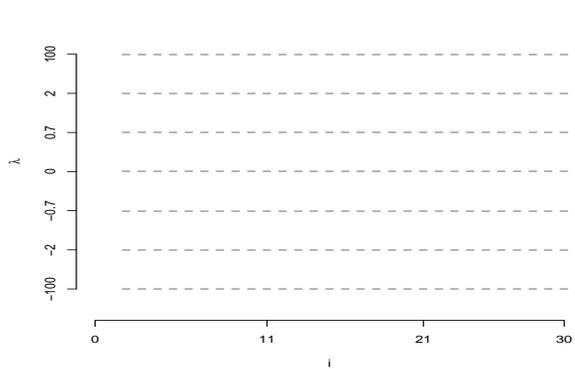
As Figuras 3.18 e 3.19 apresentam os resultados do cenário sem mudança quando  $n = 30$  e  $n = 90$ , respectivamente.



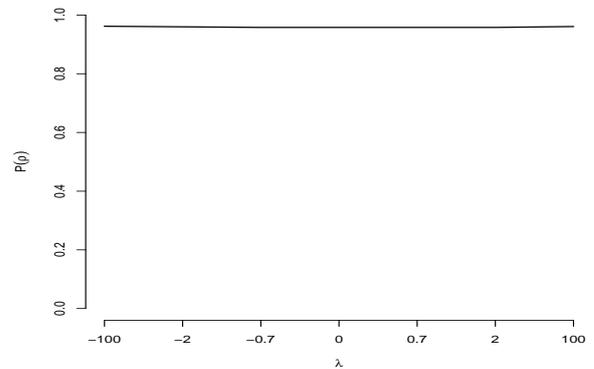
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$



(c) Partição mais provável= $\rho'$

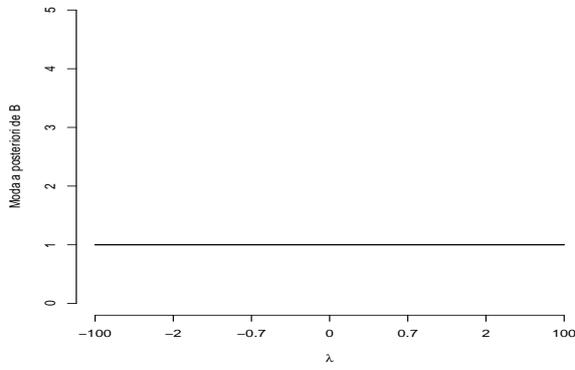


(d)  $P(\rho = \rho')$

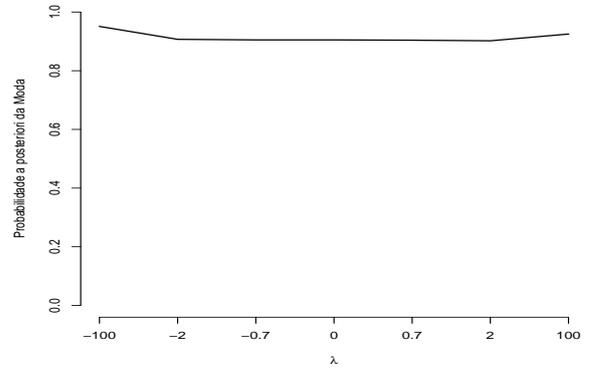
Figura 3.18: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário sem mudança,  $n = 30$ .

No cenário sem mudança, observa-se na Figura 3.18 que não houve discrepância na inferência *a posteriori* sobre o número de blocos para os valores de  $\lambda$  considerados nesse estudo de sensibilidade. Neste caso, o número correto de blocos ( $B = 1$ ) foi identificado com a mesma probabilidade *a posteriori* em todos os casos para  $n = 30$ . Logo, não foram observadas mudanças na inferência sobre a partição para  $n = 30$ . Conseqüentemente, a partição indicada como mais provável *a posteriori* foi a sem mudança e essa partição foi indicada com a mesma probabilidade para todos os valores de  $\lambda$ , assumindo tamanho de amostra  $n = 30$ . Na Figura 3.19 que ilustra o cenário sem mudança para o tamanho de amostra 90,  $\lambda = \pm 100$  indicaram o número de blocos e a partição corretamente

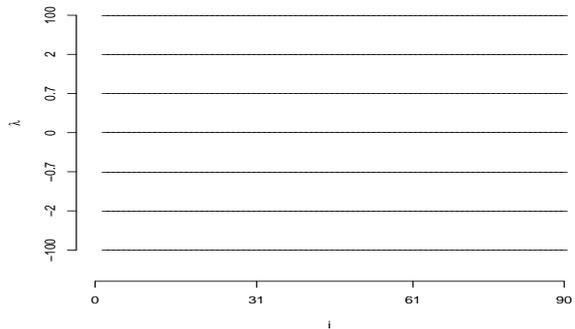
com uma probabilidade um pouco maior que os demais valores de  $\lambda$  avaliados.



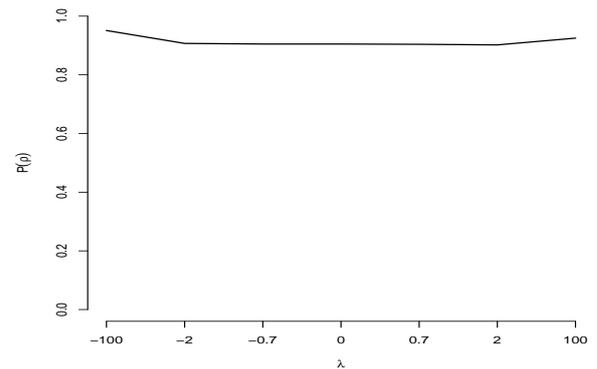
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$



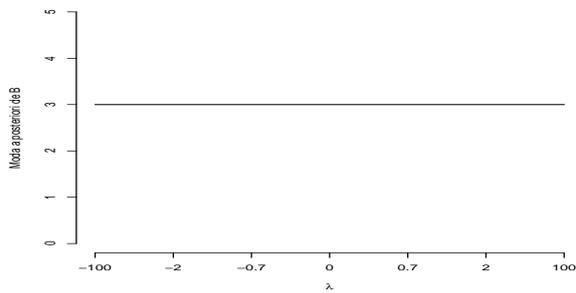
(c) Partição mais provável =  $\rho'$



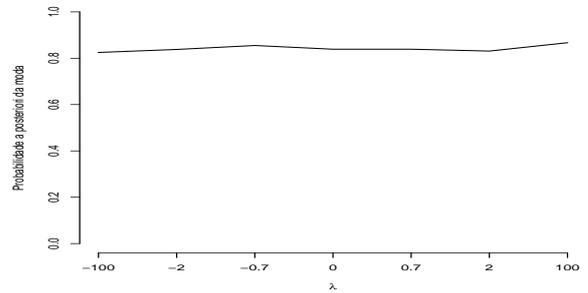
(d)  $P(\rho = \rho')$

Figura 3.19: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário sem mudança,  $n = 90$ .

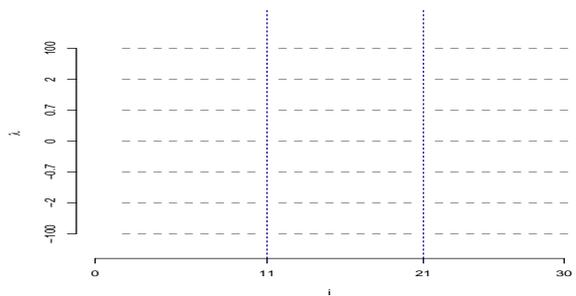
As Figuras 3.20 e 3.21 mostram os resultados do número de blocos e partição mais provável para o cenário degrau não simétrico.



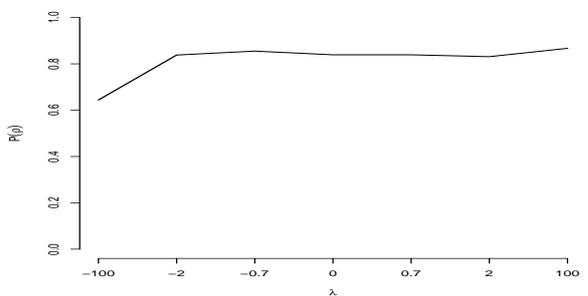
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$

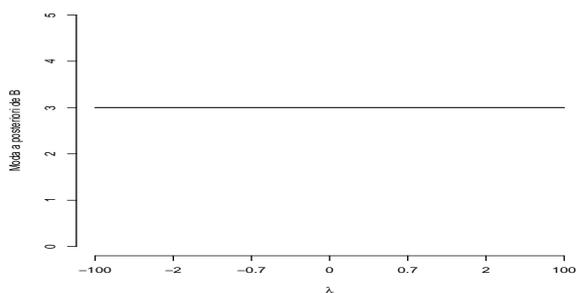


(c) Partição mais provável= $\rho'$

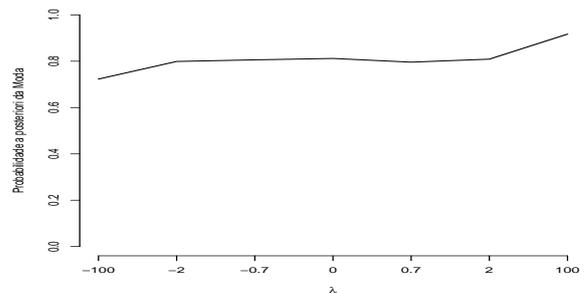


(d)  $P(\rho = \rho')$

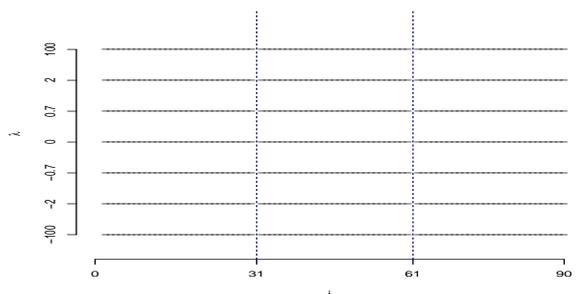
Figura 3.20: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário degrau não simétrico,  $n = 30$ .



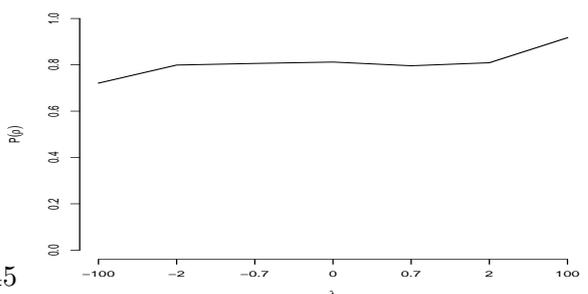
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$



(c) Partição mais provável= $\rho'$

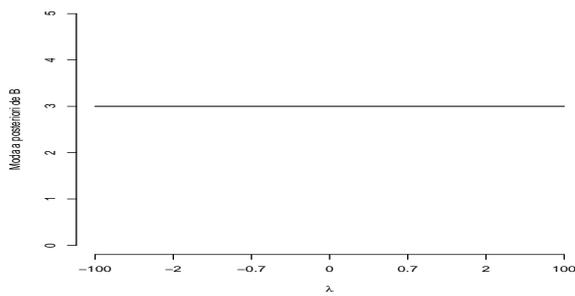


(d)  $P(\rho = \rho')$

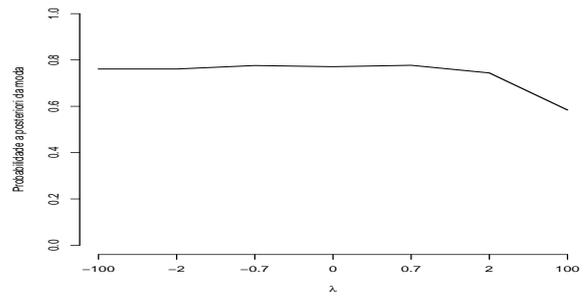
Figura 3.21: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário degrau não simétrico,  $n = 90$ .

Nas Figuras 3.20 e 3.21 que ilustram o cenário degrau não simétrico é possível observar que, para todos os valores de  $\lambda$ , o modelo indicou corretamente a partição e o número de blocos. No entanto, observa-se diferença nas probabilidades *a posteriori* obtidas para tais parâmetros. As maiores probabilidades *a posteriori* foram observadas para  $\lambda = 100$ , mas nas Figuras 3.8 e 3.9 da Seção 3.1.1 foi discutido que as estimativas produto foram comprometidas para valores de  $\lambda = \pm 100$ , ou seja, uma melhora nas estimativas do número de blocos e na partição de interesse é observada, no entanto, as estimativas das médias não são satisfatórias.

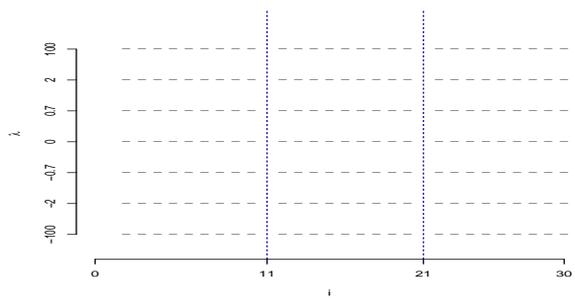
Os resultados da análise de sensibilidade para o cenário degrau positivo e negativo estão ilustrados nas Figuras 3.22, 3.23, 3.24 e 3.25 a seguir.



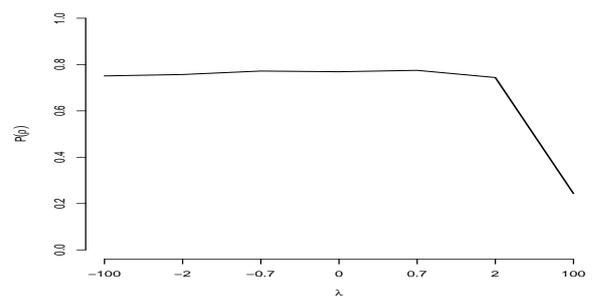
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$

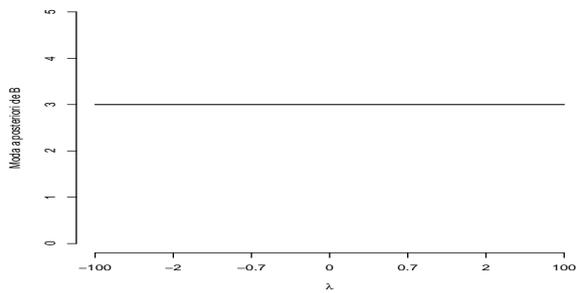


(c) Partição mais provável= $\rho'$

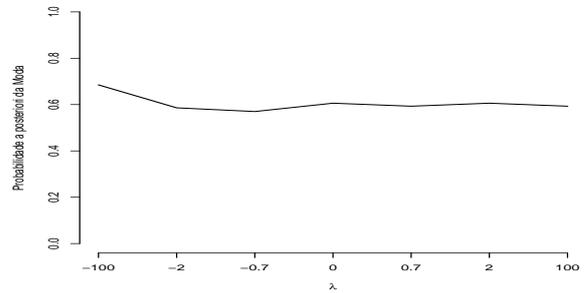


(d)  $P(\rho = \rho')$

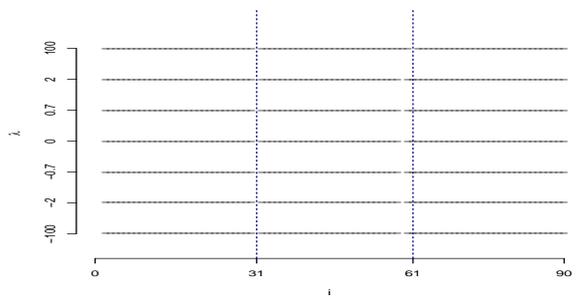
Figura 3.22: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário degrau negativo,  $n = 30$ .



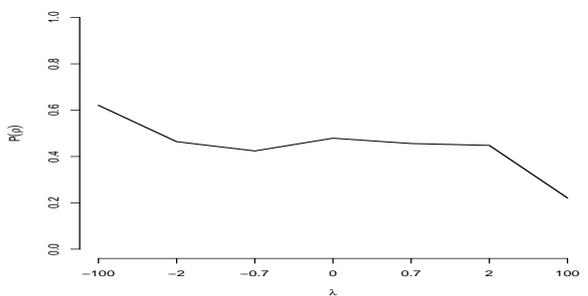
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$

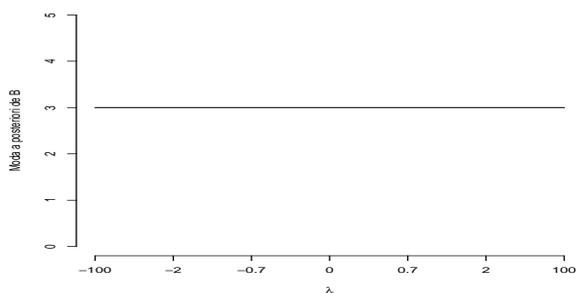


(c) Partição mais provável= $\rho'$

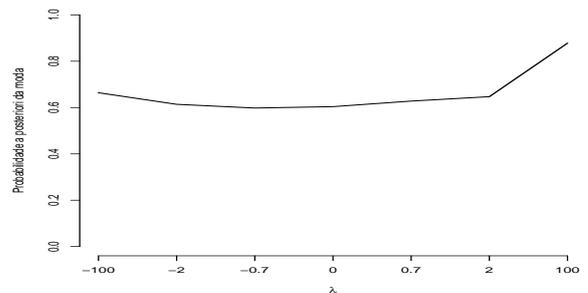


(d)  $P(\rho = \rho')$

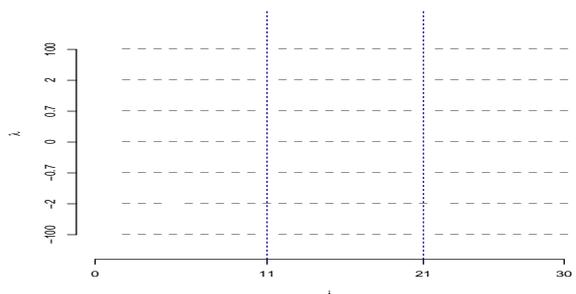
Figura 3.23: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário degrau negativo,  $n = 90$ .



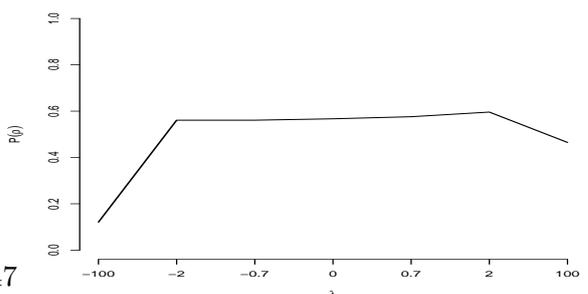
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$

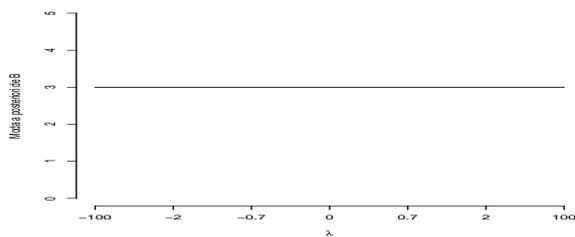


(c) Partição mais provável= $\rho'$

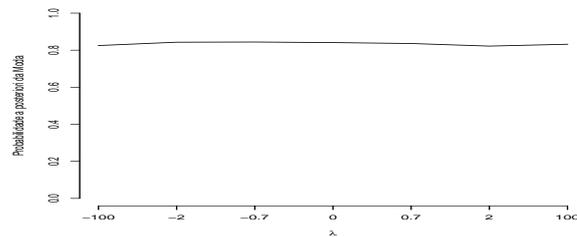


(d)  $P(\rho = \rho')$

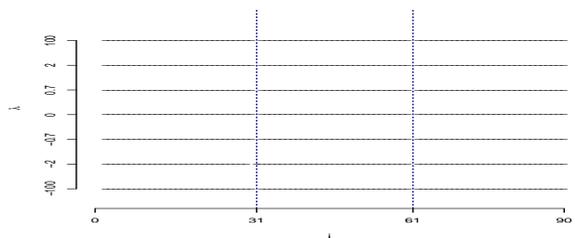
Figura 3.24: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário degrau positivo,  $n = 30$ .



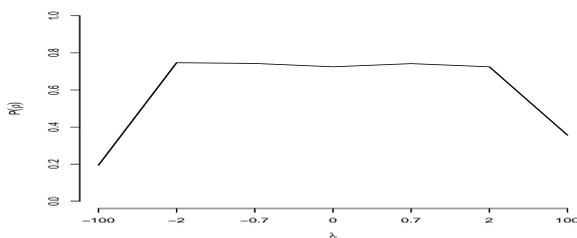
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$



(c) Partição mais provável= $\rho'$

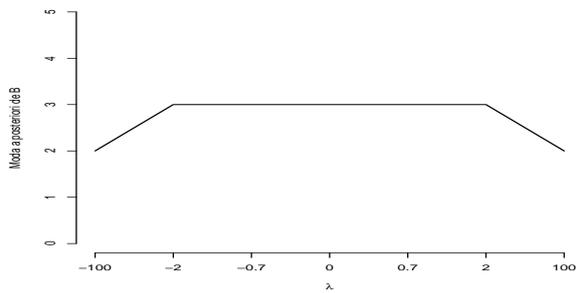


(d)  $P(\rho = \rho')$

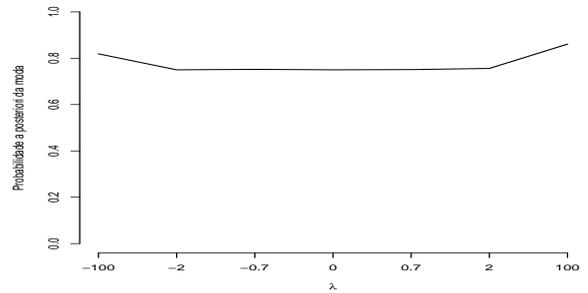
Figura 3.25: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário degrau positivo,  $n = 90$ .

Ao analisar a partição mais provável e o número de blocos *a posteriori*, no cenário degrau negativo, nota-se que para  $\lambda = 100$  observam-se as menores probabilidades independente dos tamanhos de amostra considerados nesse estudo, no entanto,  $\lambda = 100$  na Figura 3.23 é o único que identifica a partição corretamente. A Figura 3.22 ilustra que mesmo quando os dados são negativos um valor de  $\lambda = -100$  não gerou a maior probabilidade *a posteriori* para a partição mais provável, além disso, nas na Figura 3.23(c) é possível observar que para  $\lambda = (-100, -2, -0.7, 0, 0.7, 2)$  o modelo indica erroneamente a partição mais provável ( $\rho' = \{0, 30, 58\}$ ), enquanto que para  $\lambda = 100$  a partição é corretamente indicada  $\rho' = \{0, 30, 60\}$ , no entanto, com probabilidade menor. Note que no cenário degrau positivo, Figuras 3.24 e 3.25 que os valores de  $\lambda = \pm 0.7, \pm 2$  e  $\lambda = 0$  geram estimativas próximas.

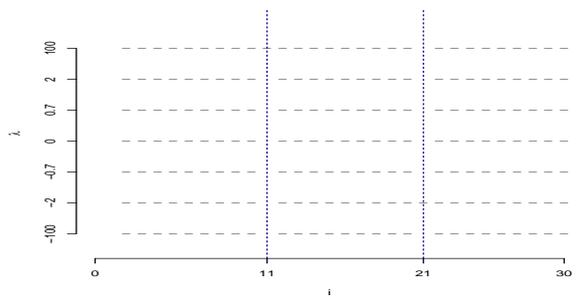
Os resultados para o cenário escada é dado nas Figuras 3.26, 3.27, 3.28, 3.29 a seguir.



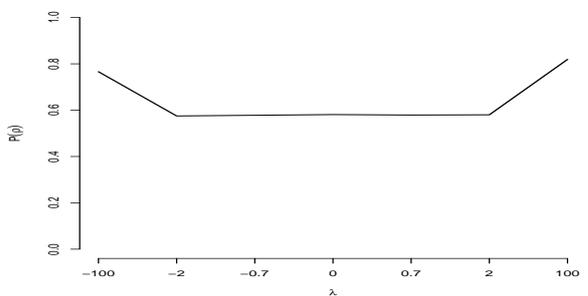
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$  -  $n=30$

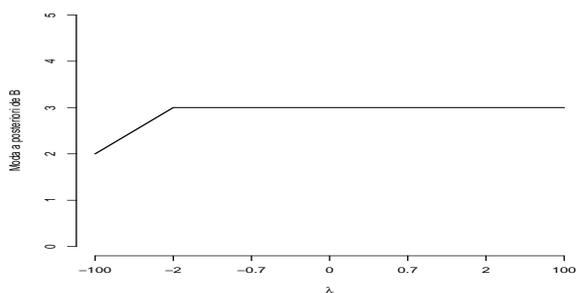


(c) Partição mais provável= $\rho'$

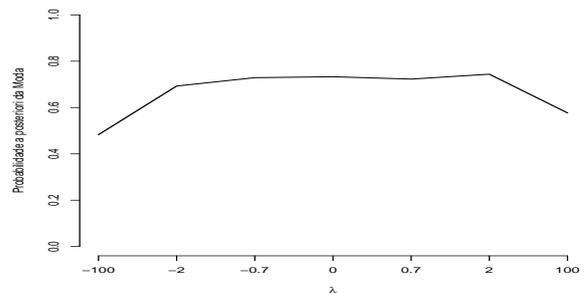


(d)  $P(\rho = \rho')$

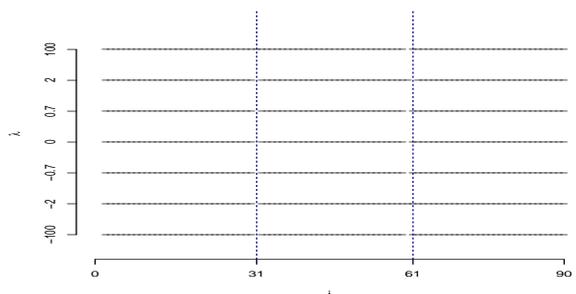
Figura 3.26: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário escada crescente,  $n = 30$ .



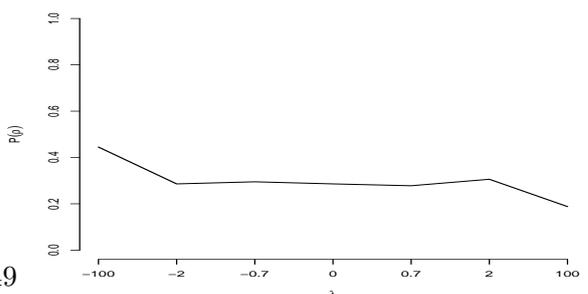
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$  -  $n=90$

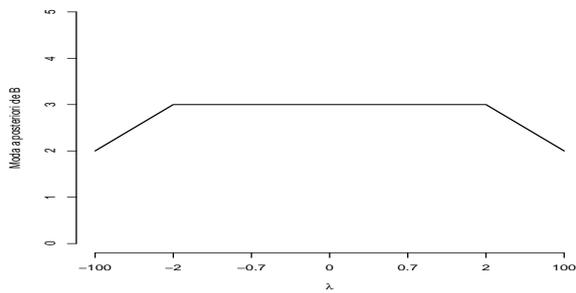


(c) Partição mais provável= $\rho'$

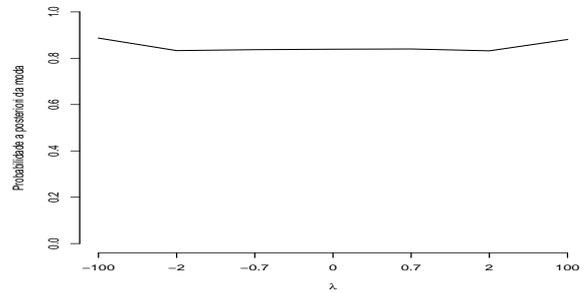


(d)  $P(\rho = \rho')$

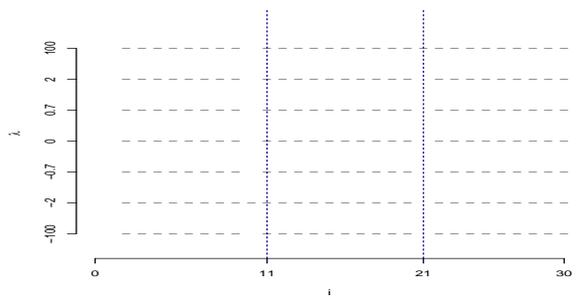
Figura 3.27: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário escada crescente,  $n = 90$ .



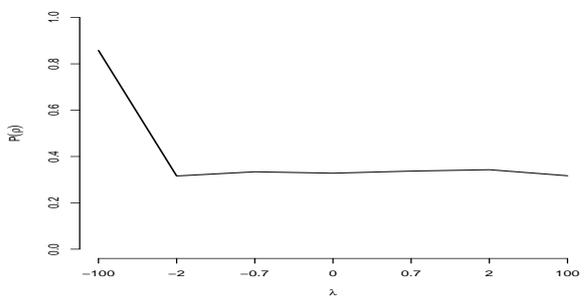
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$

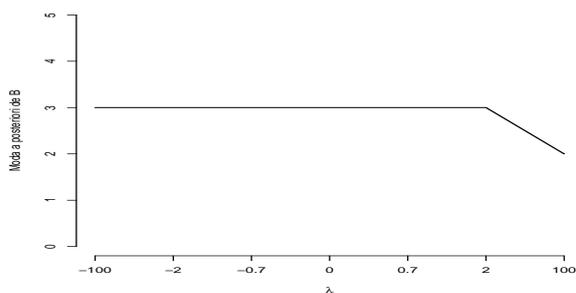


(c) Partição mais provável= $\rho'$

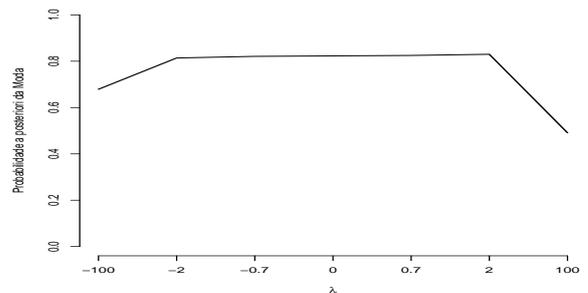


(d)  $P(\rho = \rho')$

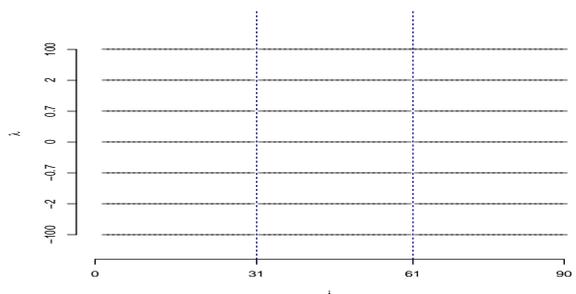
Figura 3.28: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário escada decrescente,  $n = 30$ .



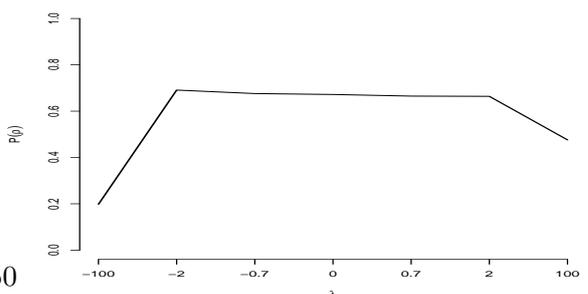
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$



(c) Partição mais provável= $\rho'$



(d)  $P(\rho = \rho')$

Figura 3.29: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário escada decrescente,  $n = 90$ .

Em relação ao número de blocos do cenário escada, quando o tamanho da amostra é 30, Figuras 3.26 e 3.28,  $\lambda = \pm 100$  resultou na maior probabilidade *a posteriori* da partição apresentar 2 blocos, no entanto, o número correto de blocos é 3. No cenário escada decrescente para  $n = 90$ ,  $\lambda = 100$  indica que o número de blocos é 2 e se o cenário é crescente  $\lambda = -100$  também indica que o número de blocos é 2.

Em contrapartida, no cenário decrescente e tamanho de amostra 90 apesar de  $\lambda = -100$  indicar corretamente o número de blocos, indica com menor probabilidade, além de fornecer estimativas ruins para as médias dos blocos como visto nas Figuras 3.13, o mesmo acontece para  $\lambda = 100$  no cenário crescente.

As Figuras 3.30 e 3.31 mostram o comportamento do número de blocos no cenário epsilon, onde a diferença entre os blocos é de apenas 0.75.

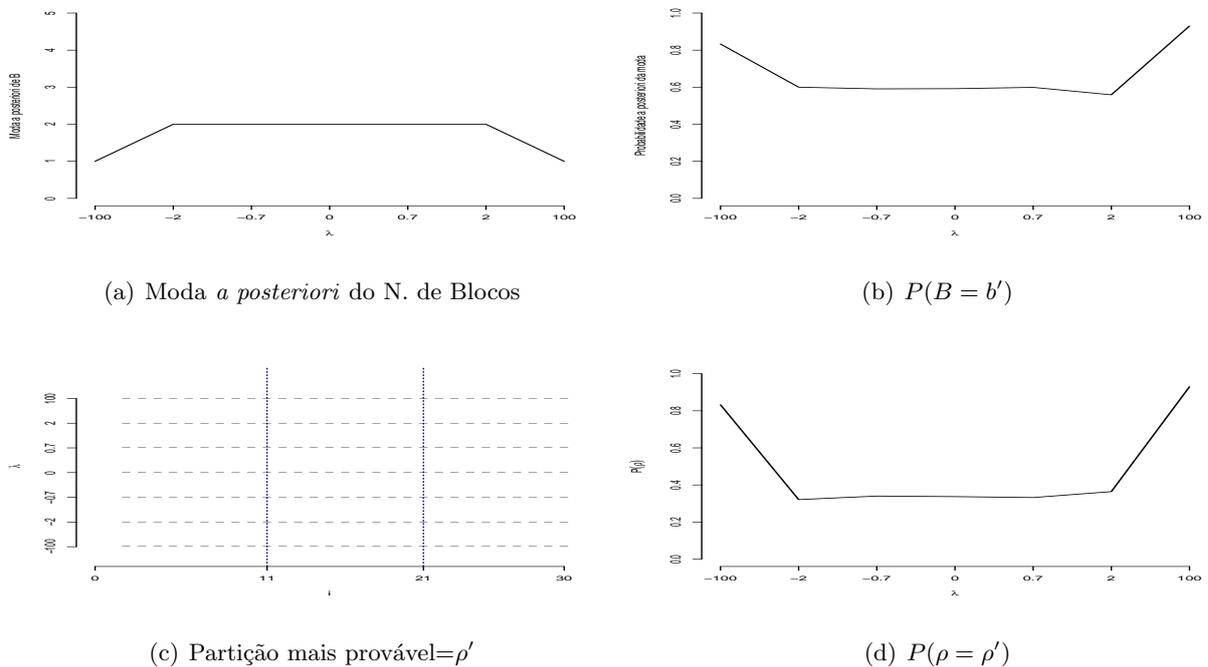
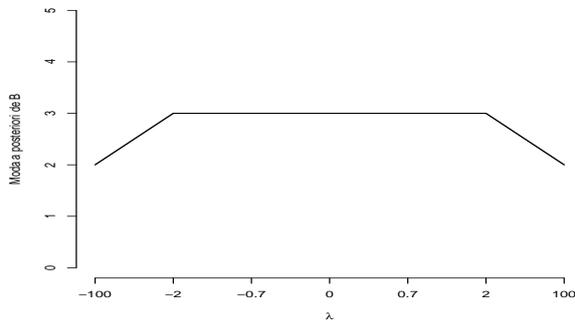
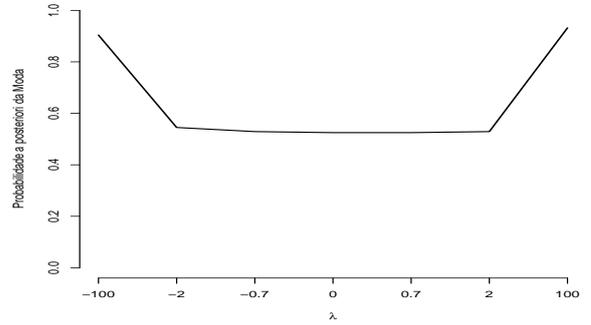


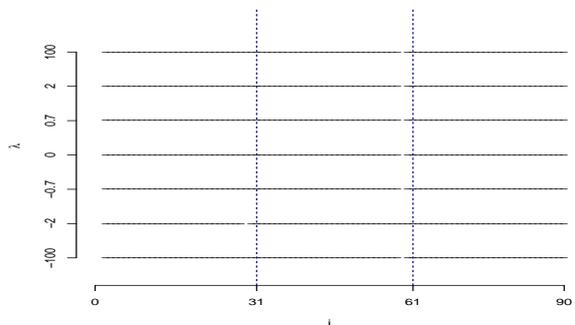
Figura 3.30: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário epsilon,  $n = 30$ .



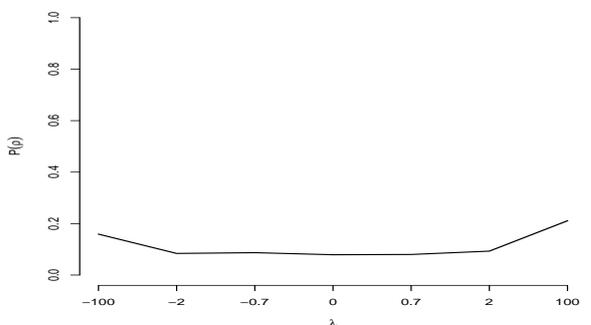
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$



(c) Partição mais provável= $\rho'$

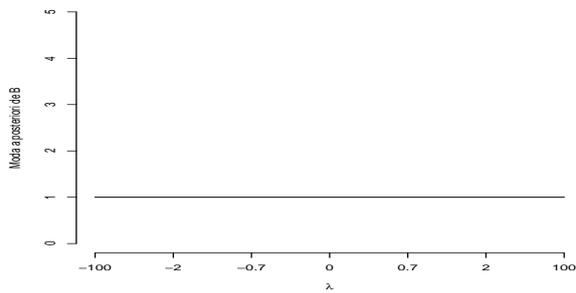


(d)  $P(\rho = \rho')$

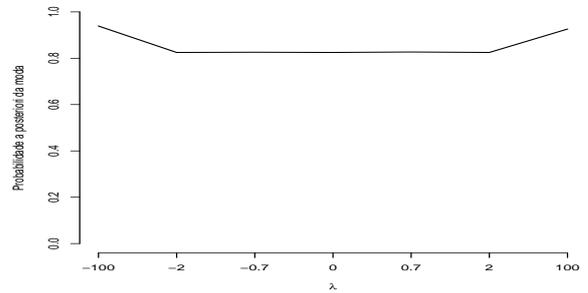
Figura 3.31: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário epsilon,  $n = 90$ .

Observa-se que o número de blocos é melhor estimado para  $\lambda = (\pm 2, \pm 0.7, 0)$  e tamanho de amostra 90, pois nesse caso o número de blocos é estimado corretamente e  $B = 3$  é indicado com 55% de probabilidade, mas no caso de  $\lambda = \pm 100$  ou o tamanho de amostra  $n = 30$  o número de blocos é subestimado.

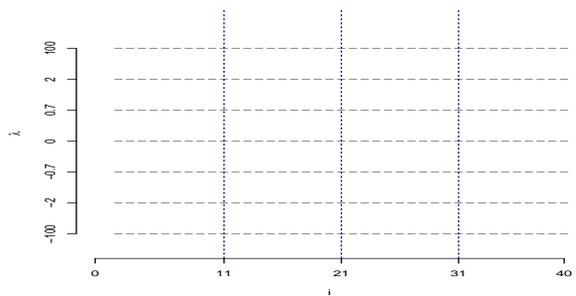
Os resultados em relação ao número de blocos para o cenário tendência estão representados nas Figuras 3.32 e 3.33 e esse cenário possui uma tendência determinística no segundo e no terceiro bloco.



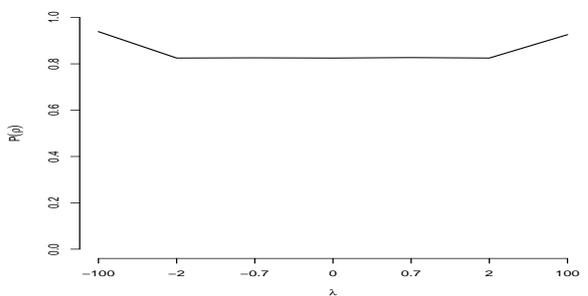
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$

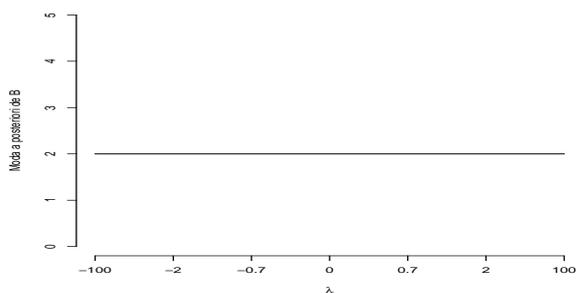


(c) Partição mais provável= $\rho'$

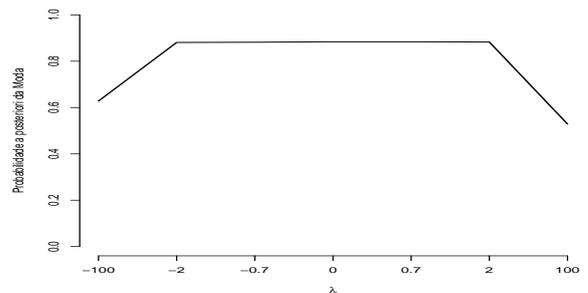


(d)  $P(\rho = \rho')$

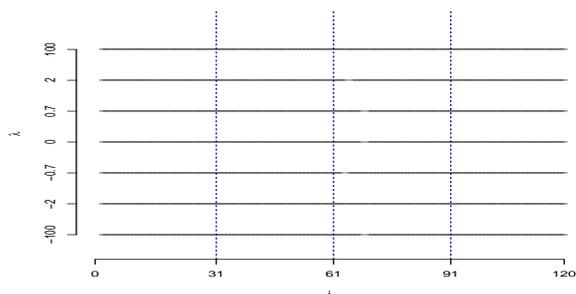
Figura 3.32: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário tendência,  $n = 30$ .



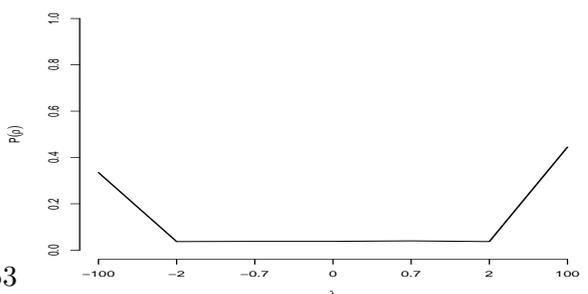
(a) Moda *a posteriori* do N. de Blocos



(b)  $P(B = b')$



(c) Partição mais provável= $\rho'$



(d)  $P(\rho = \rho')$

Figura 3.33: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ , cenário tendência,  $n = 90$ .

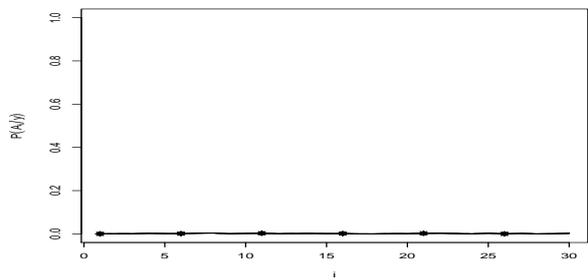
Nota-se nesse cenário que um único bloco é estimado com alta probabilidade, se o tamanho da amostra aumenta o modelo indica dois blocos, mas o verdadeiro número de blocos é 4, logo o modelo não é capaz de retratar o número correto de blocos e isso ocorre independente do valor de  $\lambda$  e tamanho de amostra.

Ao avaliar os resultados para o número de blocos, conclui-se que o número de blocos é bem estimado mesmo quando se assume uma assimetria muito forte. Com exceção do cenário escada que tende a subestimar o número de blocos para  $\lambda = \pm 100$ . Em relação a partição mais provável nota-se que para o cenário degrau positivo,  $\lambda = \pm 100$  indicou a partição mais provável com menor probabilidade e no cenário escada para  $\lambda = \pm 100$  a partição possui apenas um ponto de mudança. O cenário sem mudança foi o único que não apresentou grandes diferenças nas probabilidades *a posteriori* da partição mais provável e do número de blocos mais provável.

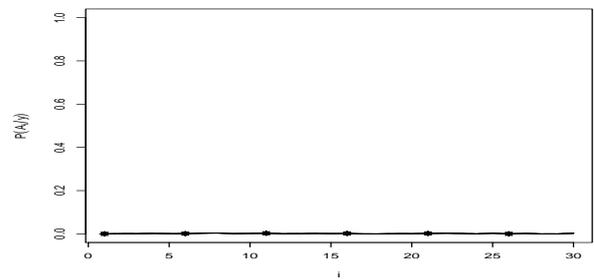
### 3.1.3 A influência de $\lambda$ e $n$ nas inferências sobre $P(A_t)$

Esta seção destina-se a mostrar os resultados obtidos em relação ao parâmetro de assimetria e o tamanho de amostra para a probabilidade de mudança em cada instante para os mesmos cenários avaliados até agora, cenário sem mudança, cenário degrau e cenário escada.

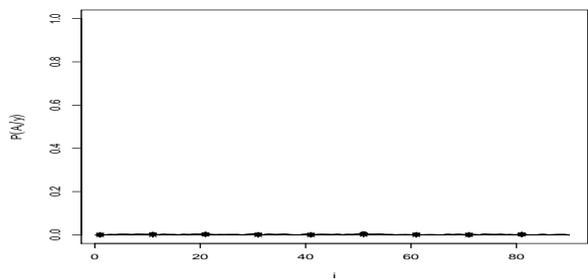
Nessa Seção os resultados são apresentados da seguinte forma, para valores de  $\lambda \leq 0$  as estimativas para cada instante ser um ponto de mudança encontram-se do lado esquerdo das figuras, e os resultados para  $\lambda \geq 0$  estão do lado direito das figuras. As Figuras 3.34 e 3.35, a seguir, apresentam os resultados para o cenário sem mudança e degrau não simétrico, respectivamente.



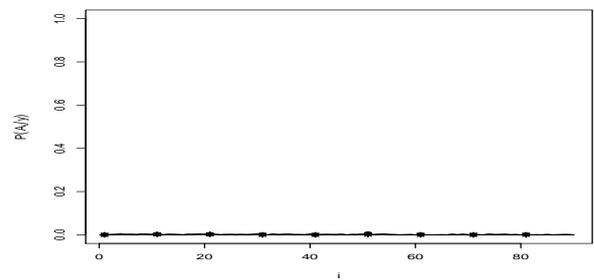
(a)  $\lambda \leq 0 - n = 30$



(b)  $\lambda \geq 0 - n = 30$

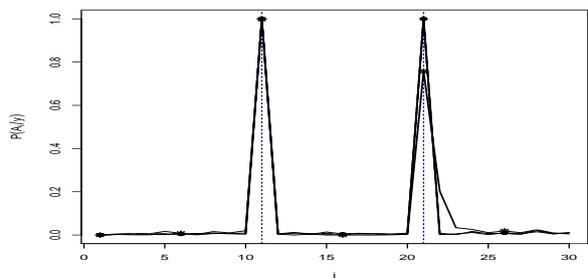


(c)  $\lambda \leq 0 - n = 90$

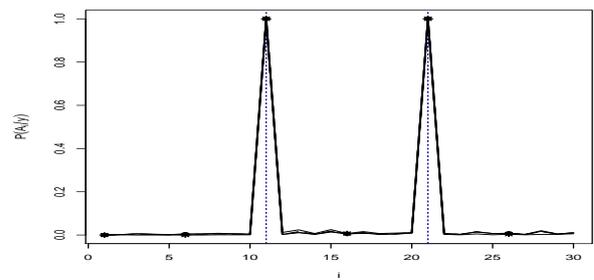


(d)  $\lambda \geq 0 - n = 90$

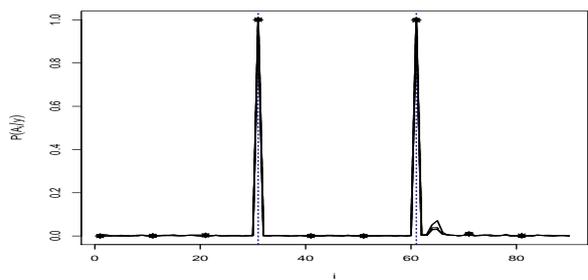
Figura 3.34: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 30$  e  $n = 90$ .



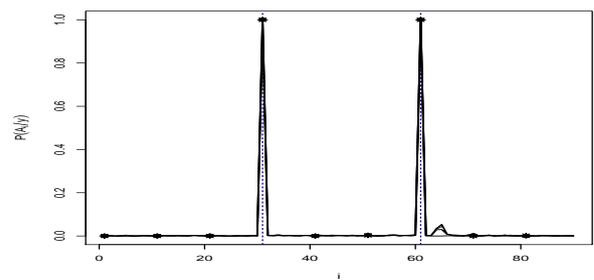
(a)  $\lambda \leq 0 - n = 30$



(b)  $\lambda \geq 0 - n = 30$



(c)  $\lambda \leq 0 - n = 90$



(d)  $\lambda \geq 0 - n = 90$

Figura 3.35: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 30$  e  $n = 90$ .

A Figura 3.34 mostra que para todos os os valores de  $\lambda$  considerados no estudo de sensibilidade informal, a probabilidade de mudança em cada instante foi bastante semelhante indicando que neste cenário sem mudança a escolha equivocada do valor de  $\lambda$  não influencia nas estimativas da probabilidade de mudança em cada instante. No cenário degrau não simétrico a inferência sobre cada instante ser mudança não foi substancialmente influenciada pelos diferentes valores de  $\lambda$ , Figura 3.35. Observa-se também, no cenário degrau não simétrico, Figura 3.35, que para  $n = 30$  a probabilidade do ponto 21 ser um ponto de mudança é menor para  $\lambda = -100$ , mas nos demais casos o comportamento é similar para todos os valores do parâmetro de assimetria avaliados.

A probabilidade de mudança em cada instante para os cenários degrau positivo e negativo são apresentadas nas Figuras 3.36 e 3.37.

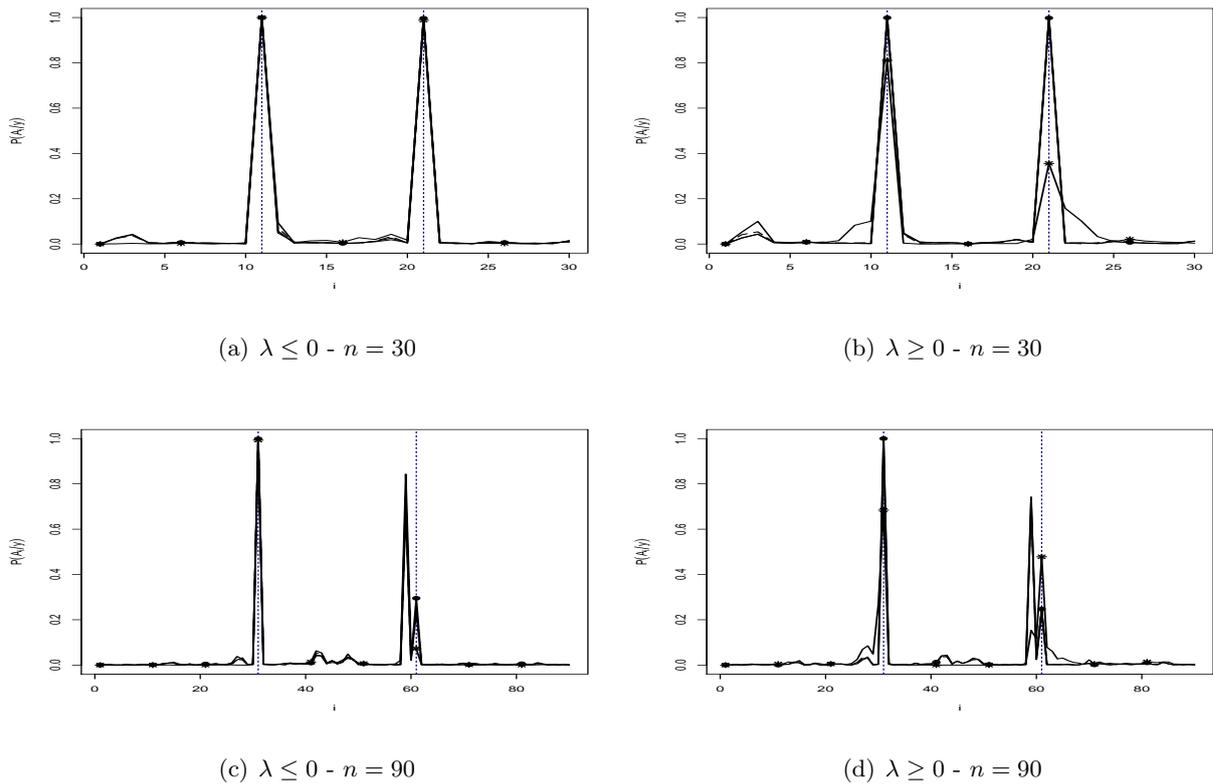


Figura 3.36: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 30$  e  $n = 90$ .

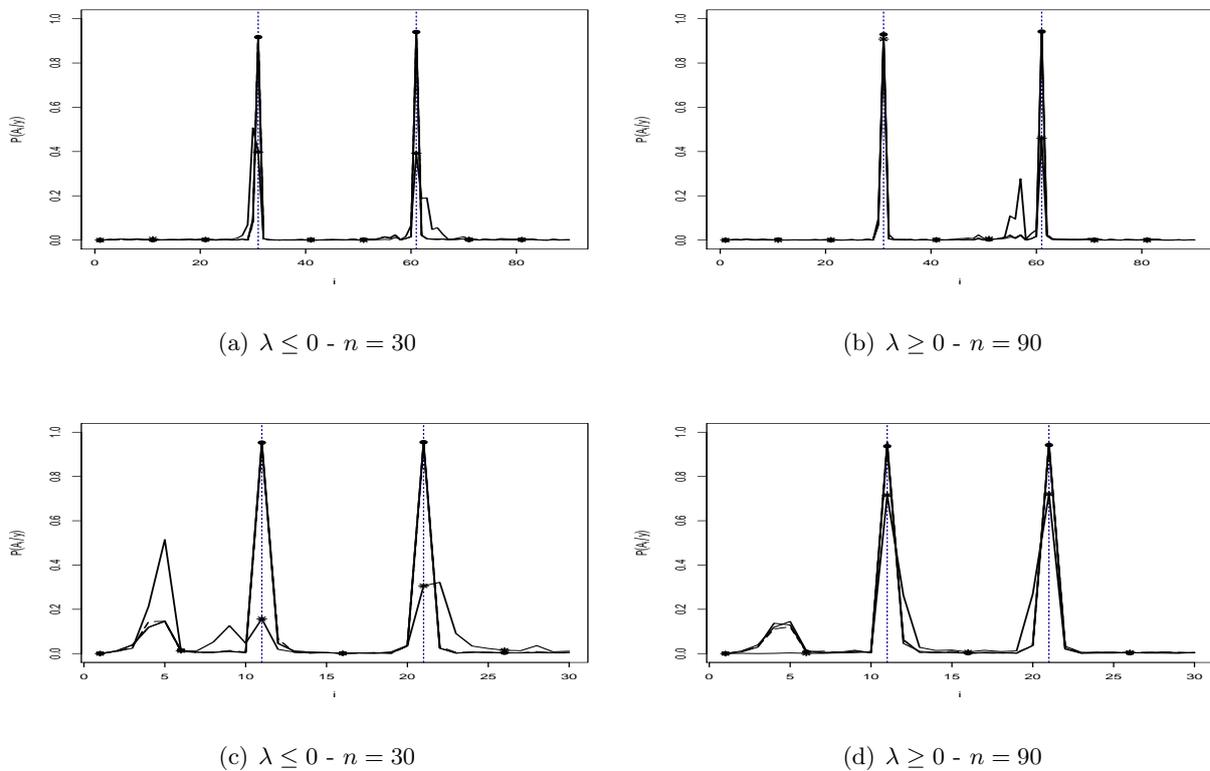
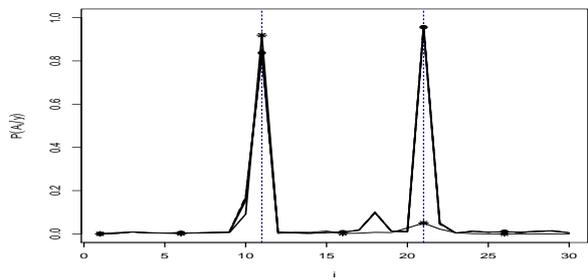


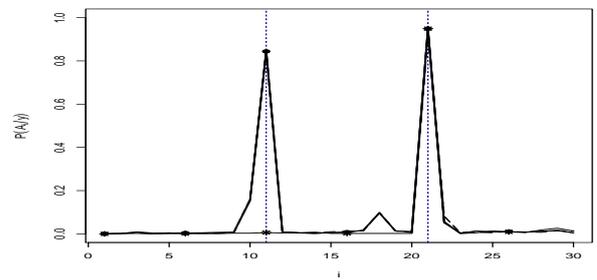
Figura 3.37: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 30$  e  $n = 90$ .

No cenário degrau negativo, a Figura 3.36 mostra que a probabilidade de mudança em cada instante, se  $\lambda = 100$ , é afetada para o tamanho de amostra 30, apesar do modelo indicar o instante correto da mudança. A maior discrepância das probabilidades de mudança em cada instante ocorrem quando  $n = 90$  para valores positivos e negativos de  $\lambda$ , pois o modelo indica como pontos de mudança os instantes 31, 59 e 61, quando o correto seria o 31 e 61. O cenário degrau positivo, Figura 3.37, apresenta comportamento similar ao degrau negativo, Figura 3.36, com exceção das estimativas para  $n = 30$  que indicam probabilidade de mudança em torno dos instantes 4 e 5, inclusive se  $\lambda = 0$ , apenas para  $\lambda = 100$  o modelo não indica probabilidade desses pontos serem pontos de mudança.

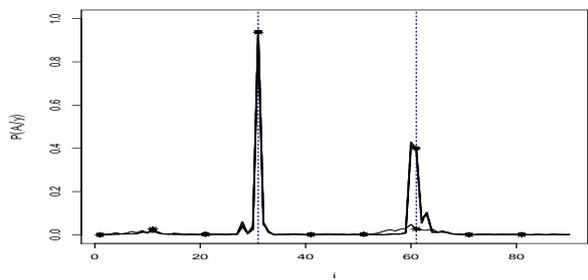
Os resultados para o cenário escada são apresentados nas Figuras 3.38 e 3.39.



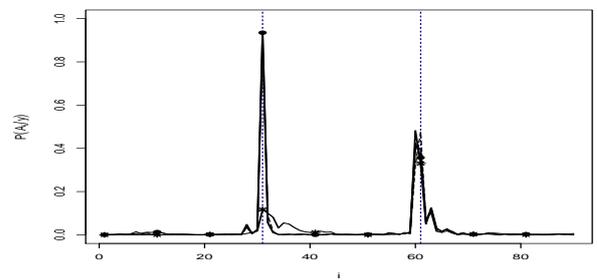
(a)  $\lambda \leq 0 - n = 30$



(b)  $\lambda \geq 0 - n = 30$

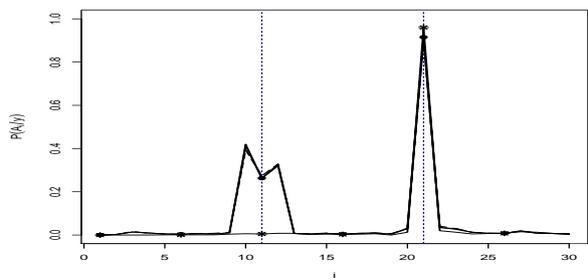


(c)  $\lambda \leq 0 - n = 90$

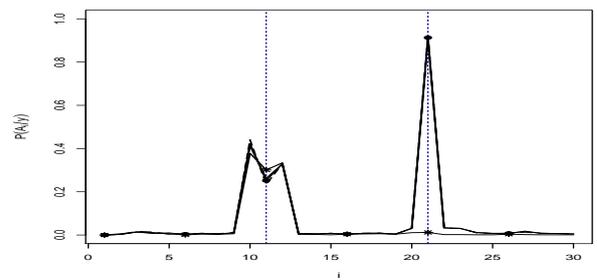


(d)  $\lambda \geq 0 - n = 90$

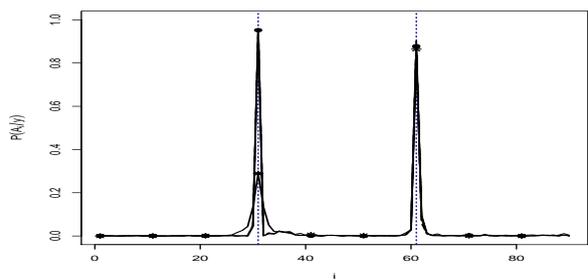
Figura 3.38: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 30$  e  $n = 90$ .



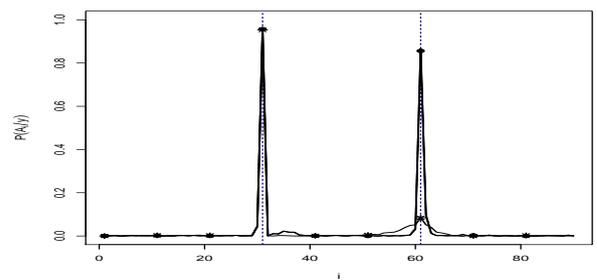
(a)  $\lambda \leq 0 - n = 30$



(b)  $\lambda \geq 0 - n = 30$



(c)  $\lambda \leq 0 - n = 90$



(d)  $\lambda \geq 0 - n = 90$

Figura 3.39: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 30$  e  $n = 90$ .

As probabilidades de mudança em cada instante para o cenário escada crescente mostradas na Figura 3.38 apresentam maior discrepância nos pontos 21 ( $n = 30$ ) e 61 ( $n = 90$ ) quando  $\lambda < 0$  e nos pontos 11 ( $n = 30$ ) e 31 ( $n = 90$ ) quando  $\lambda > 0$ . No cenário escada decrescente, Figura 3.39, as maiores discrepâncias foram verificadas nos mesmos pontos, porém com os valores de  $\lambda$  invertidos.

Os resultados para a probabilidade de mudança em cada instante para o cenário epsilon e para o cenário tendência são apresentados nas Figuras 3.40 e 3.41, respectivamente.

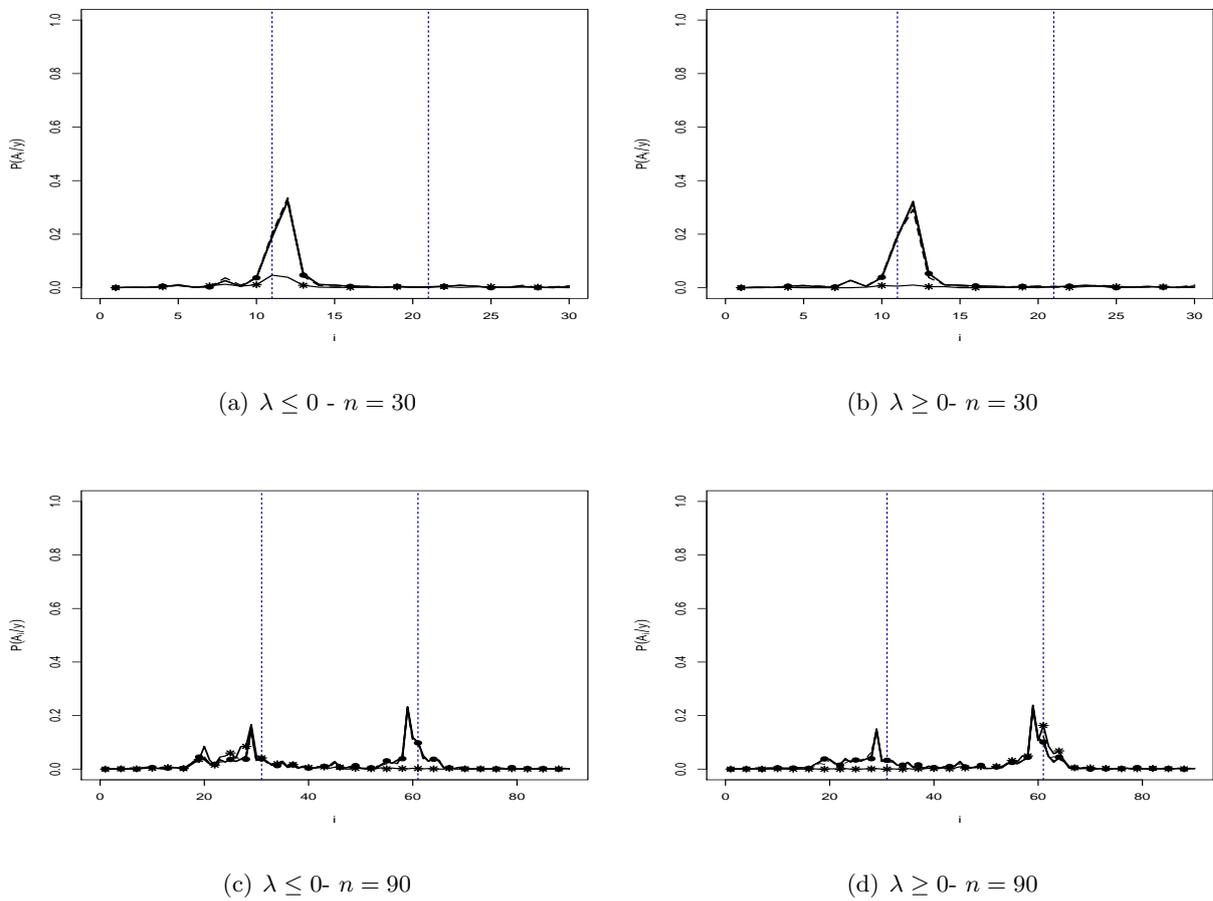
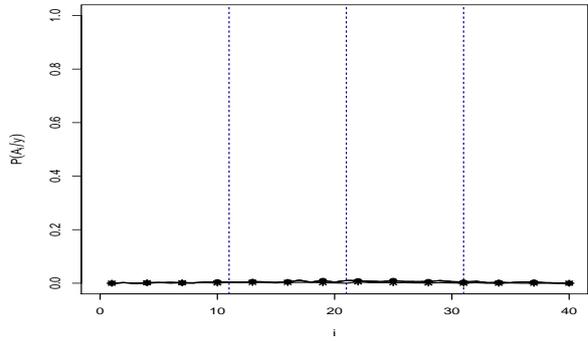
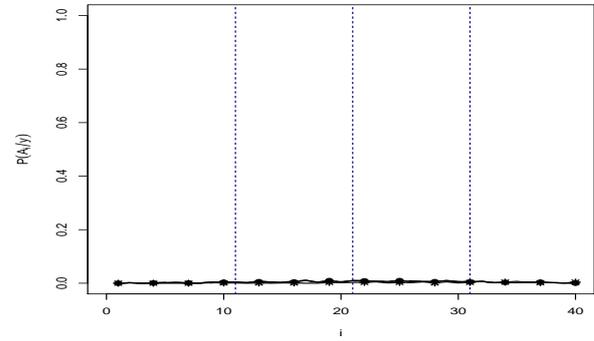


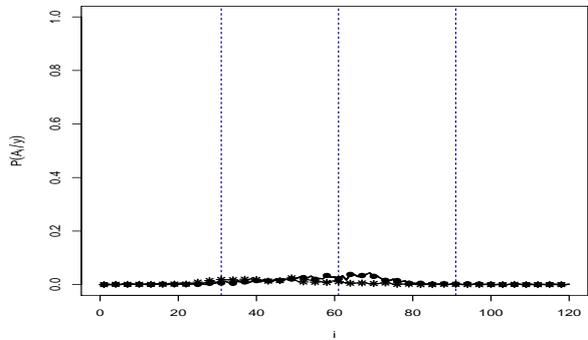
Figura 3.40: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 30$  e  $n = 90$ .



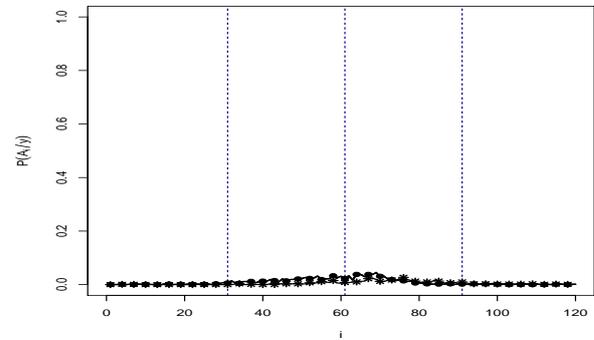
(a)  $\lambda \leq 0 - n = 40$



(b)  $\lambda \geq 0 - n = 40$



(c)  $\lambda \leq 0 - n = 120$



(d)  $\lambda \geq 0 - n = 120$

Figura 3.41: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 40$  e  $n = 120$ .

No cenário epsilon, a probabilidade de mudança em cada instante para  $n = 30$  é baixa nos instantes da mudança, menos de 30%. Além disso, só o primeiro ponto de mudança é identificado com essa probabilidade e para  $\lambda = \pm 100$ . Se o tamanho da amostra aumenta a probabilidade ainda é baixa, próxima de 20%, mas nesse caso  $\lambda = \pm 2, \pm 0.7, 0$  identificam corretamente os instantes de mudança.

No cenário tendência a probabilidade de mudança em cada instante é próxima de zero independente do tamanho da amostra e do valor do parâmetro de assimetria, mas nesse tipo de cenário o modelo partição produto tende a ter um desempenho pouco satisfatório.

De modo geral, a probabilidade de mudança em cada instante apresenta um comportamento similar ao número de blocos e partição mais provável.

A seguir será apresentado o estudo de sensibilidade informal na escolha da distribuição amostral, ou seja, como se comporta o MPP quando a contaminação é feita na distribuição dos dados. No que segue, o modelo discutido na próxima Seção será chamado de Modelo 2.

### 3.2 Sensibilidade informal na especificação da distribuição amostral

O estudo de sensibilidade apresentado nesta seção, contempla o modelo apresentado na Seção 2.5, ou seja, quando a função de distribuição dos dados pertence à classe de distribuição normal assimétrica.

Os resultados apresentados a seguir são baseados em uma única amostra, onde os dados foram gerados da distribuição normal assimétrica  $Y_i \sim \mathcal{SN}(\mu_i, \sigma_y^2, \lambda)$ , onde o parâmetro de escala  $\sigma_y^2$  é conhecido e fixo,  $\lambda$  é fixado nos valores  $\lambda = (-100, -2, -0.7, 0, 0.7, 2, 100)$ . A distribuição *a priori* para o parâmetro  $\mu_i$  é uma distribuição normal  $\mathcal{N}(\mu_0, \sigma_\mu^2)$ . Os hiperparâmetros  $\mu_0$  e  $\sigma_\mu^2$  foram considerados fixos e iguais a média dos dados e 1000, respectivamente. Assume-se *a priori*, que  $p \sim \text{Beta}(1.1, 1.1)$ .

Para o MCMC, considerou-se uma cadeia de tamanho 60000 iterações e o período de *burn-in* foi de 50000 iterações com um lag de 10 observações, logo o tamanho da amostra *a posteriori* final foi de 1000 observações. Para avaliar a sensibilidade foram fixados 5 valores do parâmetro de assimetria  $\lambda = (-100, 2, -0.7, 0, 0.7, 2$  e  $100)$  e foram considerados 2 tamanhos amostrais diferentes ( $n = 30, 90$ ). Todos os resultados foram obtidos por meio do software R (R Development Core Team, 2012).

Nesse estudo de sensibilidade informal foram considerados os cenários descritos na Tabela 3.3 a seguir, no entanto, considere que a  $\mathcal{SN}(\mu, \sigma_y^2, \lambda)$ , está representando a parametrização centrada da distribuição normal assimétrica em que  $\mu$  representa o valor da verdadeira média gerada e  $\sigma_y^2$  o

verdadeiro valor da variância e não o valor do parâmetro de locação e escala.

Tabela 3.3: Cenários avaliados

Cenário	Dist. do 1 <sup>o</sup> bloco.	Dist. do 2 <sup>o</sup> bloco	Dist. do 3 <sup>o</sup> bloco	Dist. do 4 <sup>o</sup> bloco
Sem mudança	$\mathcal{SN}(0, 1, \lambda)$	$\mathcal{SN}(0, 1, \lambda)$	$\mathcal{SN}(0, 1, \lambda)$	–
Degrau simétrico positivo	$\mathcal{SN}(5, 1, \lambda)$	$\mathcal{SN}(8, 1, \lambda)$	$\mathcal{SN}(5, 1, \lambda)$	–
Degrau simétrico negativo	$\mathcal{SN}(-5, 1, \lambda)$	$\mathcal{SN}(-8, 1, \lambda)$	$\mathcal{SN}(-5, 1, \lambda)$	–
Degrau não simétrico	$\mathcal{SN}(-3, 1, \lambda)$	$\mathcal{SN}(7, 1, \lambda)$	$\mathcal{SN}(0, 1, \lambda)$	–
Escada crescente	$\mathcal{SN}(7, 1, \lambda)$	$\mathcal{SN}(10, 1, \lambda)$	$\mathcal{SN}(13, 1, \lambda)$	–
Escada decrescente	$\mathcal{SN}(13, 1, \lambda)$	$\mathcal{SN}(10, 1, \lambda)$	$\mathcal{SN}(7, 1, \lambda)$	–
Epsilon	$\mathcal{SN}(1.75, 1, \lambda)$	$\mathcal{SN}(2.5, 1, \lambda)$	$\mathcal{SN}(3.25, 1, \lambda)$	–
Tendência	$\mathcal{SN}(0, 0.01, \lambda)$	$\mathcal{SN}(0.3 + 0.008 \times i, 0.1, \lambda)$	$\mathcal{SN}(0.6 + 0.018 \times i, 0.2, \lambda)$	$\mathcal{SN}(1.1, 0.2, \lambda)$

A seguir serão apresentados os resultados das estimativas *a posteriori* considerando os cenários apresentados na Tabela 3.3. Para facilitar a leitura do texto, como os resultados são apresentados por meio de vários gráficos o texto foi dividido em três seções. As estimativas produto e variância de  $\mu_i$ , serão apresentadas na Seção 3.2.1. As estimativas *a posteriori* para a partição e número de blocos são discutidas na Seção 3.2.2 e por fim, na Seção 3.2.3 as estimativas *a posteriori* para a probabilidade de mudança em cada instante são exibidas.

### 3.2.1 A influência de $\lambda$ e $n$ nas estimativas produto e na variância *a posteriori* das médias

Para o estudo de sensibilidade nesta seção, faz-se necessário gerar a série para um valor de  $\lambda$  e ajustar o modelo assumindo todos os valores de  $\lambda$  sugeridos nessa análise de sensibilidade. Por exemplo, é gerada uma série assumindo que o verdadeiro valor de  $\lambda = -100$ , logo, ajusta-se o modelo assumindo que  $\lambda = (-100, 2, -0.7, 0, 0.7, 2$  e  $100)$ . Assim sendo, cada valor de  $\lambda$  gerado implica em sete ajustes realizados. O objetivo é verificar quais valores de  $\lambda$  produzem estimativas

satisfatórias para os cenários simulados.

Desta forma, nesta seção são apresentadas as estimativas produto e a variância *a posteriori* de  $\mu_i$ . Ao lado esquerdo das figuras, os gráficos (a), (c), (e) e (g) apresentam as estimativas produto e os gráficos (b), (d), (f) e (h) mostram a variância *a posteriori* de  $\mu_i$ .

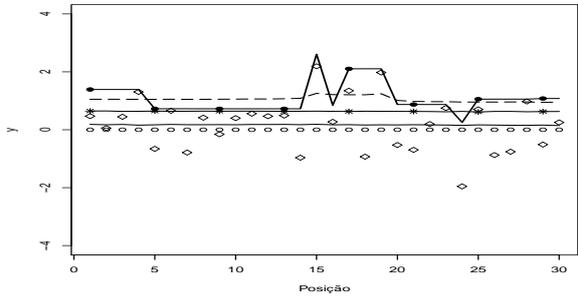
As Figuras 3.42 e 3.43 ilustram os resultados para as estimativas produto e variância *a posteriori* de  $\mu_i$  para o cenário sem mudança, assumindo tamanho de amostra 30 e 90.

Nota-se, no cenário sem mudança, que se os dados são gerados assumindo um valor de assimetria baixo  $\lambda = \pm 0.7$  ou  $\lambda = 0$ , ao ajustar um valor alto para o parâmetro de assimetria ( $\lambda = \pm 100$ ), as estimativas produto tendem a oscilar mais e acompanhar um número maior de pontos ao invés de suavizar as estimativas em torno da verdadeira média dos blocos.

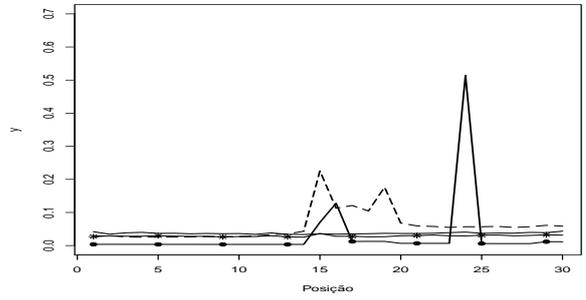
Nota-se, também, nas Figuras 3.43 e 3.45 que quando os dados são gerados assumindo  $\lambda \leq 0$  e ajusta-se o modelo com  $\lambda \geq 0$  as estimativas produto são subestimadas e observa-se que mesmo com tamanho de amostra 90 a subestimação persiste.

As estimativas das variâncias apresentam maiores variações quando o modelo gerado assume  $\lambda = \pm 0.7$  e  $\lambda = 0$  e ajusta-se o modelo com outros valores de  $\lambda$ .

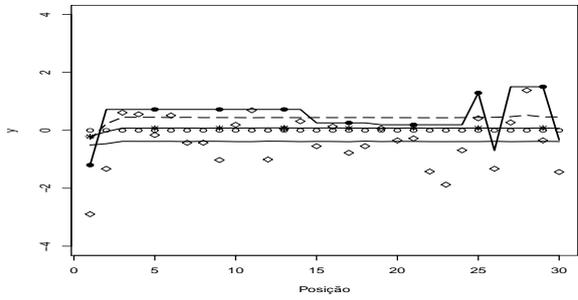
Ao longo dos cenários que serão apresentados, observa-se um padrão no comportamento da estimativas, ou seja, se os dados são simulados com um parâmetro de assimetria negativo e o modelo é ajustado com assimetria positiva, então as estimativas tendem a ser subestimadas. Se o contrário ocorre, os dados são simulados com um parâmetro de assimetria positivo e o modelo é ajustado com assimetria negativa, então as estimativas serão superestimadas. Portanto, para tornar o texto menos cansativo para o leitor, optou-se por mostrar somente os gráficos em que os dados são simulados com um parâmetro de assimetria negativo e o modelo é ajustado com assimetria positiva, os demais casos foram suprimidos desse capítulo e encontram-se no apêndice D.



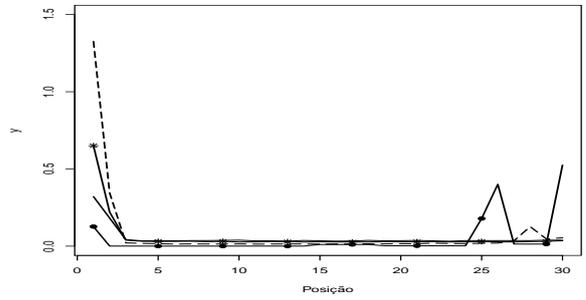
(a) Mod. gerado  $\lambda = 0$



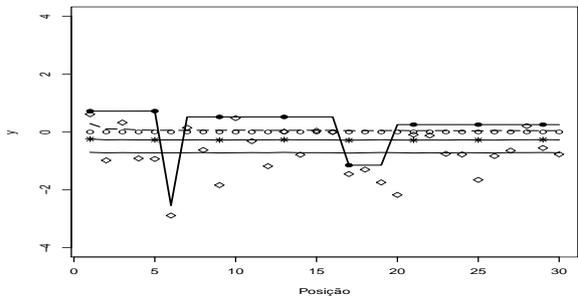
(b) Mod. gerado  $\lambda = 0$



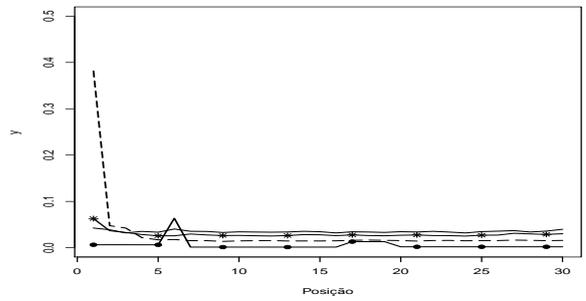
(c) Mod. gerado  $\lambda = -0.7$



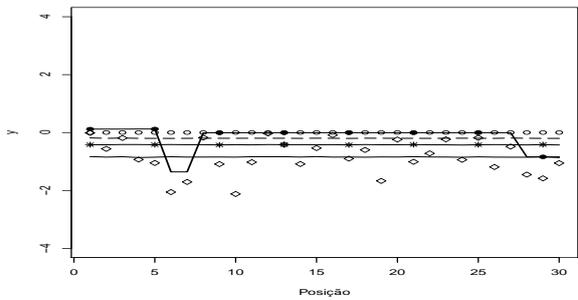
(d) Mod. gerado  $\lambda = -0.7$



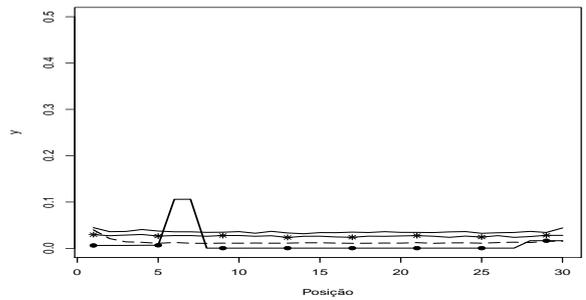
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

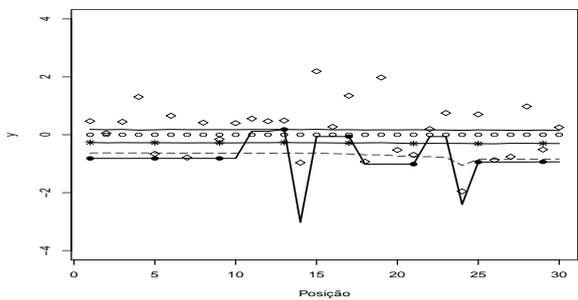


(g) Mod. gerado  $\lambda = -100$

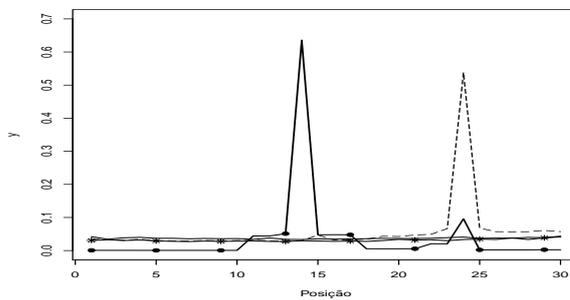


(h) Mod. gerado  $\lambda = -100$

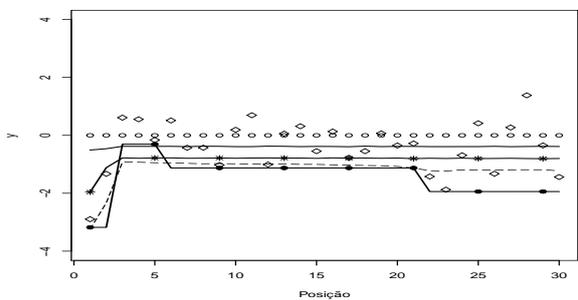
Figura 3.42: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 30$



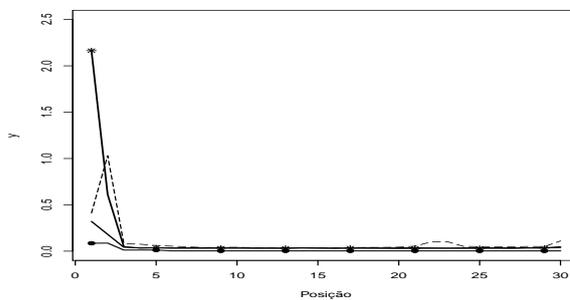
(a) Mod. gerado  $\lambda = 0$



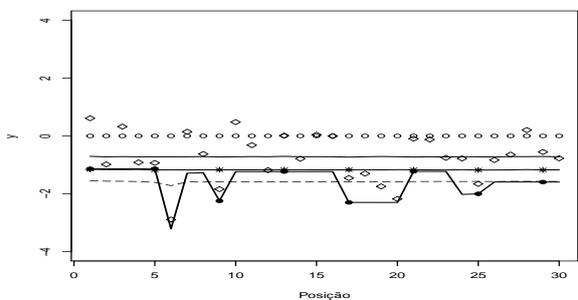
(b) Mod. gerado  $\lambda = 0$



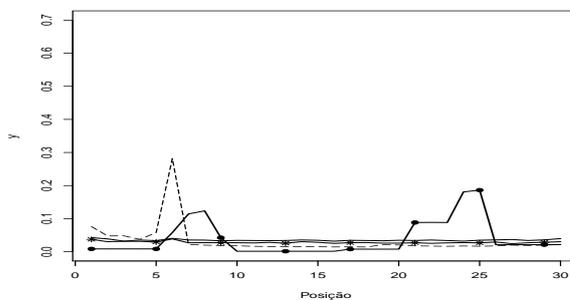
(c) Mod. gerado  $\lambda = -0.7$



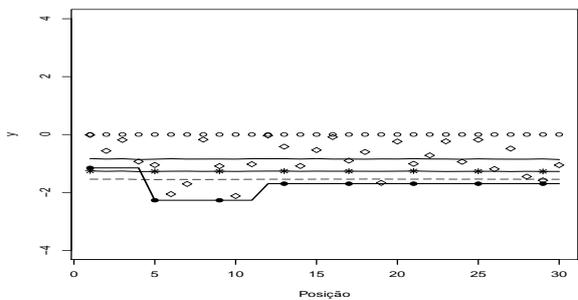
(d) Mod. gerado  $\lambda = -0.7$



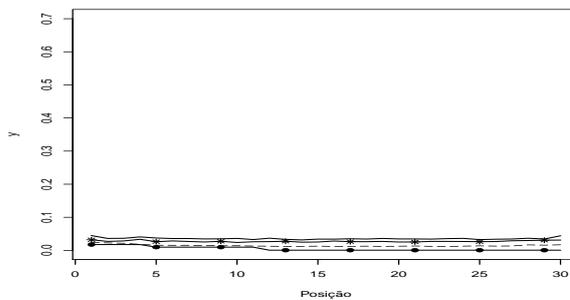
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

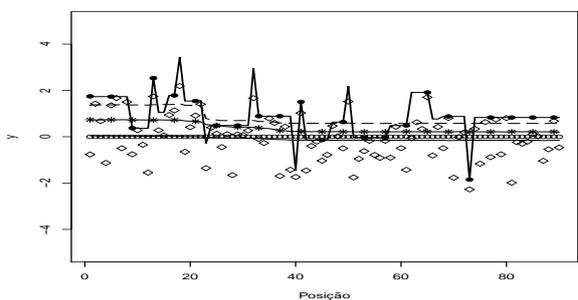


(g) Mod. gerado  $\lambda = -100$

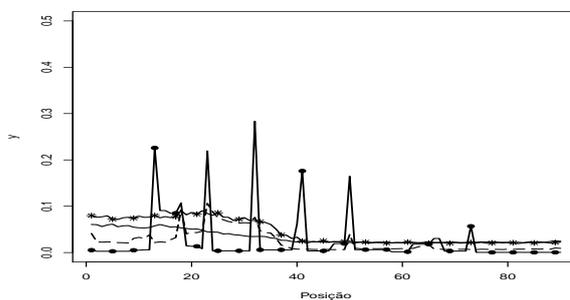


(h) Mod. gerado  $\lambda = -100$

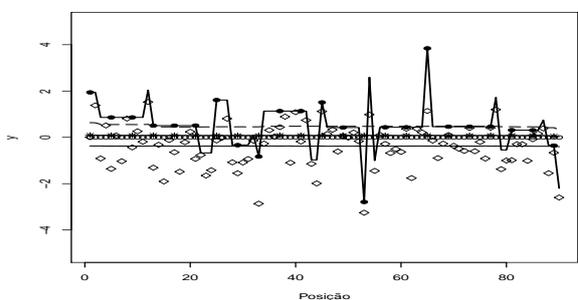
Figura 3.43: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 30$ .



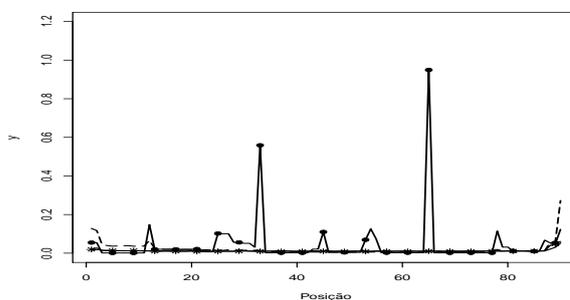
(a) Mod. gerado  $\lambda = 0$



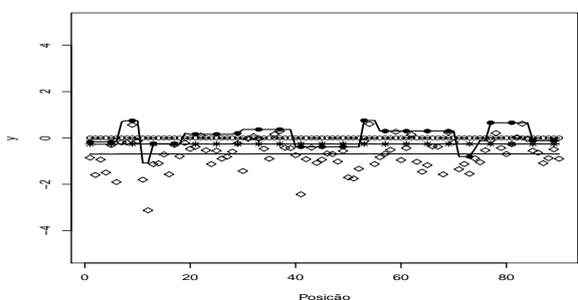
(b) Mod. gerado  $\lambda = 0$



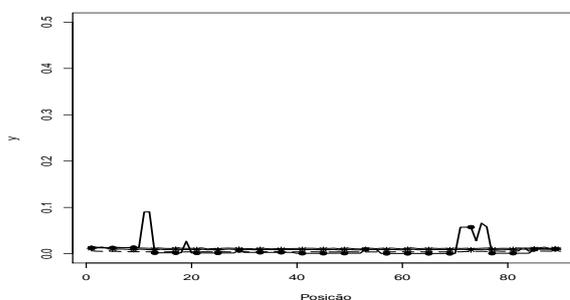
(c) Mod. gerado  $\lambda = -0.7$



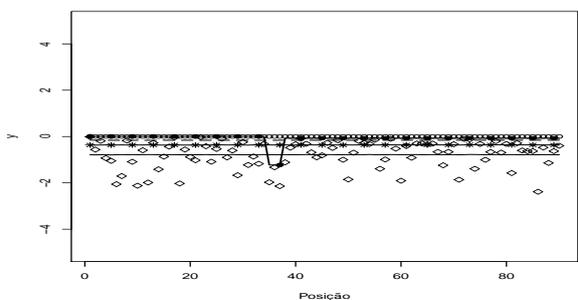
(d) Mod. gerado  $\lambda = -0.7$



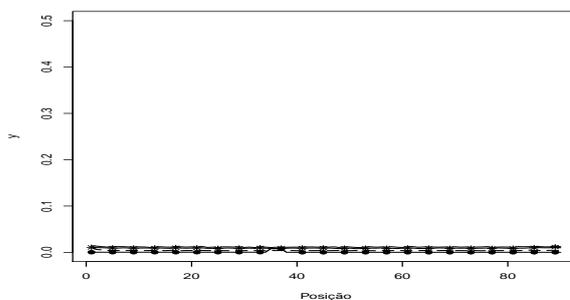
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

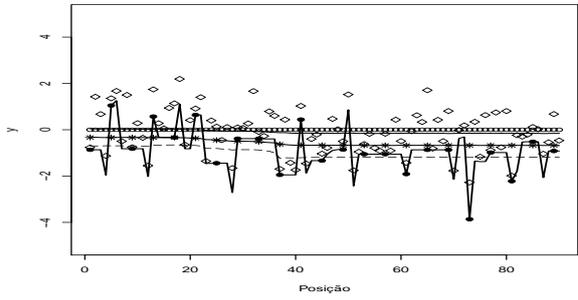


(g) Mod. gerado  $\lambda = -100$

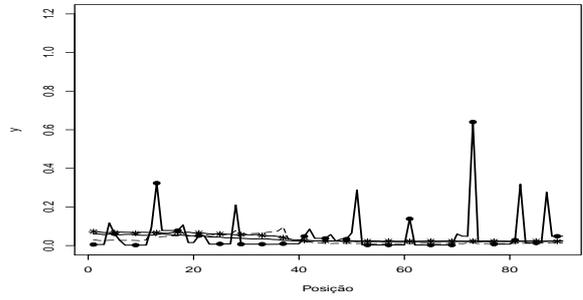


(h) Mod. gerado  $\lambda = -100$

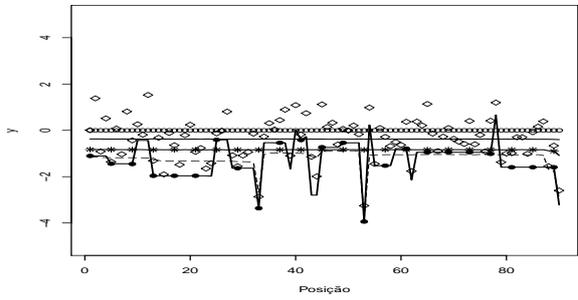
Figura 3.44: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 90$ .



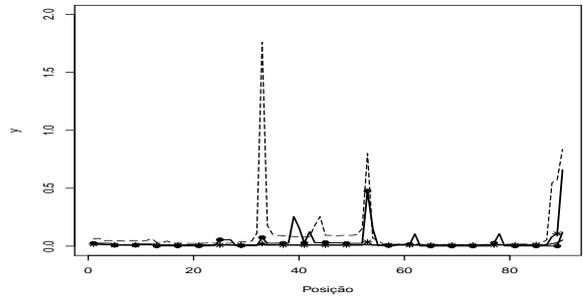
(a) Mod. gerado  $\lambda = 0$



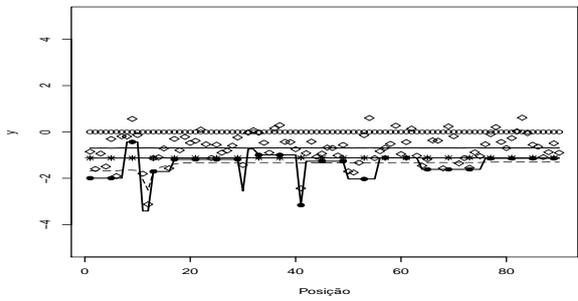
(b) Mod. gerado  $\lambda = 0$



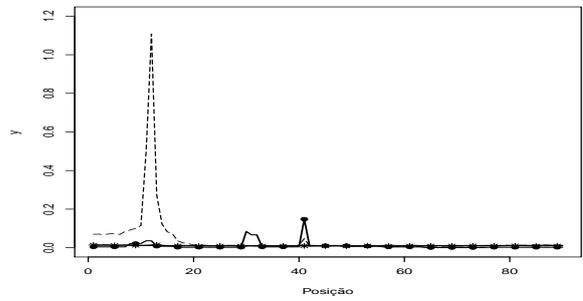
(c) Mod. gerado  $\lambda = -0.7$



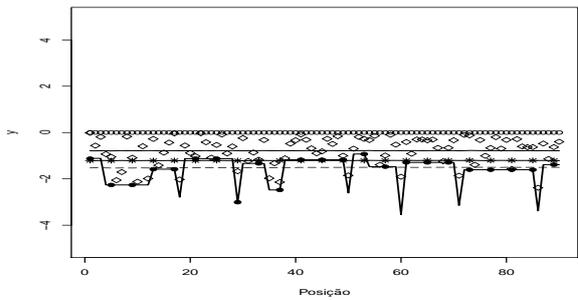
(d) Mod. gerado  $\lambda = -0.7$



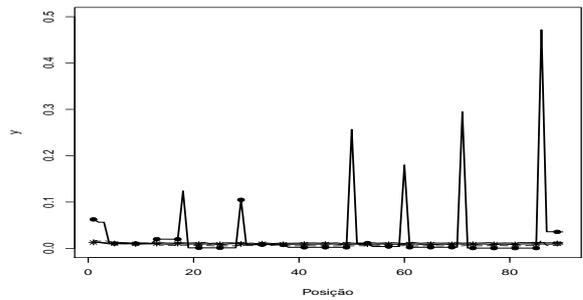
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$



(g) Mod. gerado  $\lambda = -100$



(h) Mod. gerado  $\lambda = -100$

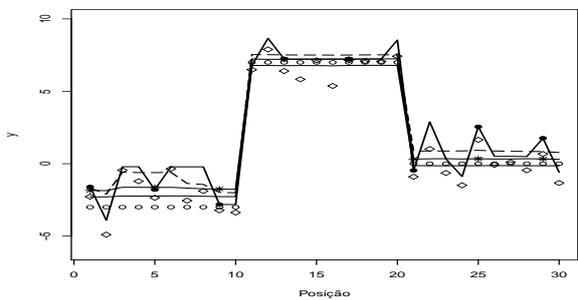
Figura 3.45: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 90$ .

O cenário degrau não simétrico é constituído por blocos negativos e positivos de tal modo que as diferenças entre as médias dos blocos não são constantes, ou seja, tal diferença é de 10 unidades do primeiro para o segundo bloco e de 7 unidades do segundo para o terceiro bloco. Os resultados desse cenário encontram-se nas Figuras 3.46,3.47,3.48,3.49.

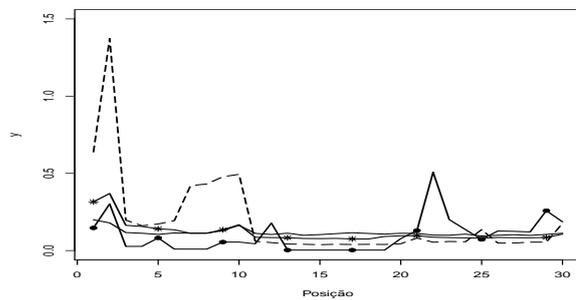
Nota-se, nesse cenário, que apesar do conjunto de dados dispor de observações positivas e negativas, quando o modelo é gerado com um parâmetro de assimetria nulo ou próximo de zero, e ajusta-se um modelo com  $\lambda = \pm 100$  as estimativas produto oscilam mais, independente dos tamanhos de amostra avaliados.

Observa-se nas Figuras 3.47 e 3.49 que quando o modelo é gerado com um parâmetro de assimetria positivo e ajusta-se o modelo assumindo que a assimetria é negativa as estimativas tendem a ser superestimadas. Se a situação for o inverso, ou seja, gera-se os dados de  $\lambda < 0$  e ajusta-se  $\lambda > 0$  as estimativas são subestimadas.

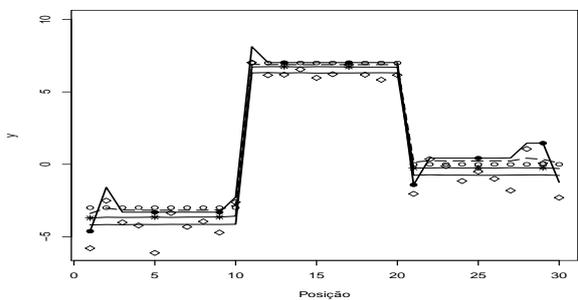
Em relação as variâncias *a posteriori* de  $\mu_i$  nota-se que as maiores variações nas estimativas ocorrem próximas aos pontos de mudança e quando os dados são gerados de  $\lambda = 0$  e  $\pm 0.7$  e ajustados com  $\lambda = \pm 2$ .



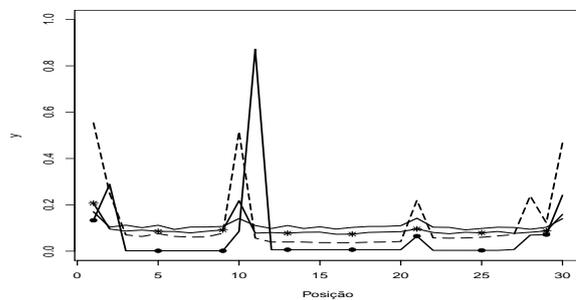
(a) Mod. gerado  $\lambda = 0$



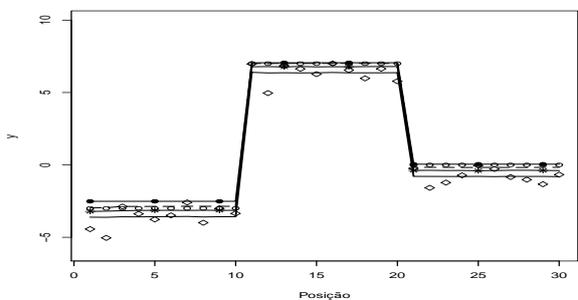
(b) Mod. gerado  $\lambda = 0$



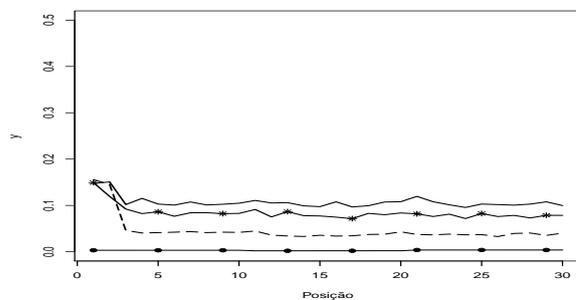
(c) Mod. gerado  $\lambda = -0.7$



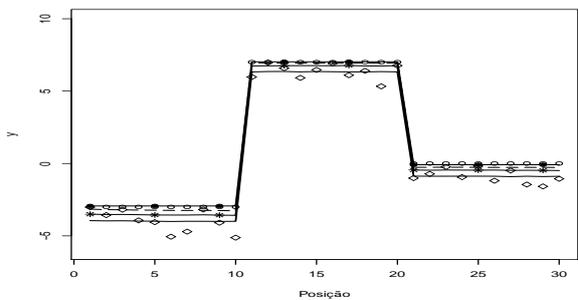
(d) Mod. gerado  $\lambda = -0.7$



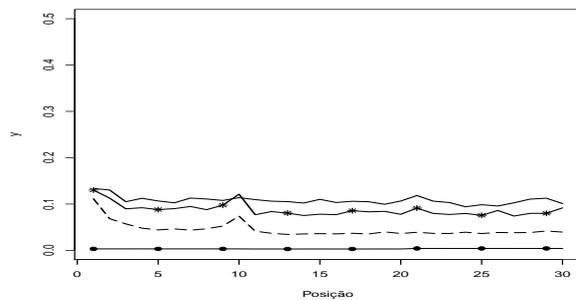
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

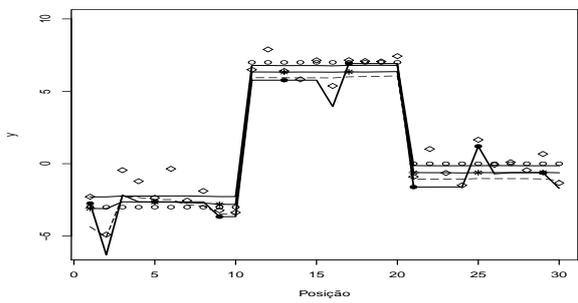


(g) Mod. gerado  $\lambda = -100$

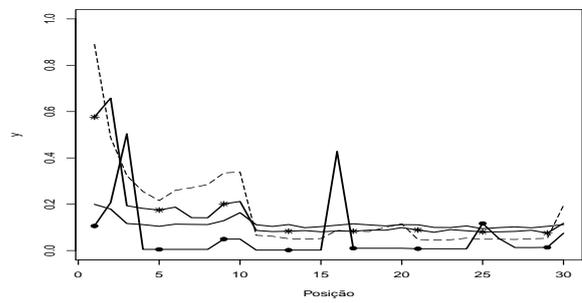


(h) Mod. gerado  $\lambda = -100$

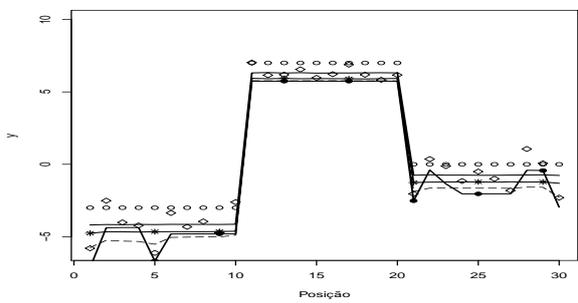
Figura 3.46: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 30$



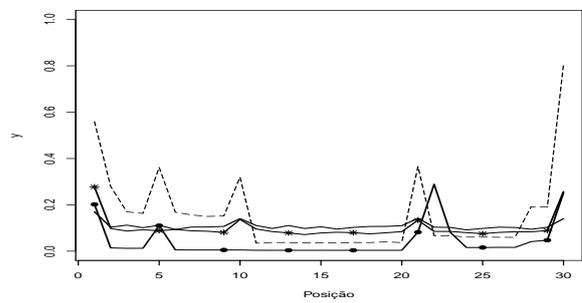
(a) Mod. gerado  $\lambda = 0$



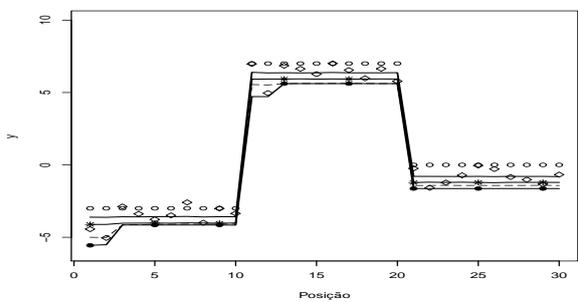
(b) Mod. gerado  $\lambda = 0$



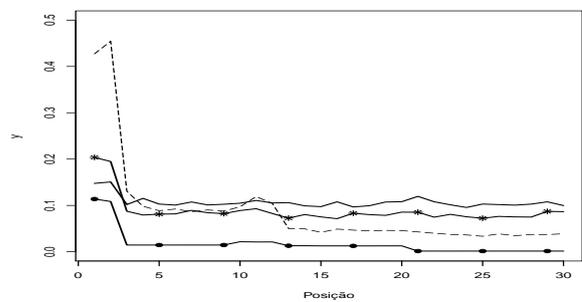
(c) Mod. gerado  $\lambda = -0.7$



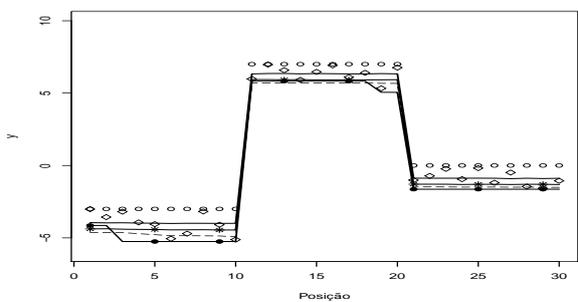
(d) Mod. gerado  $\lambda = -0.7$



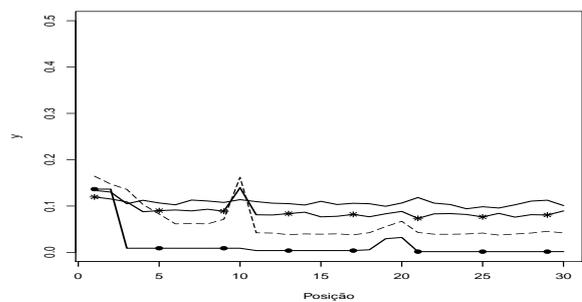
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

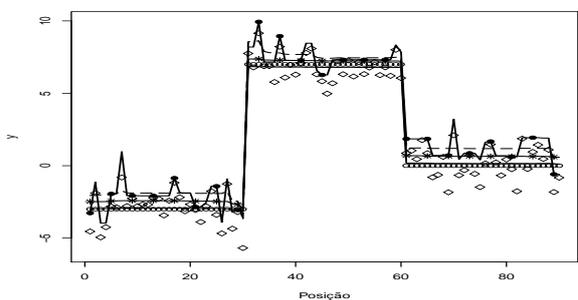


(g) Mod. gerado  $\lambda = -100$

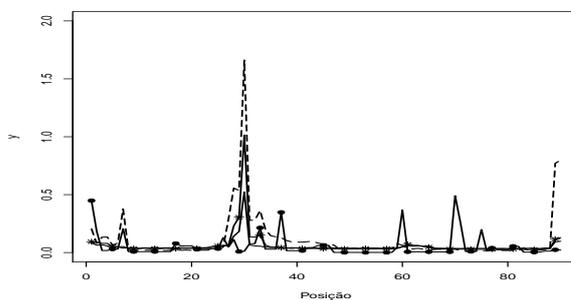


(h) Mod. gerado  $\lambda = -100$

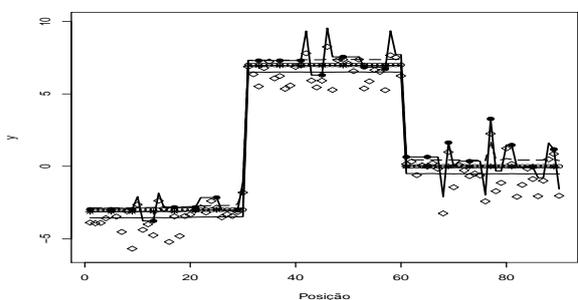
Figura 3.47: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 30$ .



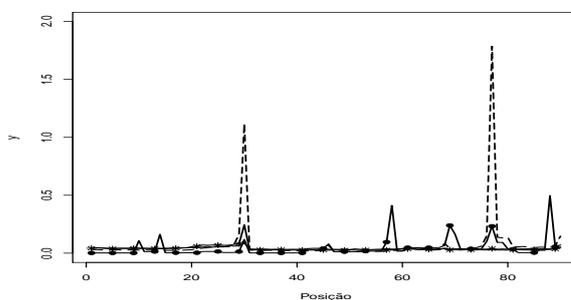
(a) Mod. gerado  $\lambda = 0$



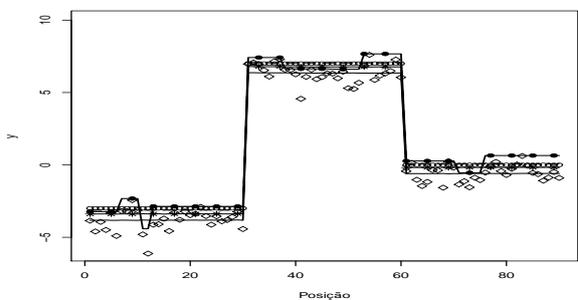
(b) Mod. gerado  $\lambda = 0$



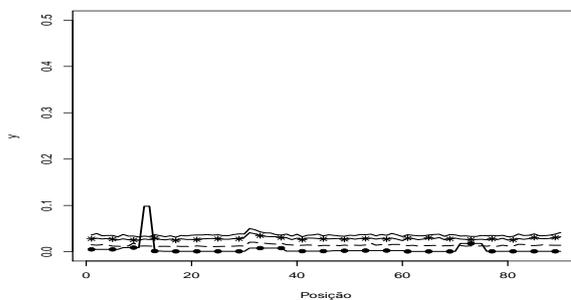
(c) Mod. gerado  $\lambda = -0.7$



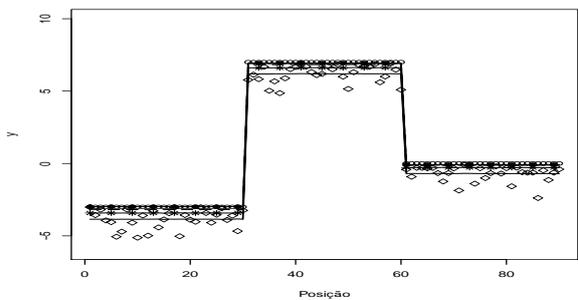
(d) Mod. gerado  $\lambda = -0.7$



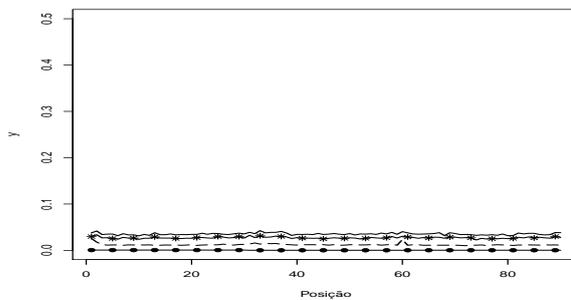
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

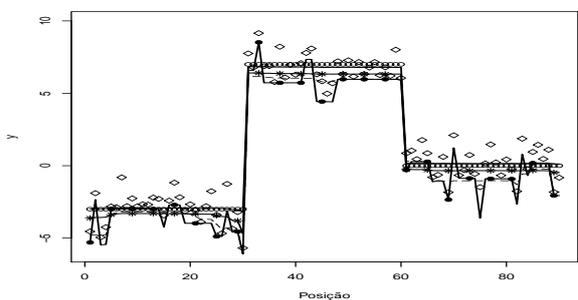


(g) Mod. gerado  $\lambda = -100$

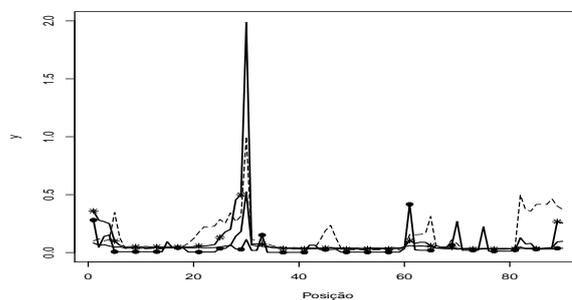


(h) Mod. gerado  $\lambda = -100$

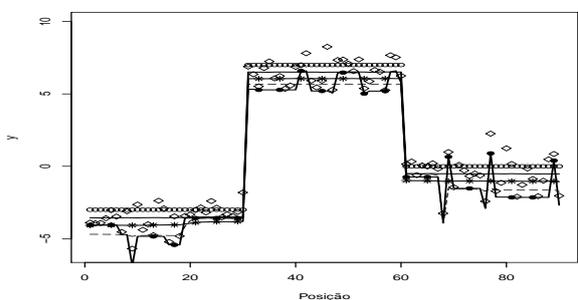
Figura 3.48: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 90$ .



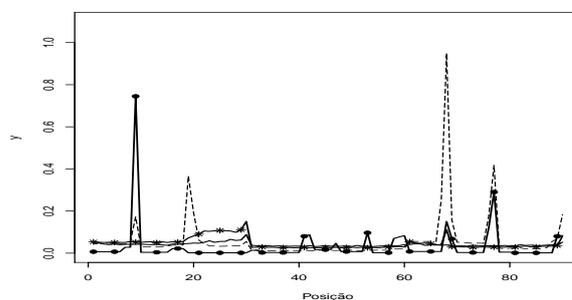
(a) Mod. gerado  $\lambda = 0$



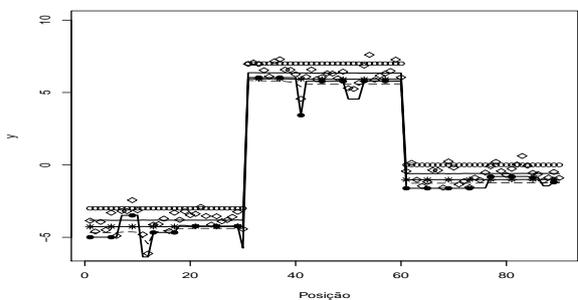
(b) Mod. gerado  $\lambda = 0$



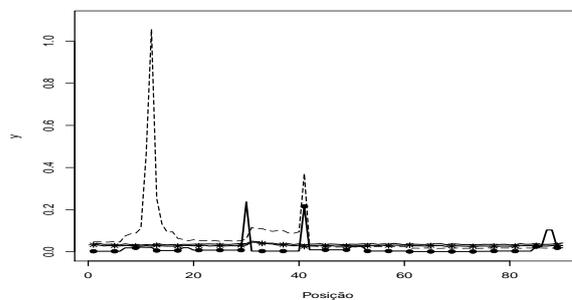
(c) Mod. gerado  $\lambda = -0.7$



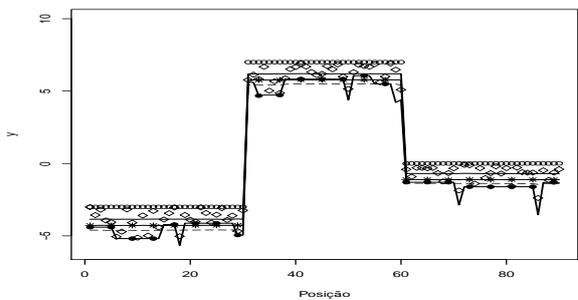
(d) Mod. gerado  $\lambda = -0.7$



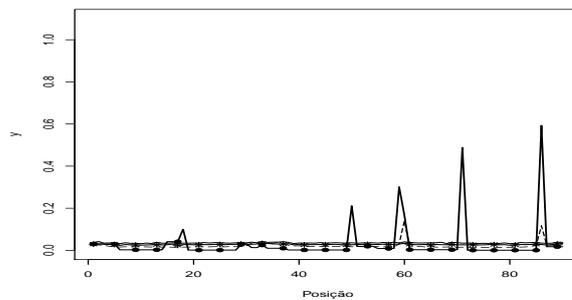
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$



(g) Mod. gerado  $\lambda = -100$



(h) Mod. gerado  $\lambda = -100$

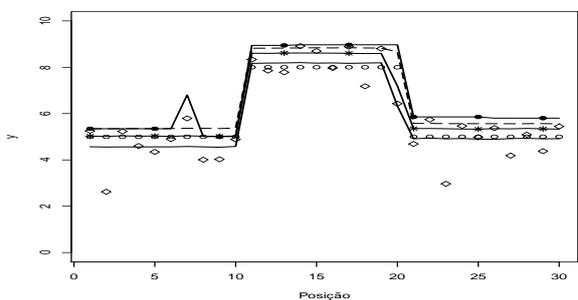
Figura 3.49: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 90$ .

Os resultados para o cenário degrau positivo, cujas as diferenças entre os blocos são de 3 unidades, são apresentados nas Figuras 3.50, 3.51, 3.52, 3.53.

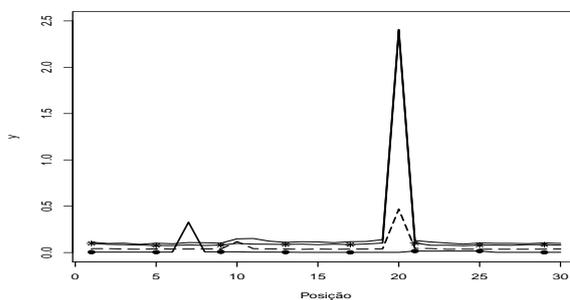
Nota-se de forma geral, um comportamento similar ao comportamento do cenário degrau não simétrico, ou seja, estimativas ruins para dados gerados com assimetria baixa ou nula e ajustados com  $\lambda = \pm 100$ . A subestimação e a superestimação também ocorrem quando o modelo é ajustado com parâmetro de assimetria contrário aos dados gerados.

Em relação as variância *a posteriori* de  $\mu_i$  nota-se que as maiores variações são em torno dos pontos de mudança ou observações que se distanciam da média do bloco, como é o caso do ponto 73 na Figura 3.52 (c) e (d).

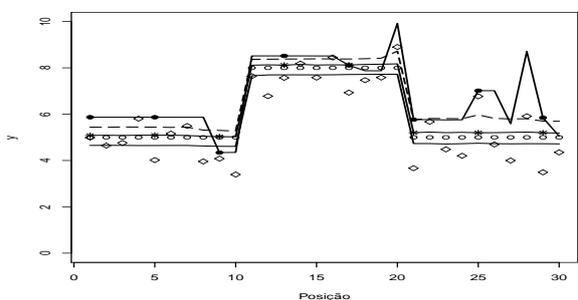
Os resultados para o cenário degrau negativo são similares ao cenário degrau positivo e encontram-se no Anexo D.



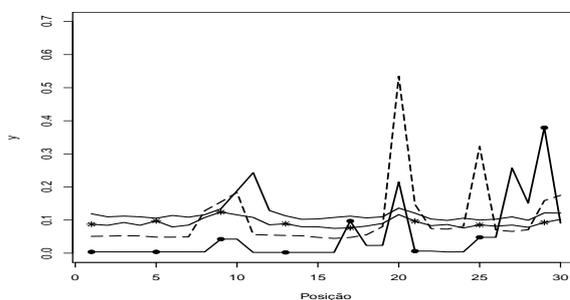
(a) Mod. gerado  $\lambda = 0$



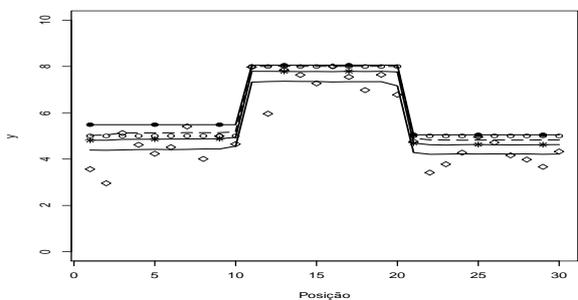
(b) Mod. gerado  $\lambda = 0$



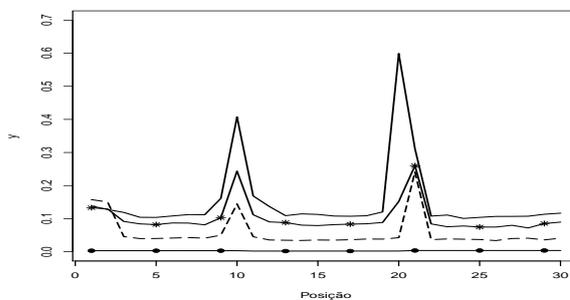
(c) Mod. gerado  $\lambda = -0.7$



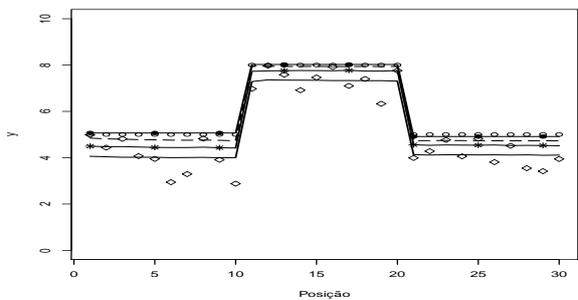
(d) Mod. gerado  $\lambda = -0.7$



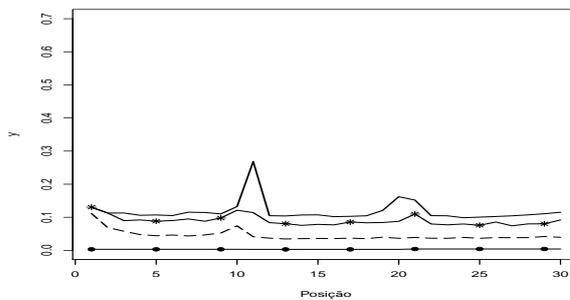
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

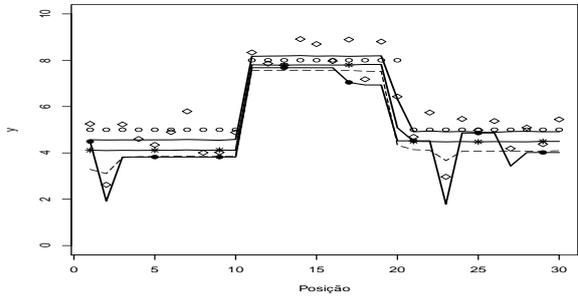


(g) Mod. gerado  $\lambda = -100$

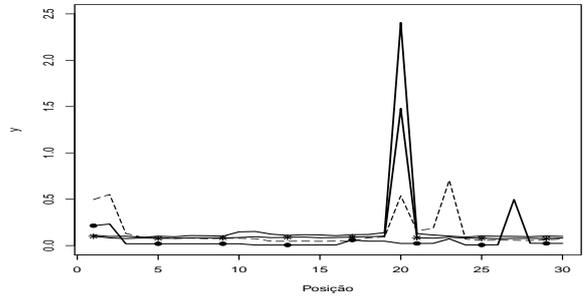


(h) Mod. gerado  $\lambda = -100$

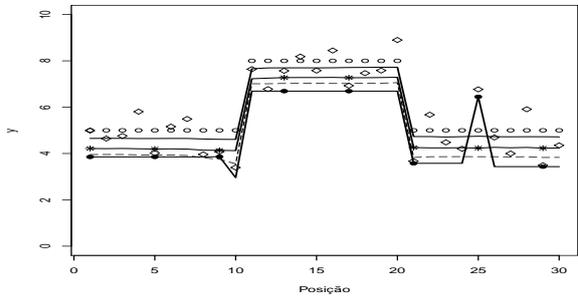
Figura 3.50: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 30$



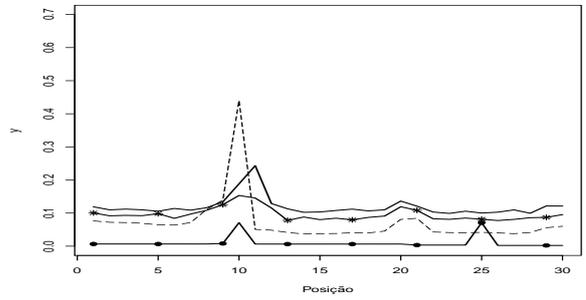
(a) Mod. gerado  $\lambda = 0$



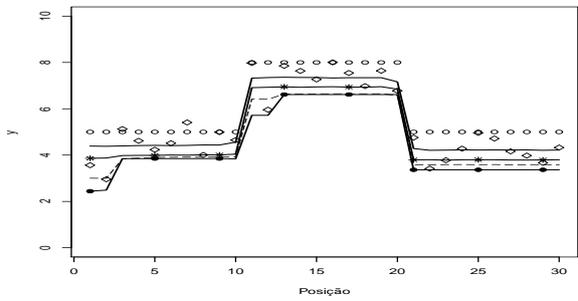
(b) Mod. gerado  $\lambda = 0$



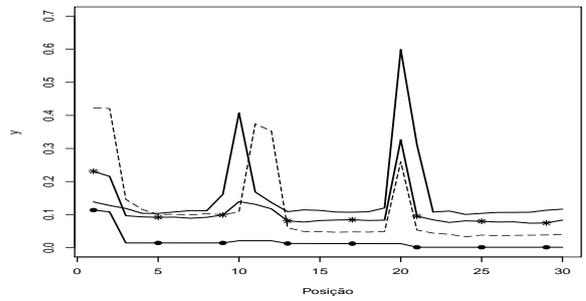
(c) Mod. gerado  $\lambda = -0.7$



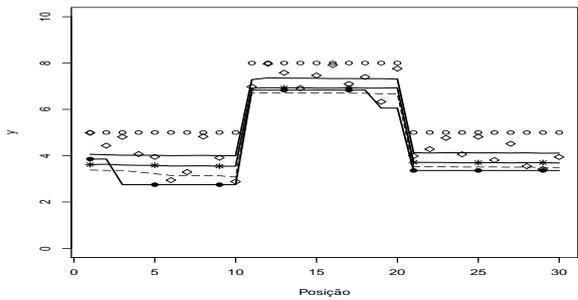
(d) Mod. gerado  $\lambda = -0.7$



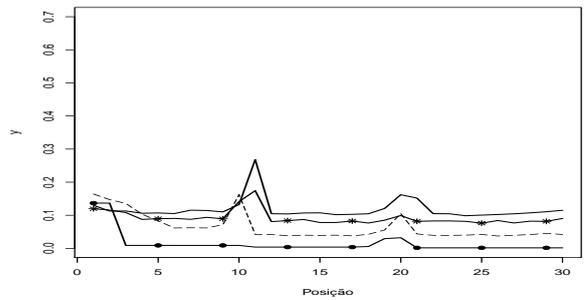
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

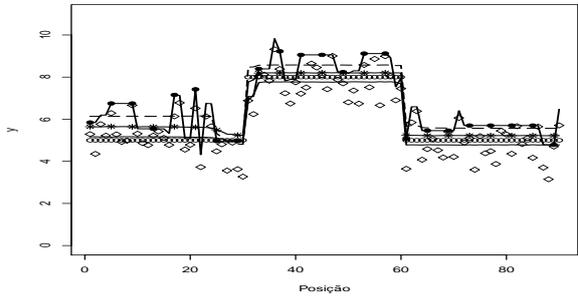


(g) Mod. gerado  $\lambda = -100$

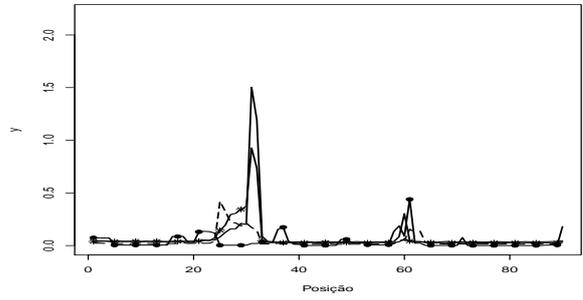


(h) Mod. gerado  $\lambda = -100$

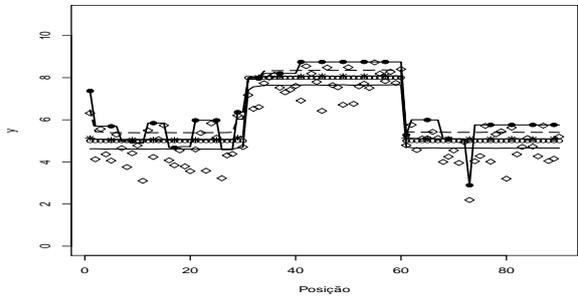
Figura 3.51: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 30$ .



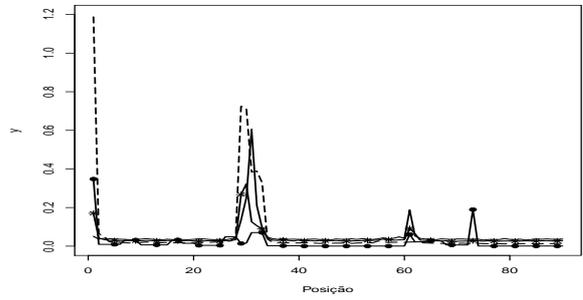
(a) Mod. gerado  $\lambda = 0$



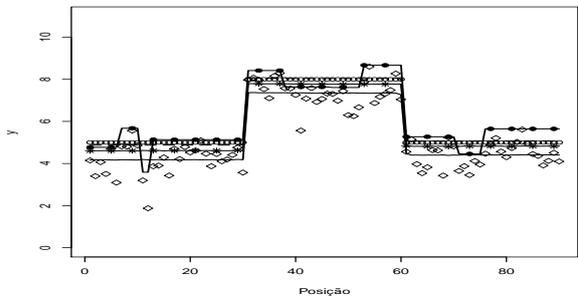
(b) Mod. gerado  $\lambda = 0$



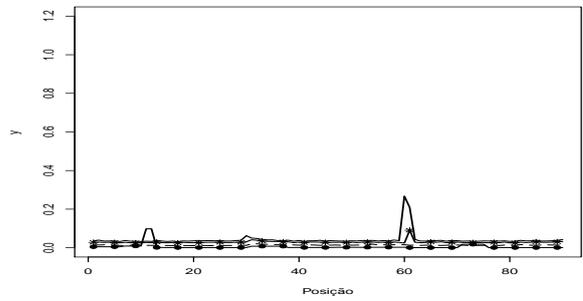
(c) Mod. gerado  $\lambda = -0.7$



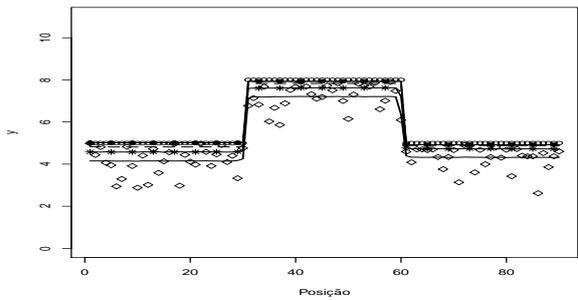
(d) Mod. gerado  $\lambda = -0.7$



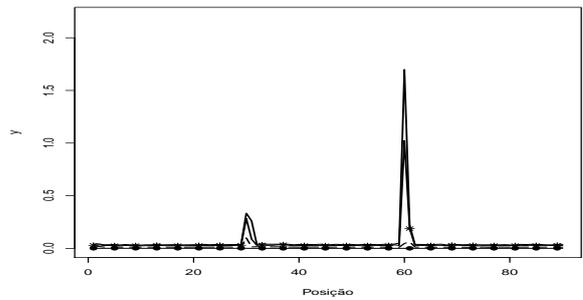
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

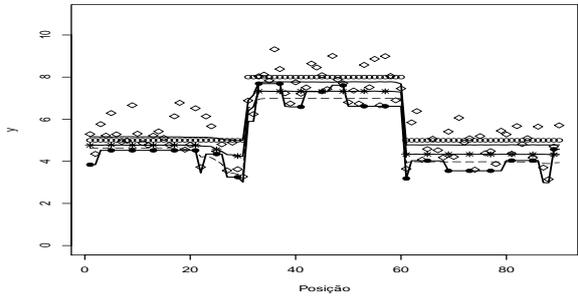


(g) Mod. gerado  $\lambda = -100$

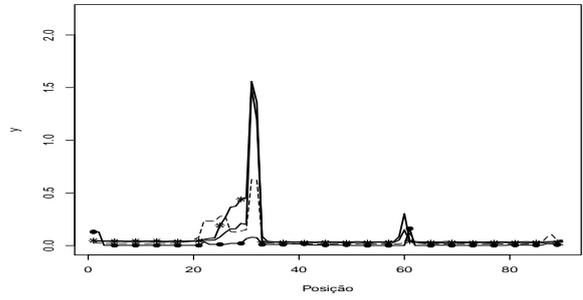


(h) Mod. gerado  $\lambda = -100$

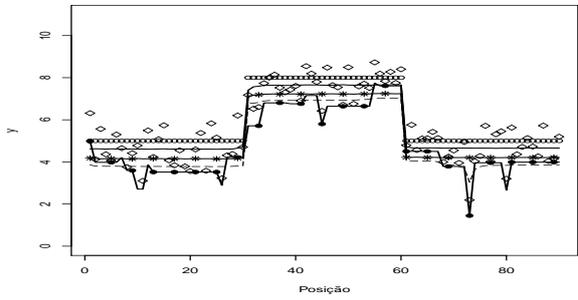
Figura 3.52: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 90$ .



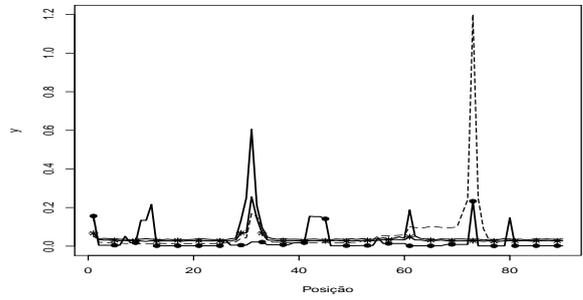
(a) Mod. gerado  $\lambda = 0$



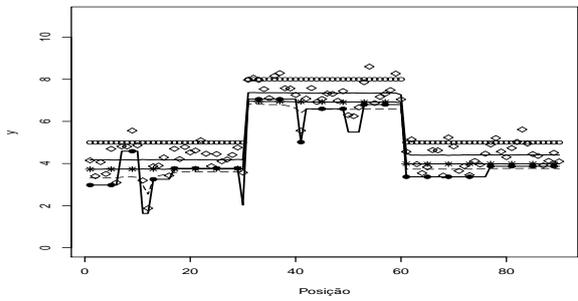
(b) Mod. gerado  $\lambda = 0$



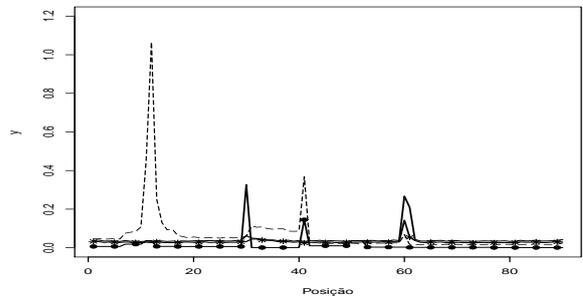
(c) Mod. gerado  $\lambda = -0.7$



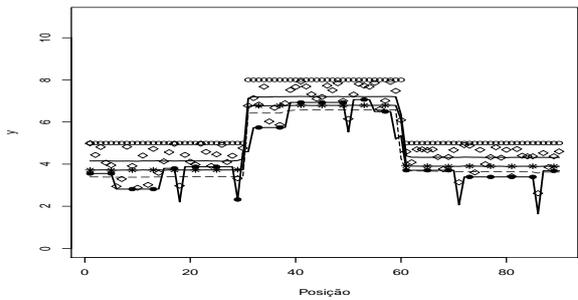
(d) Mod. gerado  $\lambda = -0.7$



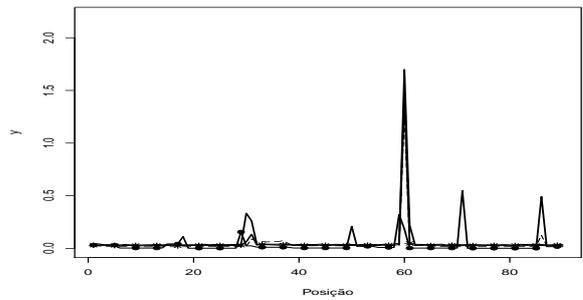
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$



(g) Mod. gerado  $\lambda = -100$



(h) Mod. gerado  $\lambda = -100$

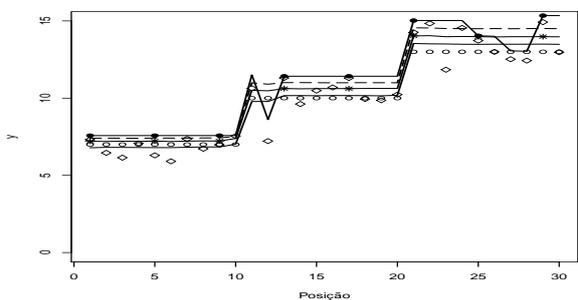
Figura 3.53: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 90$ .

No cenário escada crescente foram considerados dois pontos de mudança e as médias dos blocos foram consideradas sempre positivas e crescentes. A diferença entre os blocos vizinhos é de 3 unidades. As Figuras de 3.54,3.55, 3.56 e 3.57 ilustram os resultados obtidos. As figuras com os resultados do cenário escada decrescente apresentam resultados similares e estão no Anexo D.

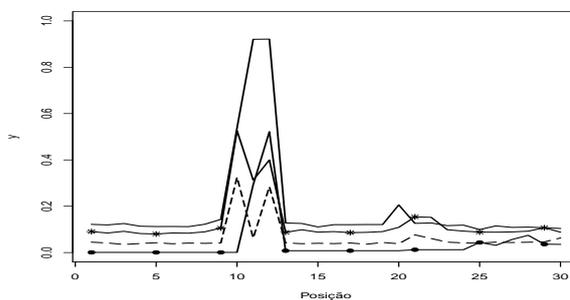
As estimativas produto tendem a oscilar mais em relação a média verdadeira quando os dados são gerados com  $\lambda = 0$  e  $\pm 0.7$  e ajustados com  $\lambda \pm 100$ , ou seja, o mesmo comportamento dos cenários anteriores em relação ao ajuste de  $\lambda = \pm 100$ . Em contrapartida, quando os dados são gerados de  $\lambda = 0$  e ajustados com  $\lambda \neq 0$  observa-se um aumento da incerteza sobre  $\mu_i$ , em relação aos cenários avaliados anteriormente.

Por fim, observa-se que as estimativas para  $\lambda = \pm 100$  quando os dados gerados assumem  $\lambda = 0$  apresentam estimativas que oscilam mais, no entanto, as estimativas são melhores que no caso quando assume-se assimetria para as médias e considera  $\lambda = \pm 100$  para a distribuição *a priori* de  $\mu_i$ , como visto na Seção 3.1.1.

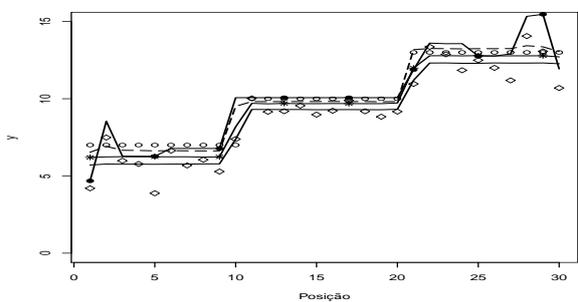
Observa-se na Figura 3.53 que para o tamanho de amostra 90, dados gerados de  $\lambda \leq 0$  e ajustados com  $\lambda = 100$  ou dados gerados de  $\lambda \geq 0$  e ajustados com  $\lambda = -100$  levam a estimativas produto que tendem a passar por diversos pontos ao invés de suavizarem a média dos dados.



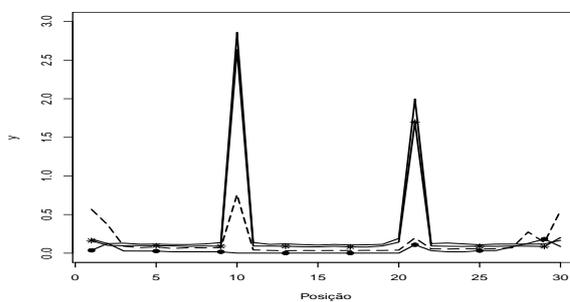
(a) Mod. gerado  $\lambda = 0$



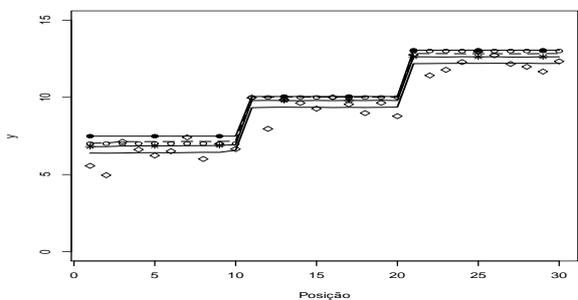
(b) Mod. gerado  $\lambda = 0$



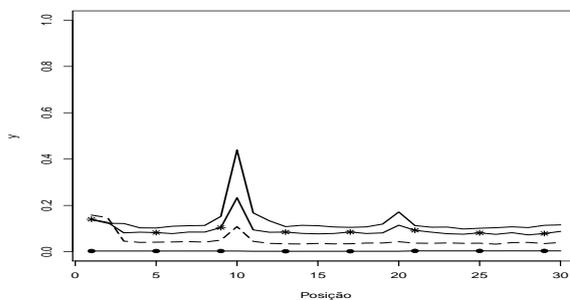
(c) Mod. gerado  $\lambda = -0.7$



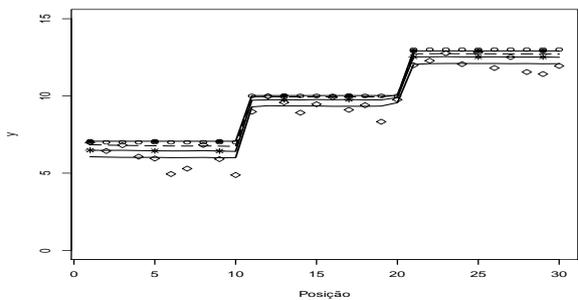
(d) Mod. gerado  $\lambda = -0.7$



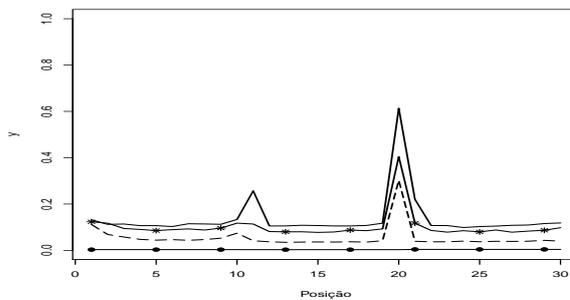
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

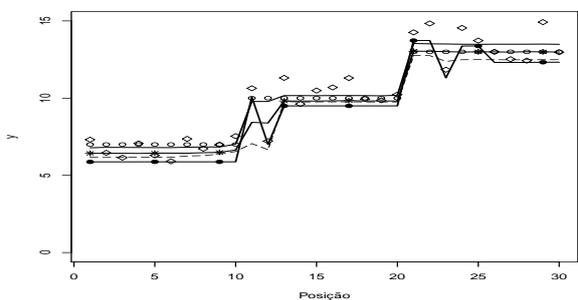


(g) Mod. gerado  $\lambda = -100$

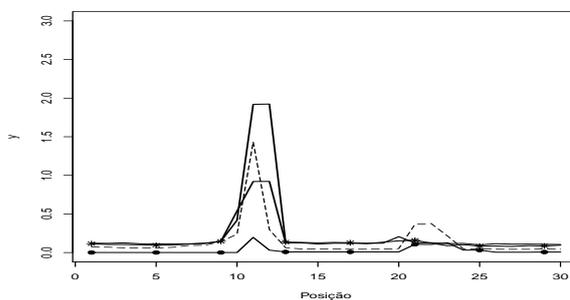


(h) Mod. gerado  $\lambda = -100$

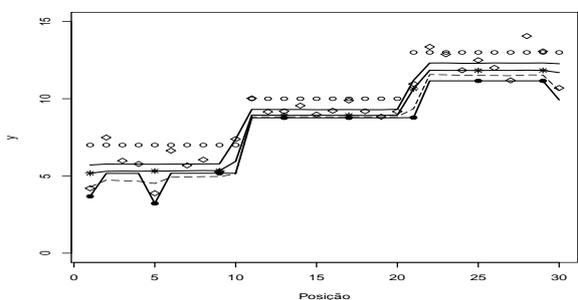
Figura 3.54: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 30$



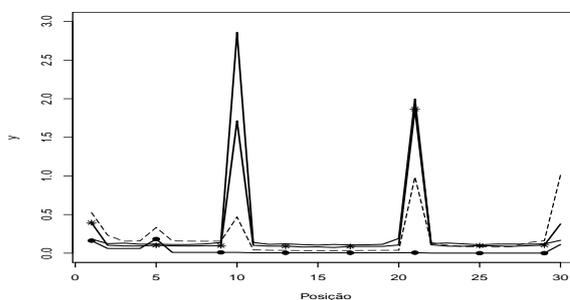
(a) Mod. gerado  $\lambda = 0$



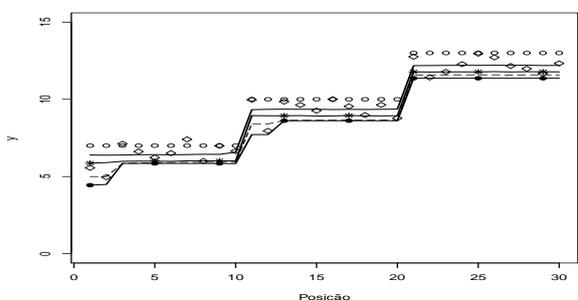
(b) Mod. gerado  $\lambda = 0$



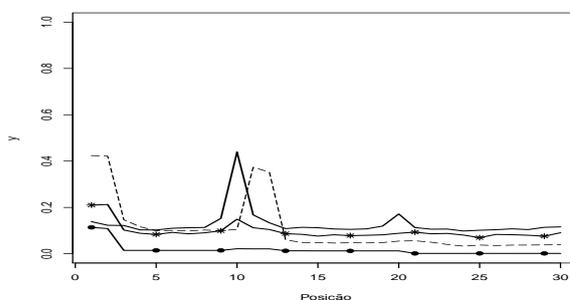
(c) Mod. gerado  $\lambda = -0.7$



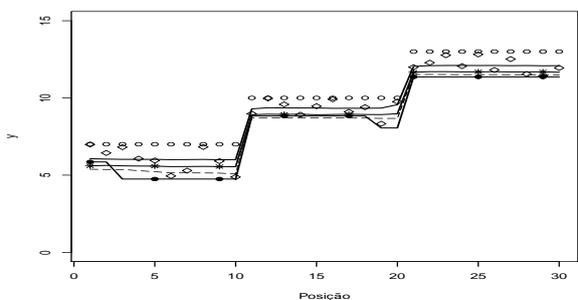
(d) Mod. gerado  $\lambda = -0.7$



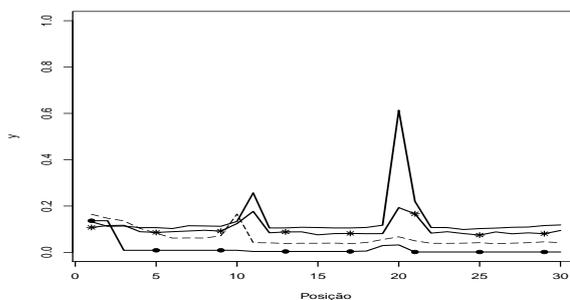
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

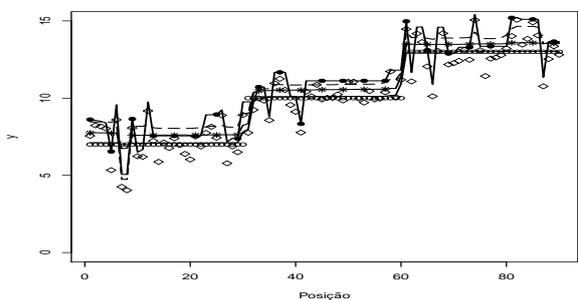


(g) Mod. gerado  $\lambda = -100$

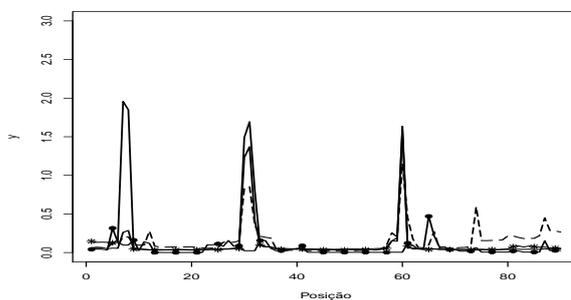


(h) Mod. gerado  $\lambda = -100$

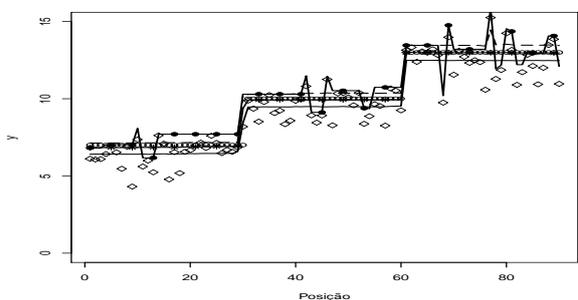
Figura 3.55: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 30$ .



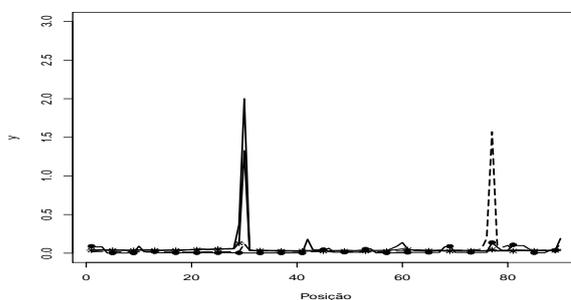
(a) Mod. gerado  $\lambda = 0$



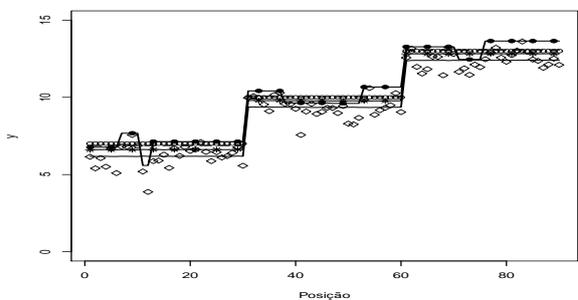
(b) Mod. gerado  $\lambda = 0$



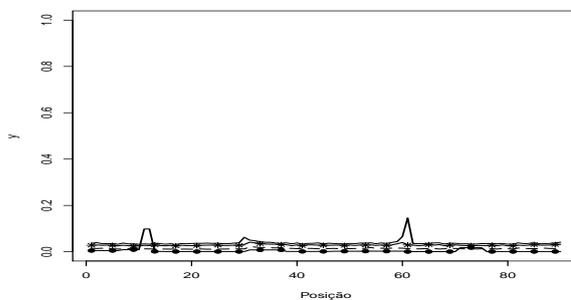
(c) Mod. gerado  $\lambda = -0.7$



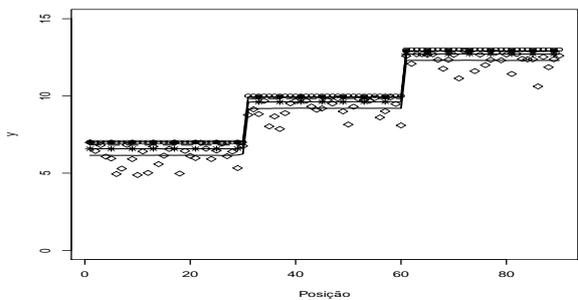
(d) Mod. gerado  $\lambda = -0.7$



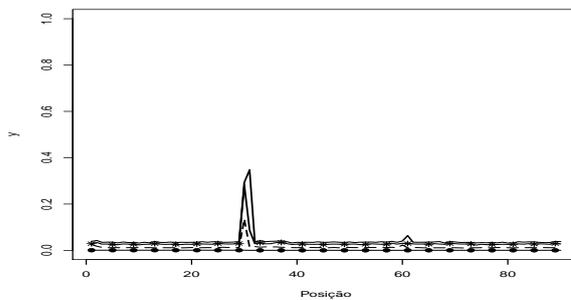
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

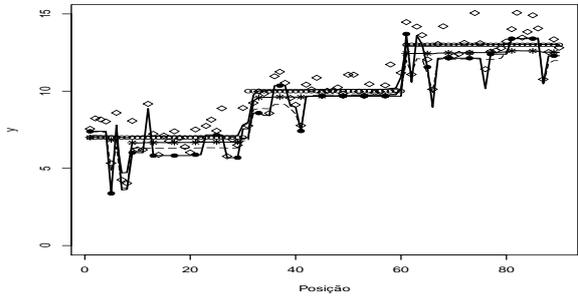


(g) Mod. gerado  $\lambda = -100$

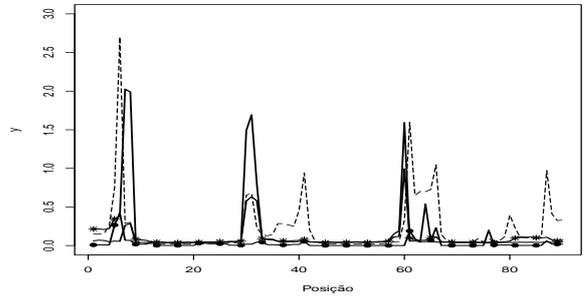


(h) Mod. gerado  $\lambda = -100$

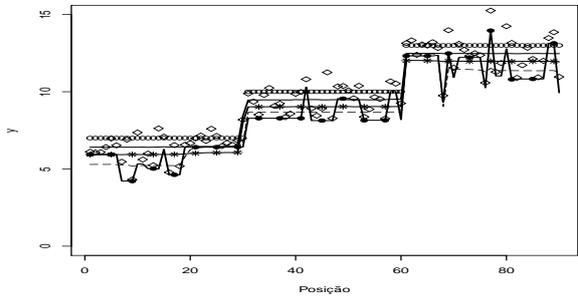
Figura 3.56: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 90$ .



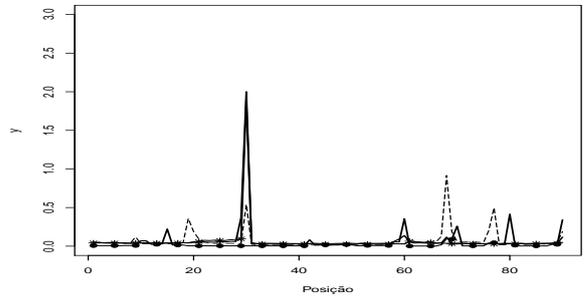
(a) Mod. gerado  $\lambda = 0$



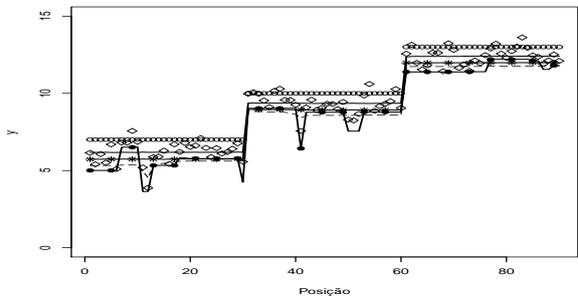
(b) Mod. gerado  $\lambda = 0$



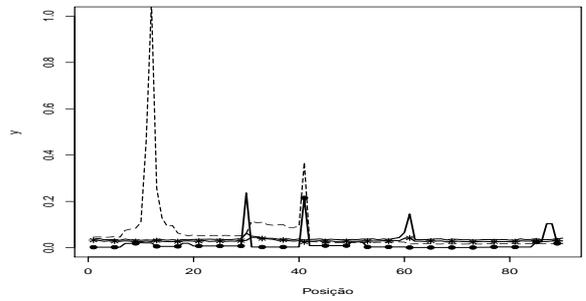
(c) Mod. gerado  $\lambda = -0.7$



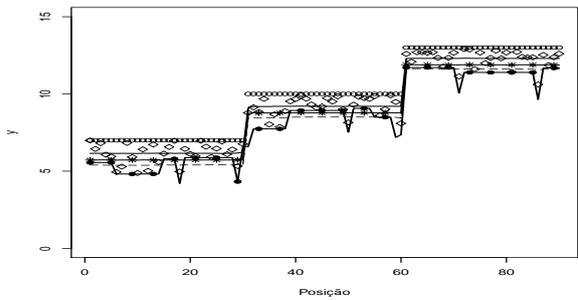
(d) Mod. gerado  $\lambda = -0.7$



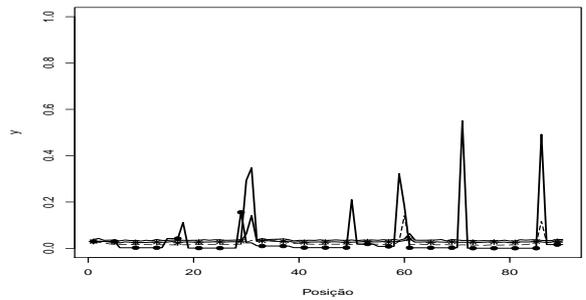
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$



(g) Mod. gerado  $\lambda = -100$



(h) Mod. gerado  $\lambda = -100$

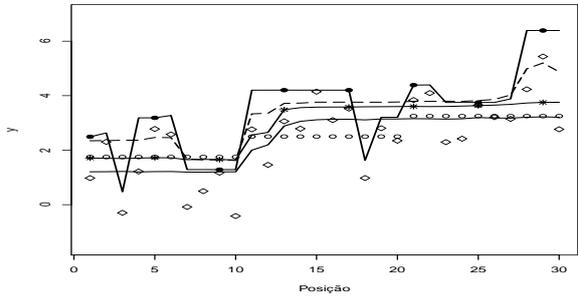
Figura 3.57: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 90$ .

O cenário epsilon ilustrado nas Figuras 3.58, 3.59, 3.60, 3.61 também possui o formato escada crescente, mas nesse caso a diferença entre as médias dos blocos é de 0.75. Nesse cenário,  $\lambda = 0$  não foi capaz de estimar corretamente as médias, mas nos casos em que os dados foram estimados com valores de  $\lambda$  em desacordo com a verossimilhança,  $\lambda = 0$  apresentou estimativas melhores que os demais valores de  $\lambda$  em discordância com os dados. No entanto, se os dados são gerados e estimados com o valor correto de assimetria, mesmo com uma diferença pequena entre os blocos, as estimativas são mais próximas as verdadeiras médias.

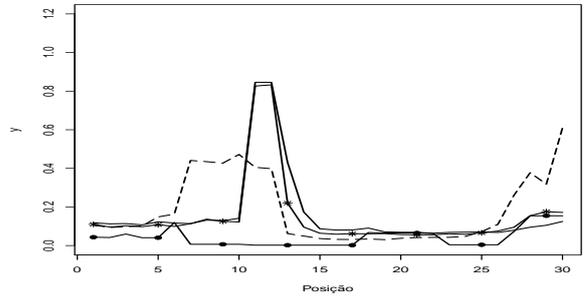
As Figuras 3.62, 3.63, 3.65, 3.64 mostram o comportamento das estimativas produto para o cenário tendência, no qual o segundo e o terceiro bloco possuem uma tendência determinística esse cenário apresenta quatro blocos, ou seja um a mais em relação aos outros cenários avaliados.

Observa-se no cenário tendência para  $n = 40$  que nenhum valor de  $\lambda$  foi capaz de estimar corretamente as médias, mesmo se  $\lambda$  foi gerado e ajustado com o valor correto.

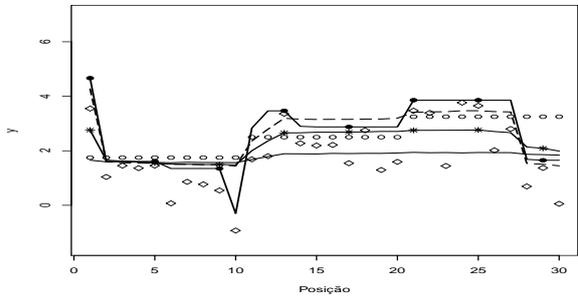
Quando o número de observações aumenta,  $n = 120$ , apesar de  $\lambda = 0$  suavizar as estimativas, modular as observações assumindo  $\lambda = 0$  produz melhores estimativas para as médias, mesmo quando os dados são gerados e ajustados com o parâmetro de assimetria correto, mas geralmente, esse tipo de cenário não é bem estimado pelo modelo partição produto.



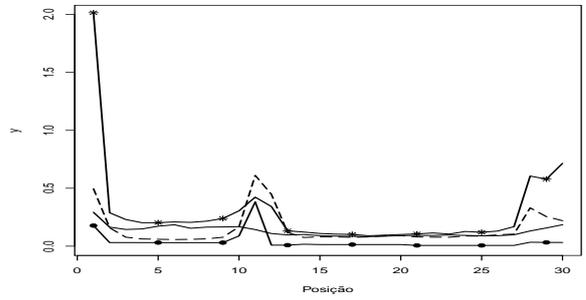
(a) Mod. gerado  $\lambda = 0$



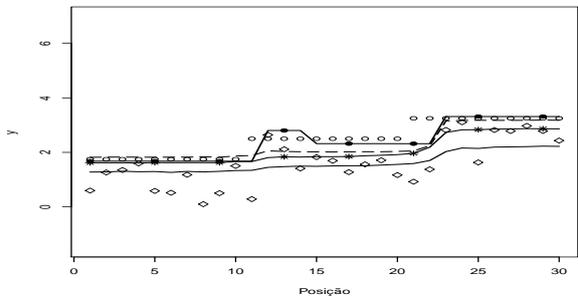
(b) Mod. gerado  $\lambda = 0$



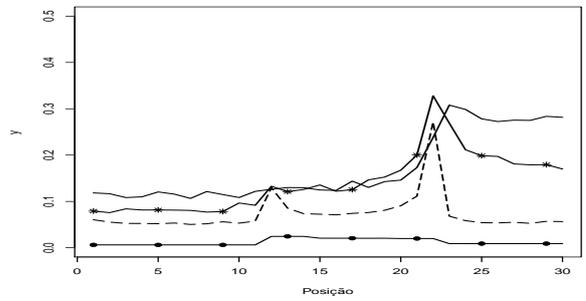
(c) Mod. gerado  $\lambda = -0.7$



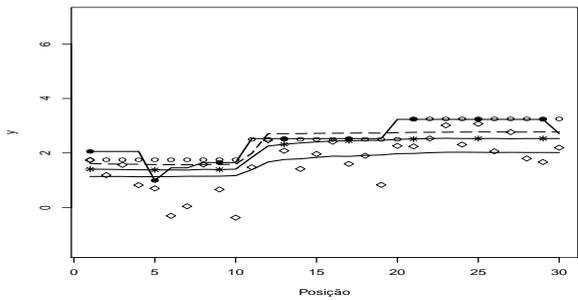
(d) Mod. gerado  $\lambda = -0.7$



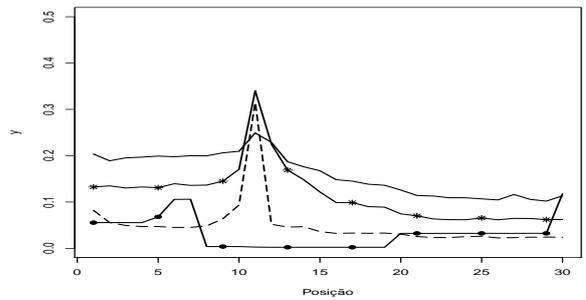
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

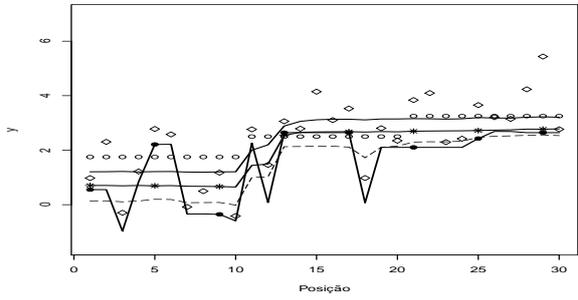


(g) Mod. gerado  $\lambda = -100$

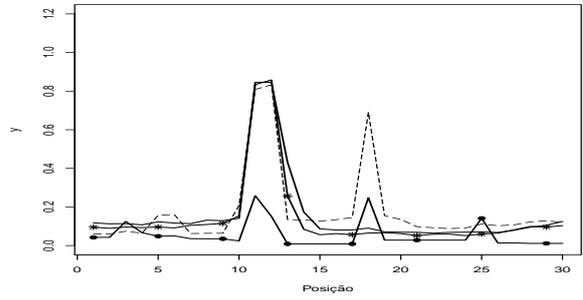


(h) Mod. gerado  $\lambda = -100$

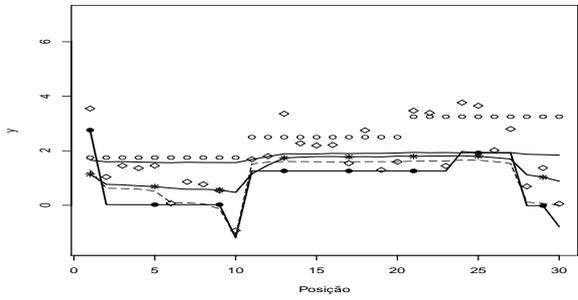
Figura 3.58: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 30$



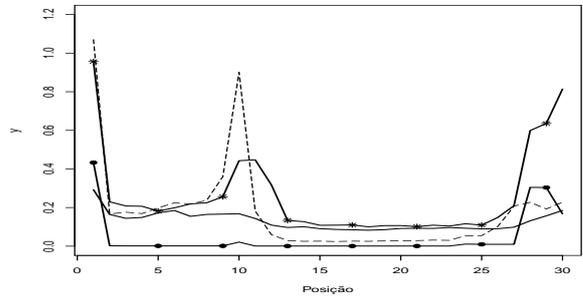
(a) Mod. gerado  $\lambda = 0$



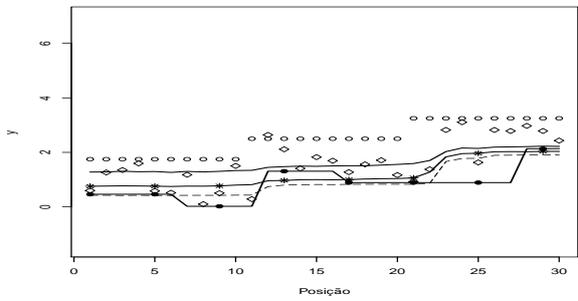
(b) Mod. gerado  $\lambda = 0$



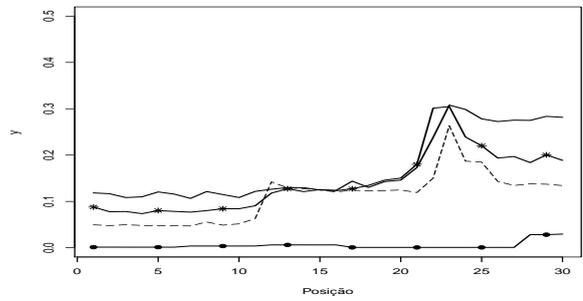
(c) Mod. gerado  $\lambda = -0.7$



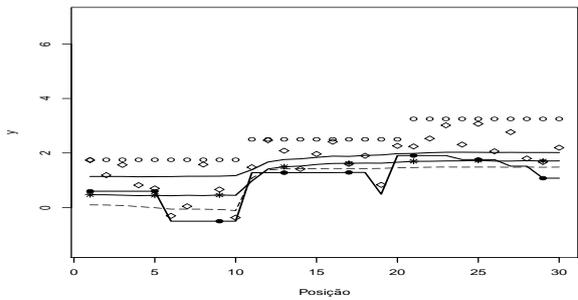
(d) Mod. gerado  $\lambda = -0.7$



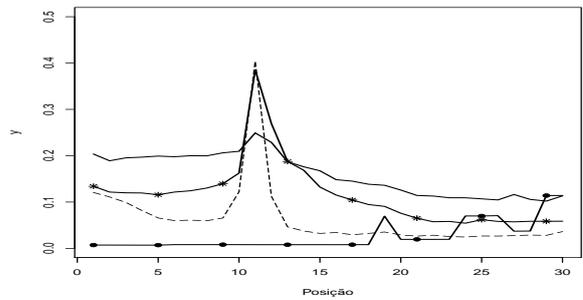
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

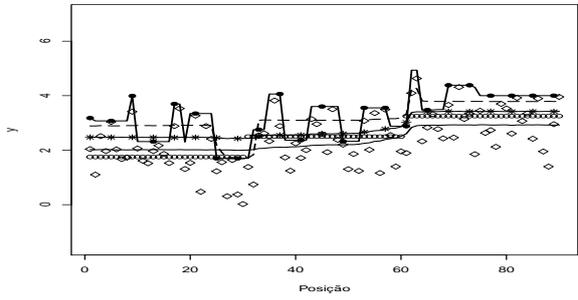


(g) Mod. gerado  $\lambda = -100$

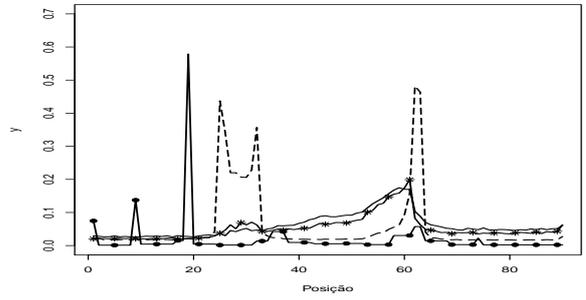


(h) Mod. gerado  $\lambda = -100$

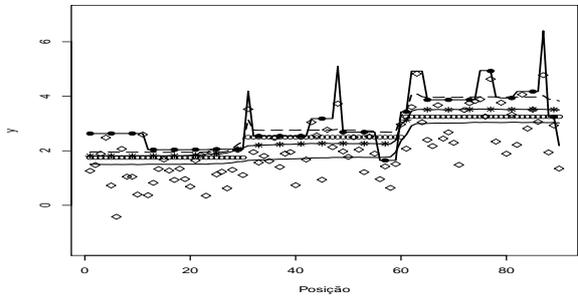
Figura 3.59: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 30$ .



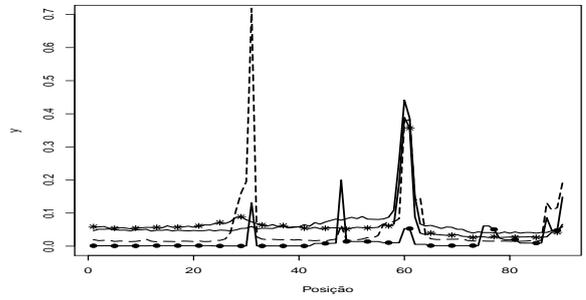
(a) Mod. gerado  $\lambda = 0$



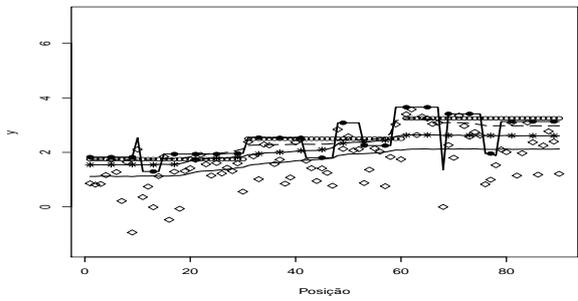
(b) Mod. gerado  $\lambda = 0$



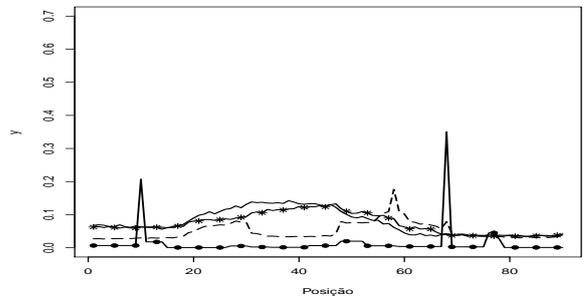
(c) Mod. gerado  $\lambda = -0.7$



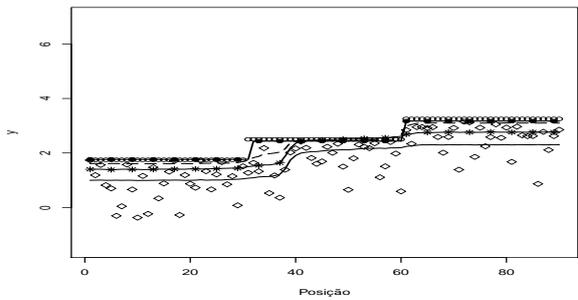
(d) Mod. gerado  $\lambda = -0.7$



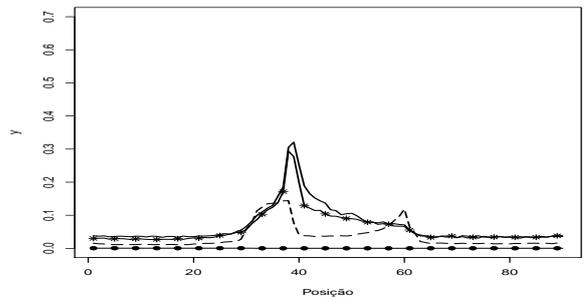
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

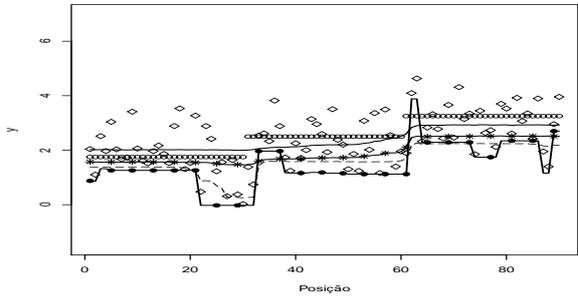


(g) Mod. gerado  $\lambda = -100$

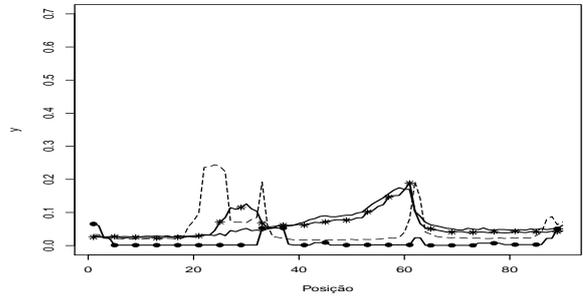


(h) Mod. gerado  $\lambda = -100$

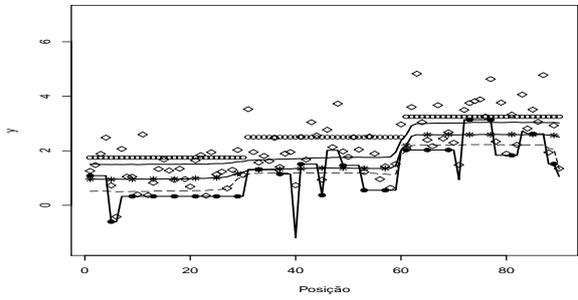
Figura 3.60: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 90$ .



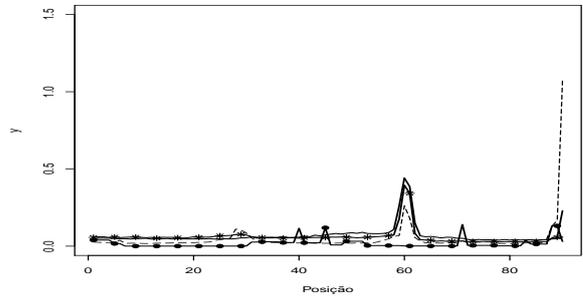
(a) Mod. gerado  $\lambda = 0$



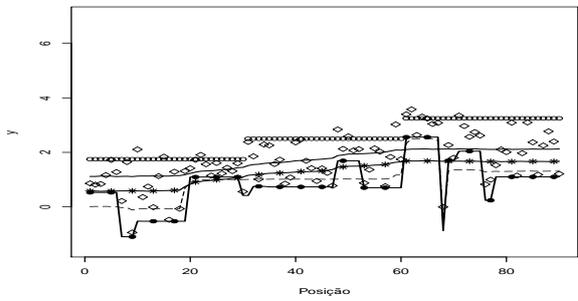
(b) Mod. gerado  $\lambda = 0$



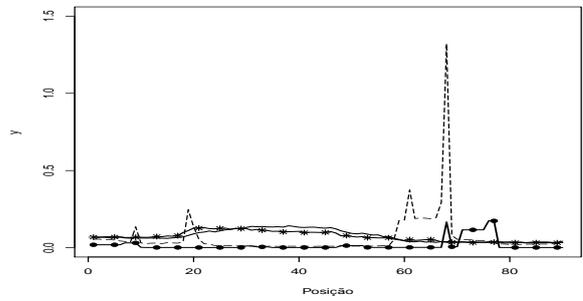
(c) Mod. gerado  $\lambda = -0.7$



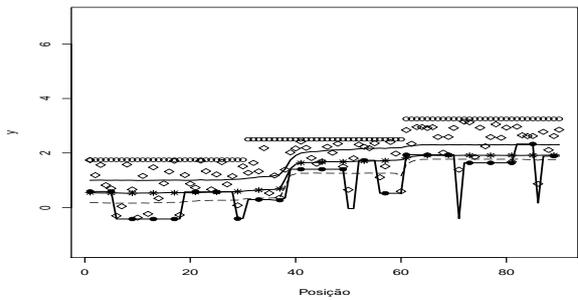
(d) Mod. gerado  $\lambda = -0.7$



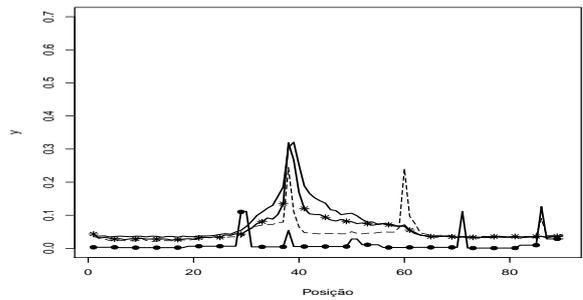
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

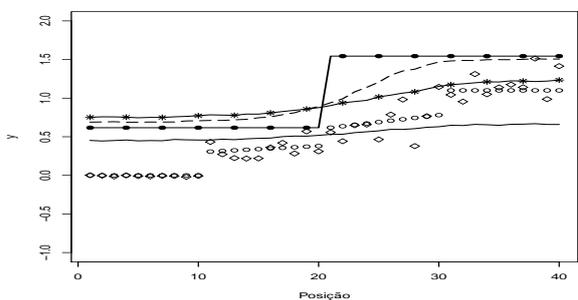


(g) Mod. gerado  $\lambda = -100$

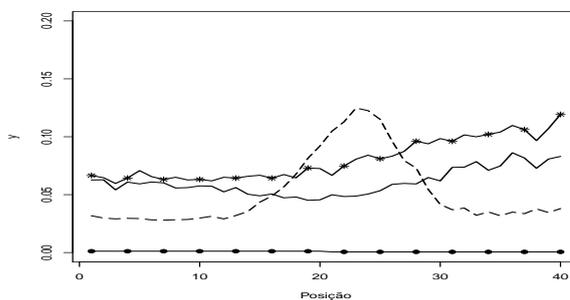


(h) Mod. gerado  $\lambda = -100$

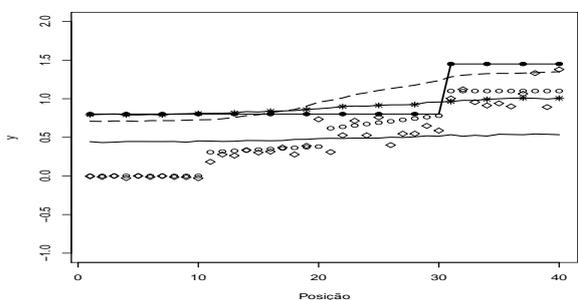
Figura 3.61: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 90$ .



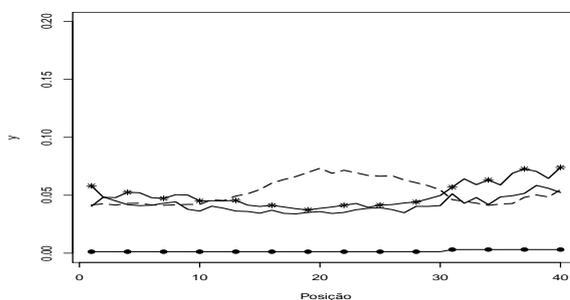
(a) Mod. gerado  $\lambda = 0$



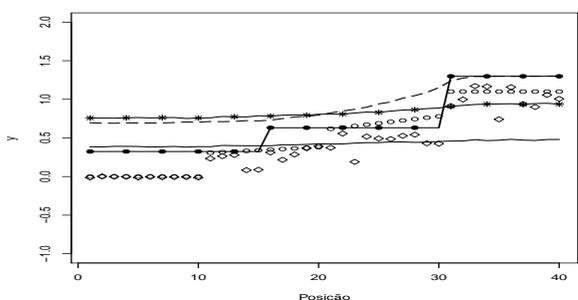
(b) Mod. gerado  $\lambda = 0$



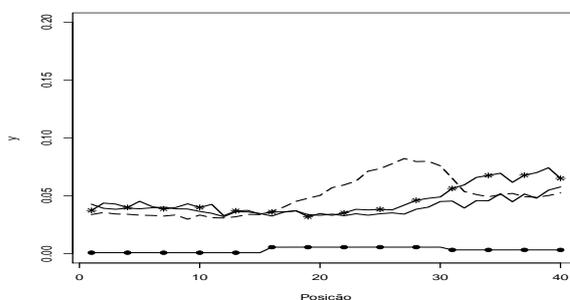
(c) Mod. gerado  $\lambda = -0.7$



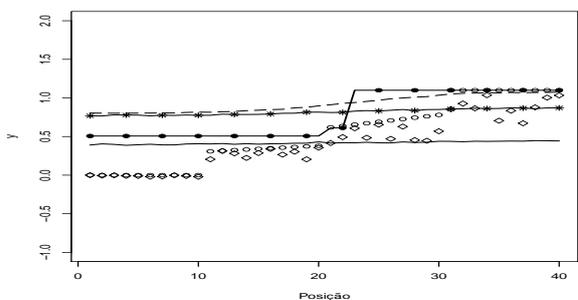
(d) Mod. gerado  $\lambda = -0.7$



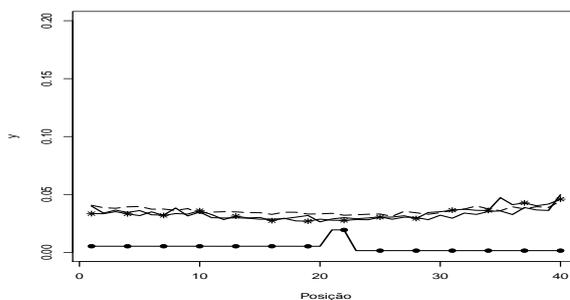
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

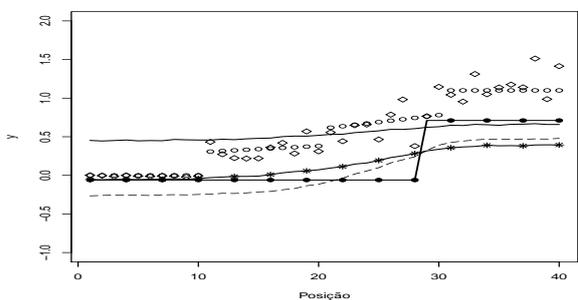


(g) Mod. gerado  $\lambda = -100$

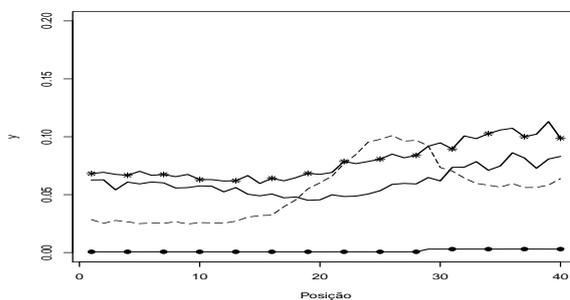


(h) Mod. gerado  $\lambda = -100$

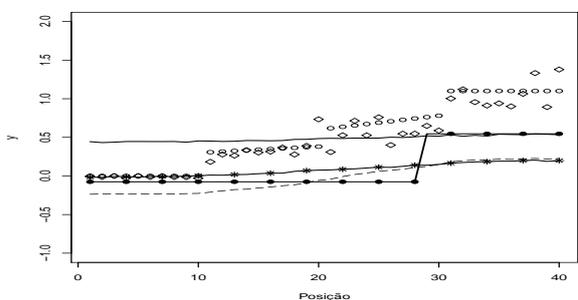
Figura 3.62: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 30$



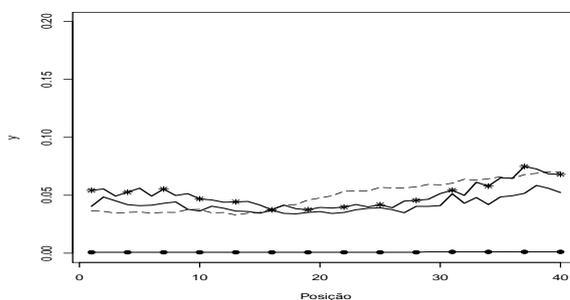
(a) Mod. gerado  $\lambda = 0$



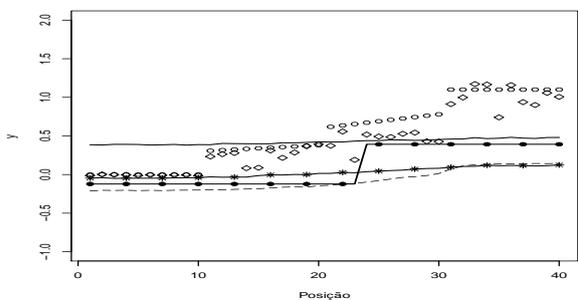
(b) Mod. gerado  $\lambda = 0$



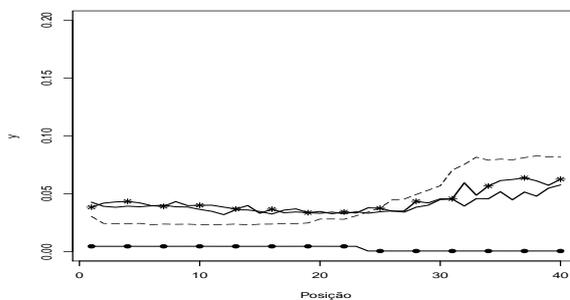
(c) Mod. gerado  $\lambda = -0.7$



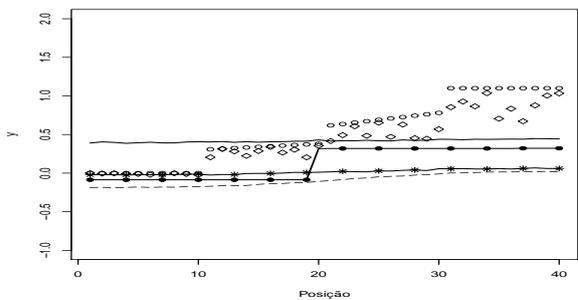
(d) Mod. gerado  $\lambda = -0.7$



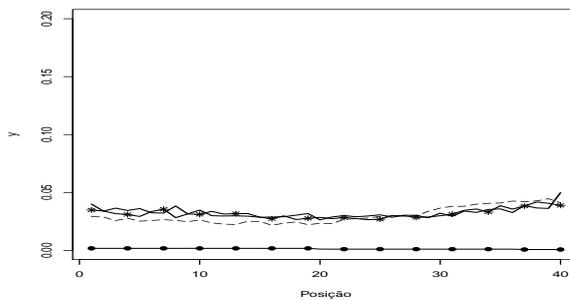
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

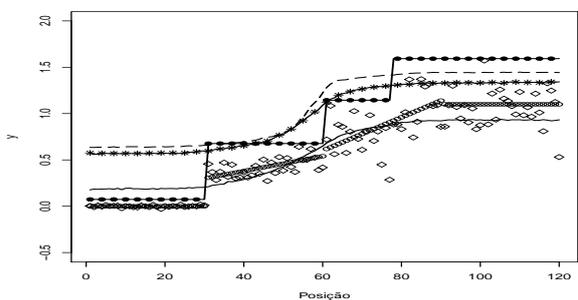


(g) Mod. gerado  $\lambda = -100$

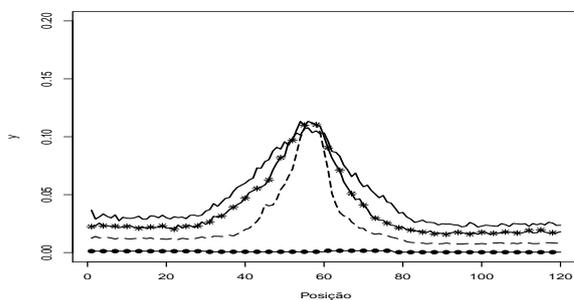


(h) Mod. gerado  $\lambda = -100$

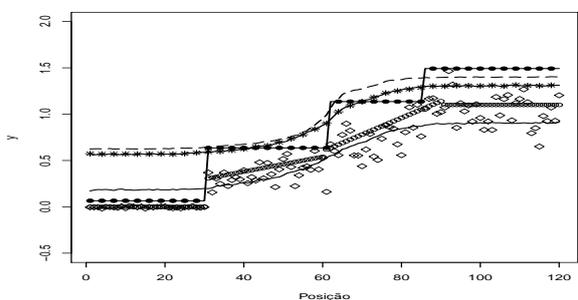
Figura 3.63: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 30$ .



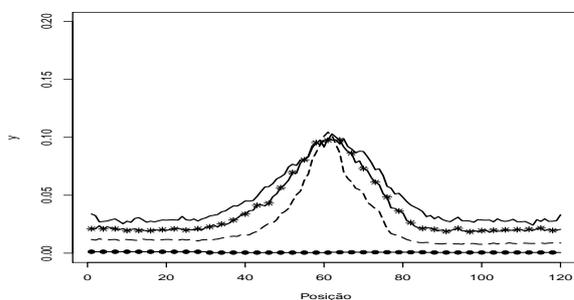
(a) Mod. gerado  $\lambda = 0$



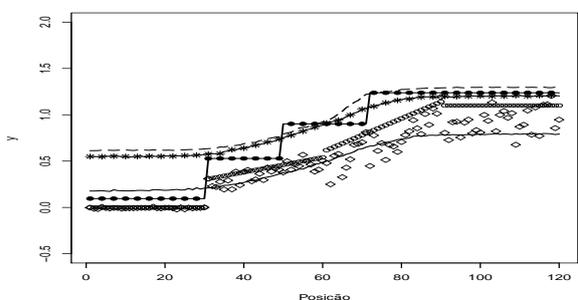
(b) Mod. gerado  $\lambda = 0$



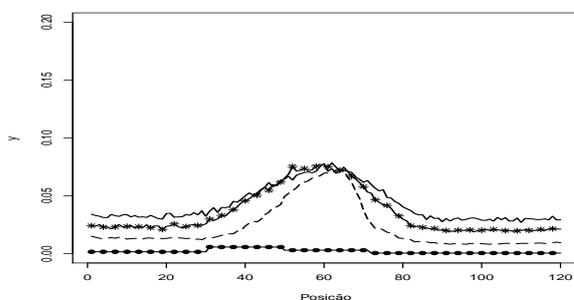
(c) Mod. gerado  $\lambda = -0.7$



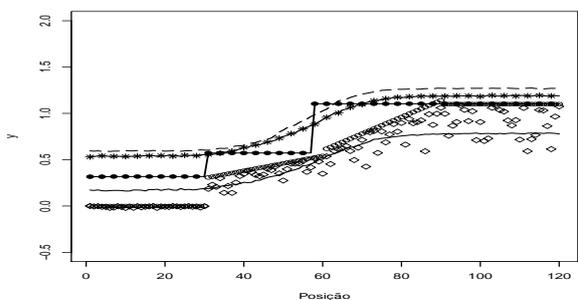
(d) Mod. gerado  $\lambda = -0.7$



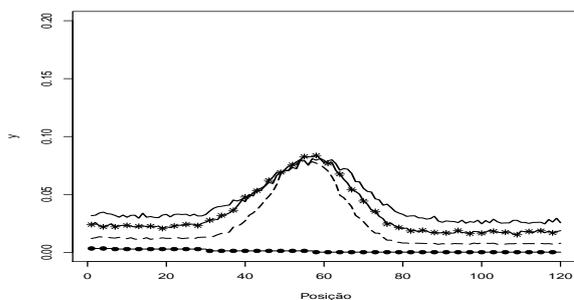
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$

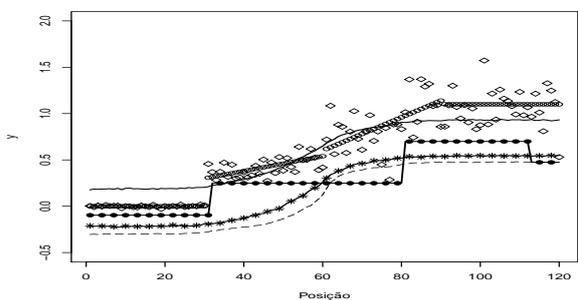


(g) Mod. gerado  $\lambda = -100$

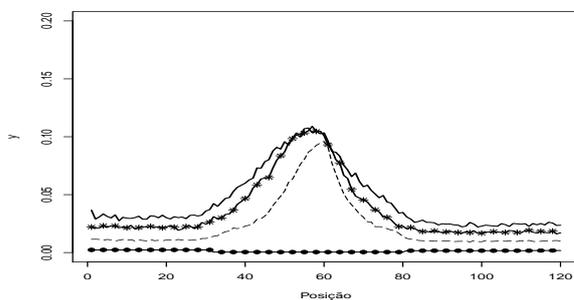


(h) Mod. gerado  $\lambda = -100$

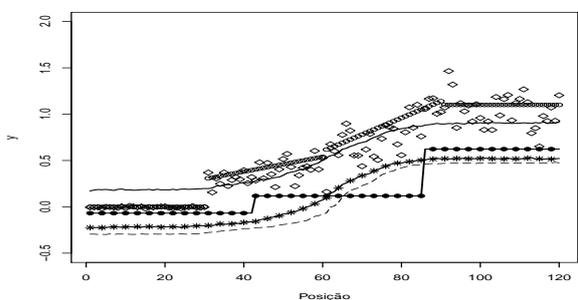
Figura 3.64: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 90$ .



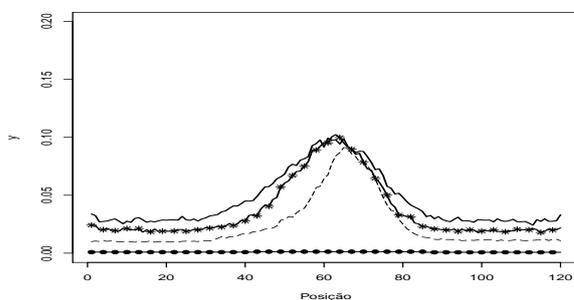
(a) Mod. gerado  $\lambda = 0$



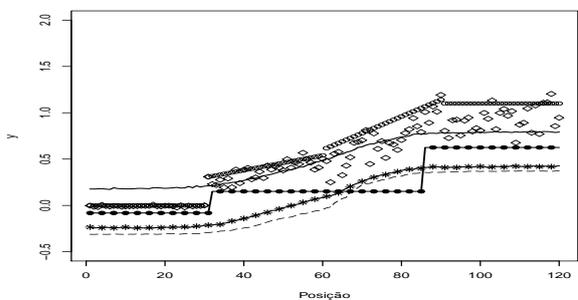
(b) Mod. gerado  $\lambda = 0$



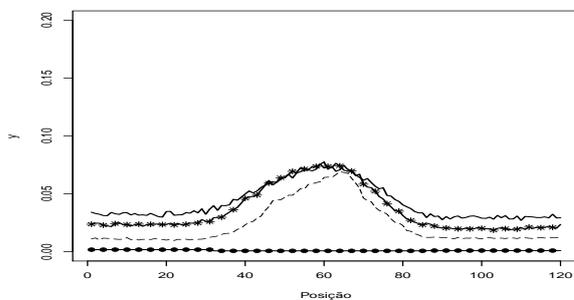
(c) Mod. gerado  $\lambda = -0.7$



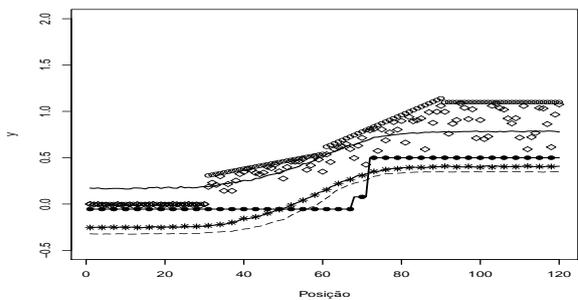
(d) Mod. gerado  $\lambda = -0.7$



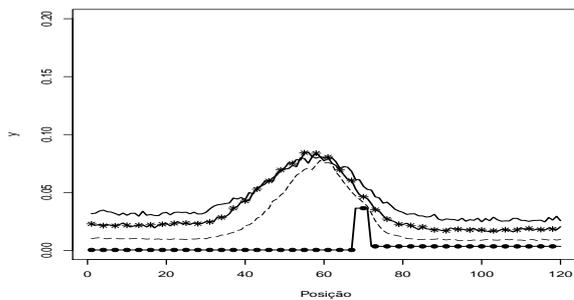
(e) Mod. gerado  $\lambda = -2$



(f) Mod. gerado  $\lambda = -2$



(g) Mod. gerado  $\lambda = -100$



(h) Mod. gerado  $\lambda = -100$

Figura 3.65: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 90$ .

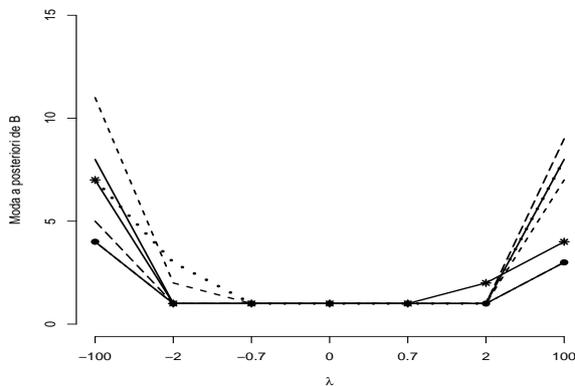
De forma geral, observa-se que as estimativas produto são subestimadas se for ajustado  $\lambda < 0$  e os dados forem gerados assumindo  $\lambda > 0$  e superestimados se os dados são gerados assumindo  $\lambda > 0$  e ajusta-se o modelo com  $\lambda < 0$ . Se o modelo ajustado assume que  $\lambda = \pm 100$ , quando na verdade o parâmetro de assimetria é  $\lambda = \pm 0.7$  ou 0, mesmo com o aumento do tamanho da amostra, as estimativas produto são ruins.

### 3.2.2 A influência de $\lambda$ e $n$ nas inferências sobre $\rho$ e $B$

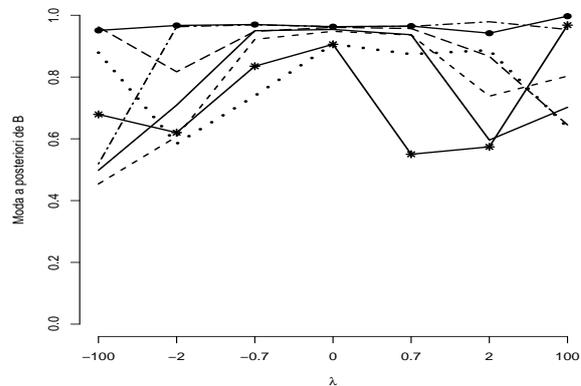
Nesta seção serão apresentados os resultados para a moda *a posteriori* do número de blocos e a probabilidade da moda *a posteriori*. Para interpretar e avaliar o número de blocos em relação a  $\lambda$  ajustado, a linha contínua indica o modelo sob o qual os dados foram gerados. Por exemplo, a linha cheia indica que os dados foram gerados assumindo  $\lambda = 0$ , assim, ao acompanhar o eixo X é possível identificar quantos blocos cada  $\lambda$  estimou assumindo que  $\lambda$  gerado é zero (linha cheia). Desta forma, na Figura 3.66(a), que ilustra o cenário sem mudança, acompanhando a linha cheia que indica que os dados foram gerados com  $\lambda = 0$ , nota-se que quando o modelo ajustado assume  $\lambda = -100$  (eixo X) o número de blocos estimado (eixo Y) é 8.

No cenário sem mudança, observa-se na Figura 3.66 que quando o tamanho de amostra é 30, o número de pontos de mudança é identificado corretamente se os dados são ajustados assumindo  $\lambda = 0, \pm 0.7$ , independente de qual tenha sido o parâmetro de assimetria dos dados simulados. No caso do tamanho de amostra 90, o número de blocos é identificado corretamente se  $\lambda = 0$  para todos os conjunto de dados simulados.

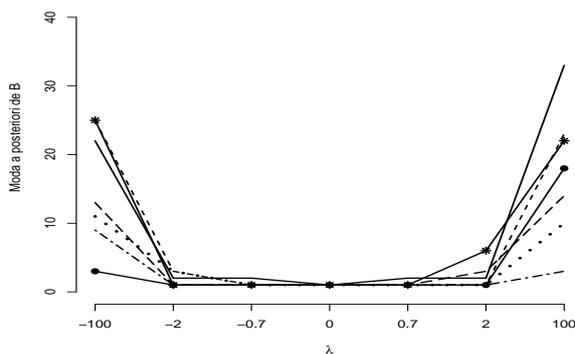
No entanto, nota-se que para dados gerados de  $\lambda = 0$  a moda *a posteriori* de  $B = 1$  só ocorre quando os dados são ajustados assumindo  $\lambda = 0$ . Para  $n = 90$ , Figura 3.66(c) e (d), observa-se que  $\lambda = -2$  resulta na maior probabilidade *a posteriori* de  $B = 2$  quando os dados foram gerados assumindo  $\lambda = 0$  (linha cheia), mas o número correto de blocos é 1, no cenário sem mudança.



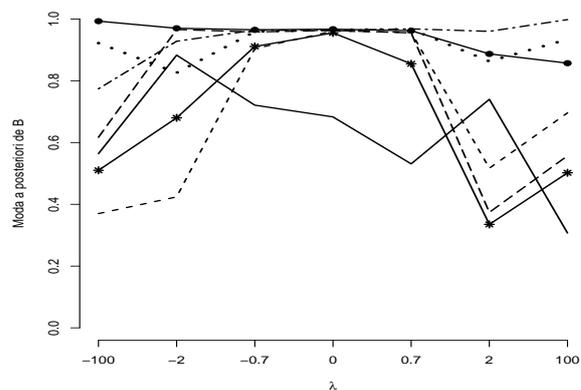
(a) Moda *a posteriori* do N. de Blocos -  $n = 30$



(b)  $P(B = b')$  -  $n = 30$



(c) Moda *a posteriori* do N. de Blocos -  $n = 90$

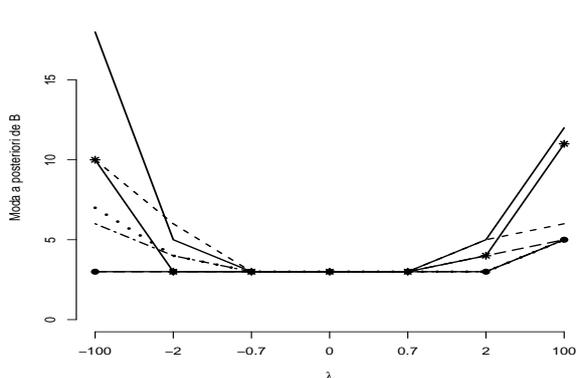


(d)  $P(B = b')$  -  $n = 90$

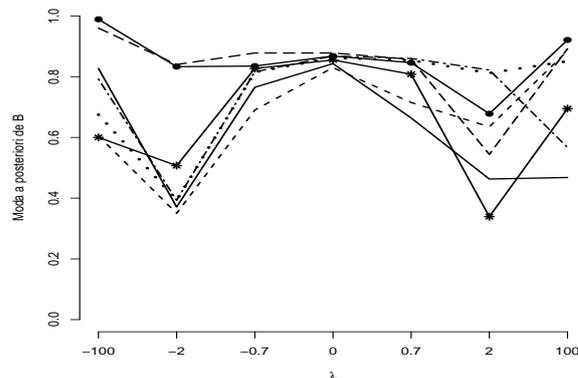
Figura 3.66: Moda *a posteriori* de  $B$  e sua probabilidade para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada). Cenário sem mudança.

A Figura 3.67 mostra os resultados do número de blocos para o cenário degrau não simétrico. Observa-se na Figura 3.67 que para o tamanho de amostra 90 que o modelo identifica corretamente o número de blocos independente do valor de  $\lambda$  gerado se os dados forem ajustados por  $\lambda = 0, -0.7$ , mas se os dados forem gerados assumindo  $\lambda = \pm 100$  o modelo que assume corretamente o valor de  $\lambda = \pm 100$  foi capaz de identificar corretamente o número de blocos  $B = 3$ , com maior probabilidade. Para  $n = 30$  e os dados gerados de  $\lambda = -100$  o modelo ajustado corretamente com  $\lambda = -100$

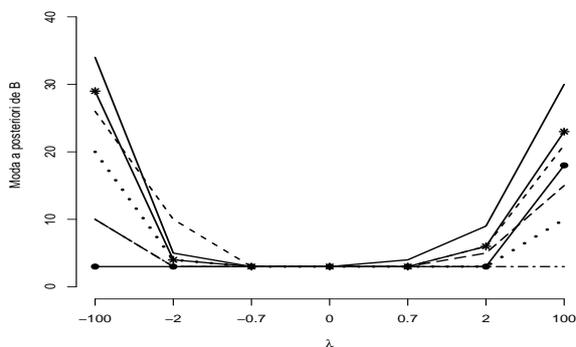
também resultou uma probabilidade superior de identificar corretamente o número de blocos quando comparado ao modelo que assume  $\lambda = 0$ , mas para o modelo gerado e ajustado corretamente com  $\lambda = 100$  o número de blocos mais provável *a posteriori* é 5.



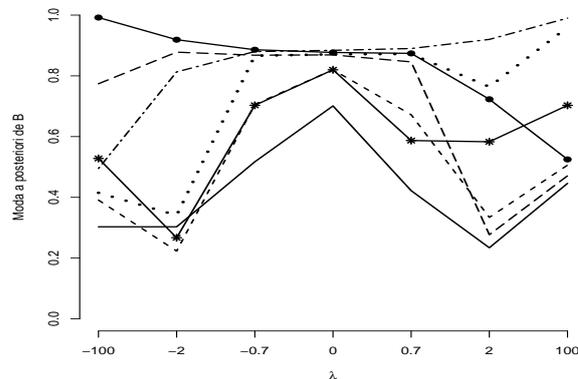
(a) Moda *a posteriori* do N. de Blocos -  $n = 30$



(b)  $P(B == b')$  -  $n = 30$



(c) Moda *a posteriori* do N. de Blocos -  $n = 90$



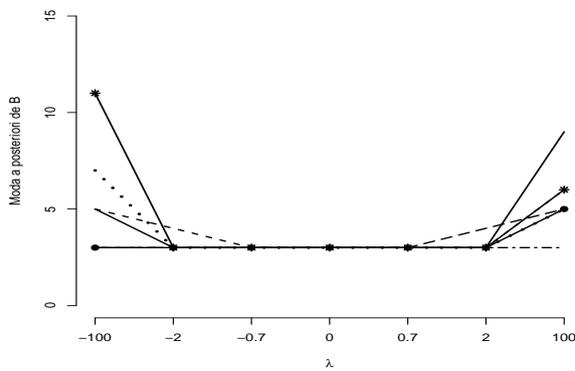
(d)  $P(B = b')$  -  $n = 90$

Figura 3.67: Moda *a posteriori* de  $B$  e sua probabilidade para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário degrau não simétrico.

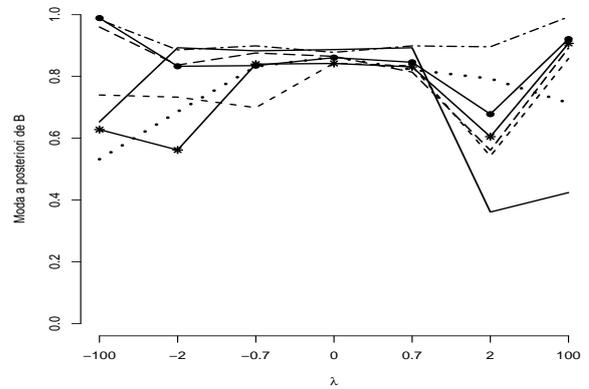
A Figura 3.68 mostra os resultados para o cenário degrau positivo, o cenário degrau negativo apresenta resultados similares e encontra-se no Anexo D. Nota-se que para o tamanho de amostra 30, a Figura 3.68(a) e (b) revela que  $\lambda = 0$  estima bem o número de blocos para todos os valores de  $\lambda$

simulados, mas a probabilidade de indicar corretamente o número de blocos é bem inferior quando comparada ao modelo que ajusta  $\lambda = 100$  aos dados que de fato foram gerados de  $\lambda = 100$ .

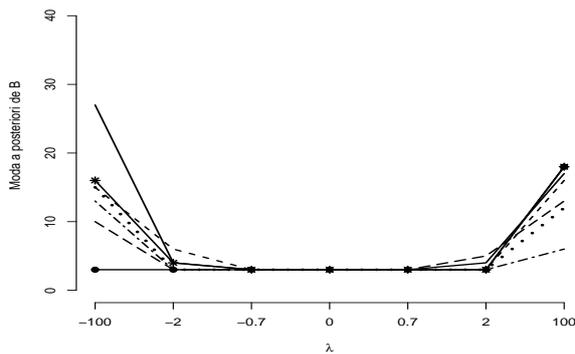
Observa-se na Figura 3.68,  $n = 90$ , que quando os dados são gerados de  $\lambda = -100$  e ajusta-se o modelo com  $\lambda = -100, -2, -0.7, 0, 0.7, 2$  o número de blocos é corretamente estimado, mas a maior probabilidade do modelo indicar  $B = 3$  é observada para o modelo que assume  $\lambda = -100$ .



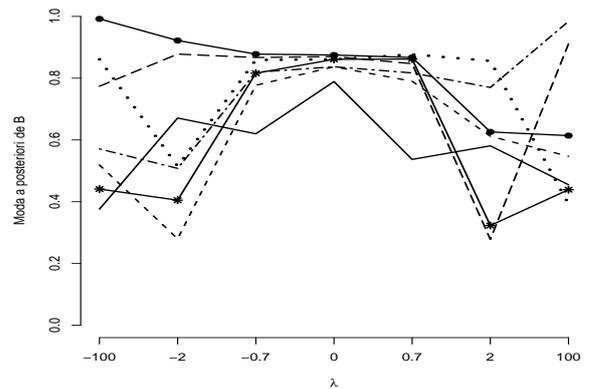
(a) Moda *a posteriori* do N. de Blocos -  $n = 30$



(b)  $P(B = b')$  -  $n = 30$



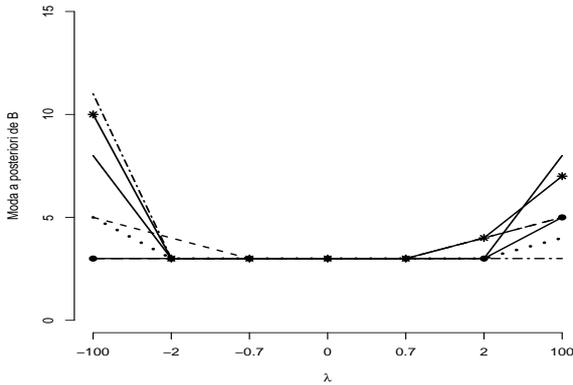
(c) Moda *a posteriori* do N. de Blocos -  $n = 90$



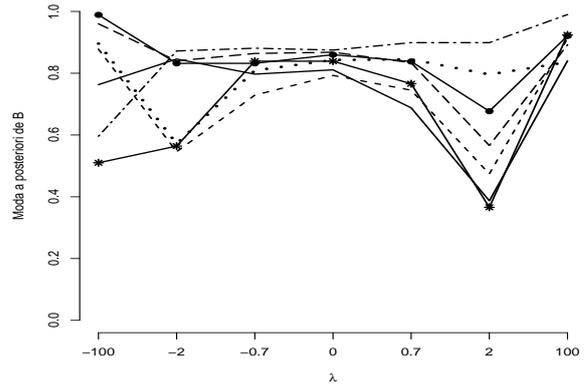
(d)  $P(B = b')$  -  $n = 90$

Figura 3.68: Moda *a posteriori* de  $B$  e sua probabilidade para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada). Cenário degrau positivo.

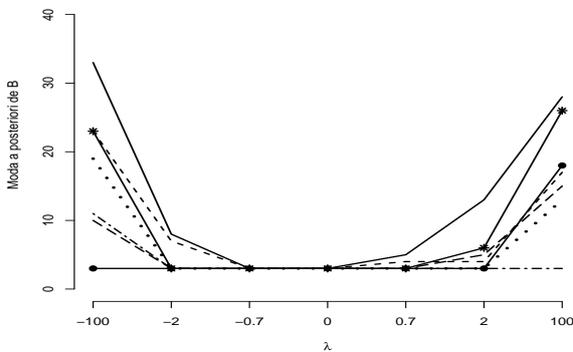
A Figura 3.69 apresenta os resultados para o cenário escada crescente e finaliza o estudo de sensibilidade para o número de blocos *a posteriori*, o cenário decrescente também foi avaliado e como os resultados foram similares ao cenário escada crescente os resultados foram suprimidos dessa seção e encontram-se no Anexo D.



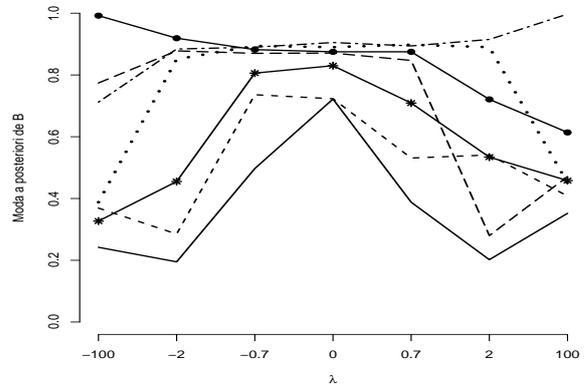
(a) Moda *a posteriori* do N. de Blocos -  $n = 30$



(b)  $P(B = b')$  -  $n = 30$



(c) Moda *a posteriori* do N. de Blocos -  $n = 90$



(d)  $P(B = b')$  -  $n = 90$

Figura 3.69: Moda *a posteriori* de  $B$  e sua probabilidade para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada). Cenário escada crescente.

Em relação ao número de blocos do cenário escada crescente, Figura 3.69, para ambos os tamanhos de amostra,  $\lambda = \pm 100$  resultou na maior probabilidade *a posteriori* de indicar corretamente o

número de blocos quando os dados foram gerados assumindo  $\lambda = \pm 100$ .

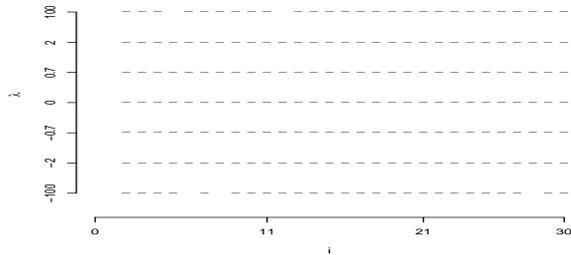
Nota-se ainda, na Figura 3.69 que para dados gerados de  $\lambda = \pm 2$ , linhas tracejada e longa e linha pontilhada, o melhor ajuste ocorre assumindo  $\lambda = 100$ , linha tracejada e pontilhada, para o tamanho de amostra 30, mas para  $n = 90$  as probabilidades para  $\lambda = -100$  e  $\lambda = -2$  quando os dados são gerados de  $\lambda = -2$  são próximas, o mesmo acontece para  $\lambda = 2$ .

Ao avaliar os resultados do número de blocos estimados em função de  $\lambda$  observa-se que mesmo quando os dados são gerados de um parâmetro de assimetria alto ( $\lambda = \pm 100$ ),  $\lambda = 0$  é capaz de estimar corretamente o número de blocos, mesmo que com uma probabilidade mais baixa em alguns casos. No entanto, se os dados são gerados assumindo  $\lambda = \pm 100$  e ajusta-se o modelo corretamente com  $\lambda = \pm 100$ , exceto para o cenário sem mudança e o cenário degrau não simétrico  $n = 30$  e degrau positivo  $n = 90$ , o número de blocos é corretamente indicado com alta probabilidade *a posteriori*.

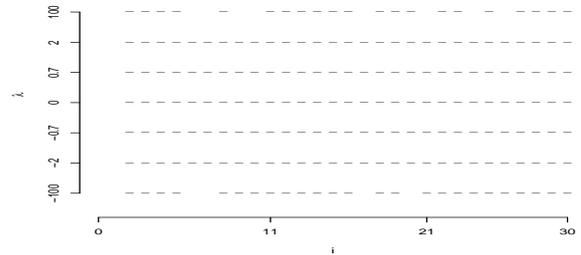
Para todos os cenários avaliados, se os dados são gerados com um parâmetro de assimetria,  $\lambda = 0, \pm 0.7, \pm 2$  e ajusta-se o modelo assumindo que  $\lambda = \pm 100$  o número de blocos é superestimado.

As Figuras 3.70, 3.71, 3.72, 3.73, 3.74, 3.75, 3.76 e 3.77 apresentam as partições mais prováveis para cada cenário discutido nessa seção, cada gráfico dentro das figuras representa um valor de  $\lambda$  simulado para aquele cenário, por exemplo, a Figura 3.70 (a) refere-se aos dados gerados assumindo  $\lambda = -100$  no cenário sem mudança, o eixo (Y) da Figura 3.70 (a) ilustra os valores de  $\lambda$  que foram ajustados e os traços dentro do gráfico indicam a partição mais provável, onde o espaço em branco representa um ponto de mudança e os traços o bloco estimado.

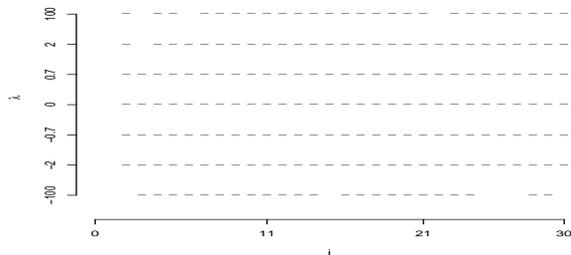
Desta forma, ao avaliar a partição mais provável quando os dados são simulados assumindo  $\lambda = 0$  e ajustados com  $\lambda = \pm 100$ , Figura 3.70 (d), nota-se que o modelo indica erroneamente que a partição possui cerca de 7 à 8 blocos quando na verdade o cenário correto é sem mudança. Com um tamanho de amostra maior,  $n = 90$ , observa-se o mesmo comportamento, uma partição com vários blocos quando na verdade o cenário é sem mudança.



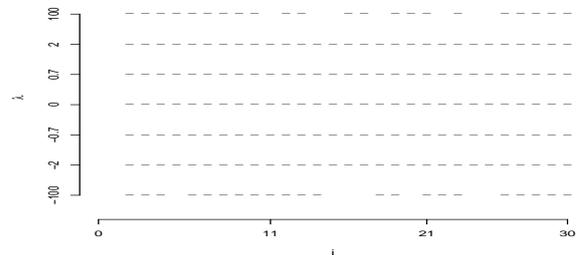
(a) Partição mais provável - Dados  $\lambda = -100$



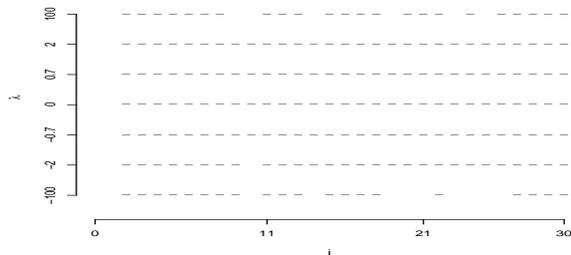
(b) Partição mais provável - Dados  $\lambda = -2$



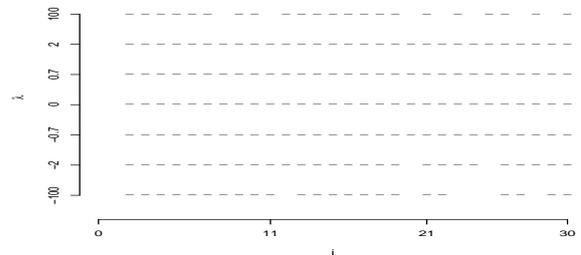
(c) Partição mais provável - Dados  $\lambda = -0.7$



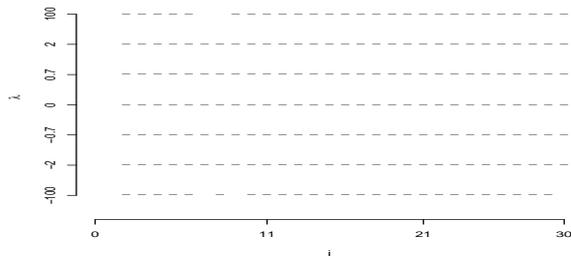
(d) Partição mais provável - Dados  $\lambda = 0$



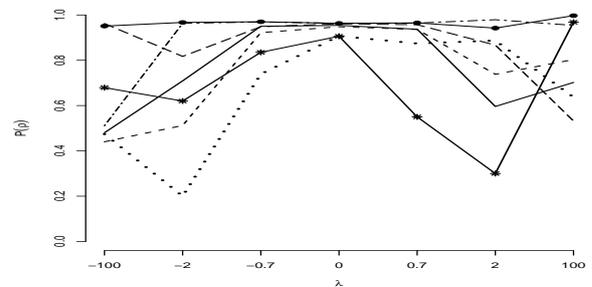
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

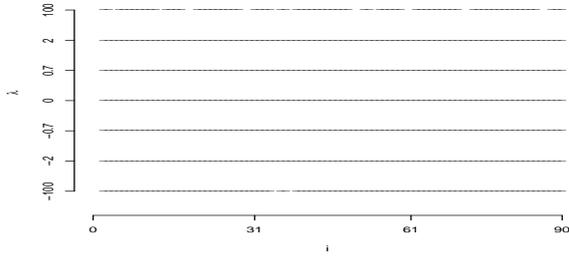


(g) Partição mais provável - Dados  $\lambda = 100$

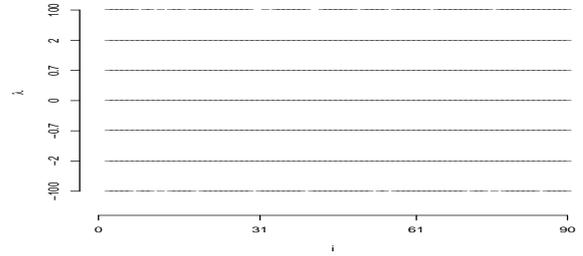


(h)  $P(\rho = \rho')$

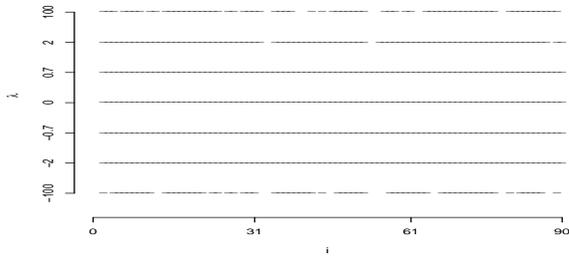
Figura 3.70: Partição mais provável *a posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário sem mudança,  $n = 30$ .



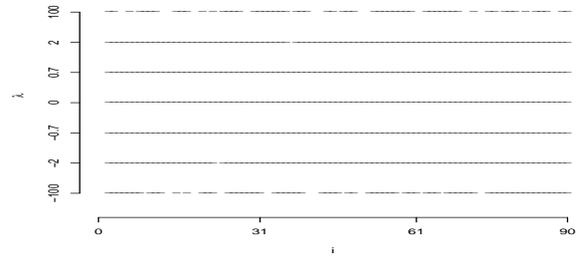
(a) Partição mais provável - Dados  $\lambda = -100$



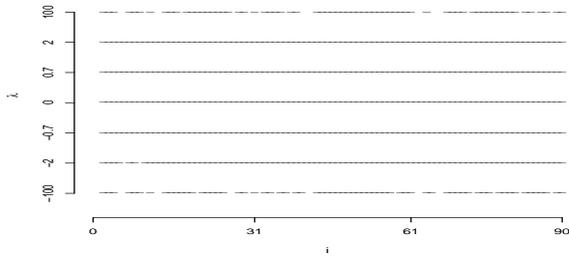
(b) Partição mais provável - Dados  $\lambda = -2$



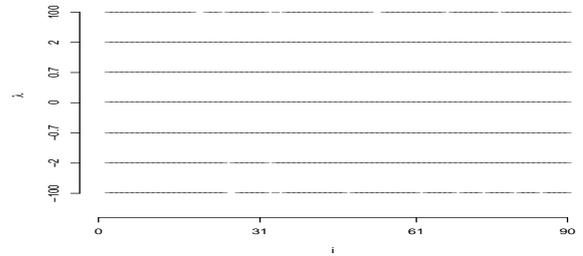
(c) Partição mais provável - Dados  $\lambda = -0.7$



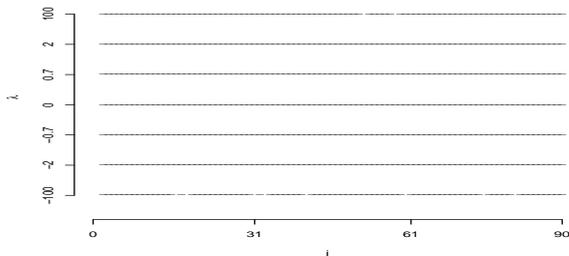
(d) Partição mais provável - Dados  $\lambda = 0$



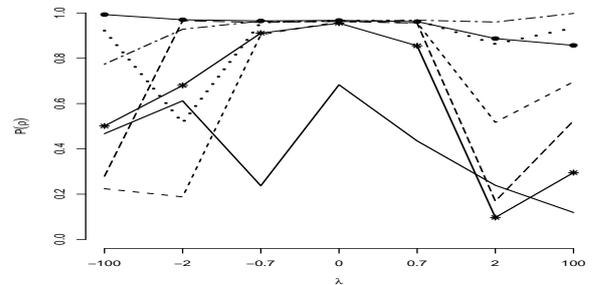
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$



(g) Partição mais provável - Dados  $\lambda = 100$



(h)  $P(\rho = \rho')$

Figura 3.71: Partição mais provável *a posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário sem mudança,  $n = 90$ .

Como foi visto na Figura 3.66 o número de blocos foi superestimado, mesmo quando o modelo foi gerado e ajustado corretamente com  $\lambda = \pm 100$ , logo a partição também foi estimada erroneamente como pode ser visto nas Figuras 3.70 e 3.71 (a) e (g), as partições indicam que existem pontos de mudança no cenário sem mudança.

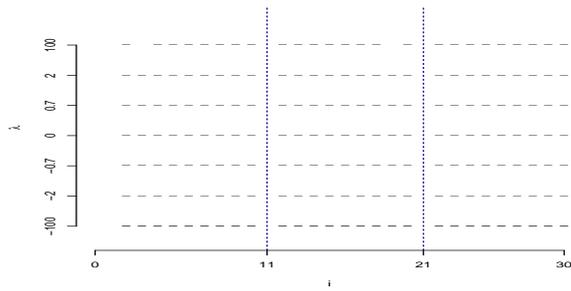
Mesmo que com uma probabilidade menor, em alguns casos, assumir o parâmetro de assimetria  $\lambda = 0, \pm 0.7$  resultou em uma estimação correta da partição de interesse no cenário sem mudança.

As Figuras 3.72 e 3.73 ilustram os resultados da partição mais provável para o cenário degrau não simétrico quando  $n = 30$  e  $n = 90$ , respectivamente.

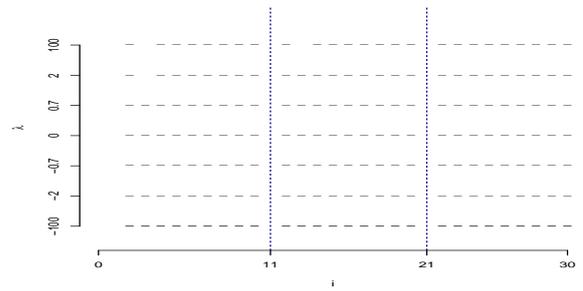
No cenário degrau não simétrico para  $n = 30$ , Figura 3.72(g), quando os dados foram gerados e ajustados assumindo  $\lambda = 100$  a partição foi identificada incorretamente com probabilidade 0.7, mas para  $n = 90$ , Figura 3.73(g), a partição foi identificada corretamente com alta probabilidade. No caso em que os dados foram gerados com  $\lambda = -100$  e modelados com  $\lambda = -100$  a partição correta é identificada como mais provável com alta probabilidade em ambos os tamanhos de amostra.

Nota-se na Figura 3.72(b) e (h) que para os dados gerados de  $\lambda = -2$  e ajustados com  $\lambda = -100$  a partição foi identificada corretamente com alta probabilidade, mas na Figura 3.73(b) e (h), observa-se que para  $n = 90$  a partição indicada como mais provável difere da partição correta.

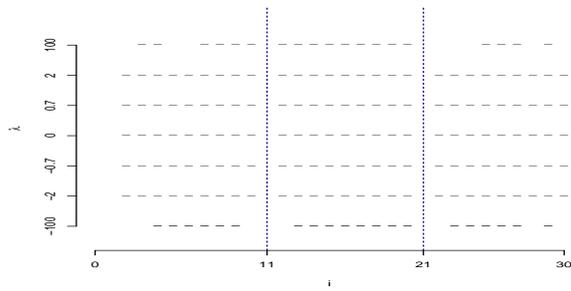
Outra característica que se pode observar nesse cenário é que se os dados são gerados assumindo  $\lambda = \pm 100$  e ajusta-se o modelo com  $\lambda = \pm 2$  a partição mais provável é identificada corretamente com uma probabilidade superior a 0.75 para ambos os tamanhos de amostra considerados.



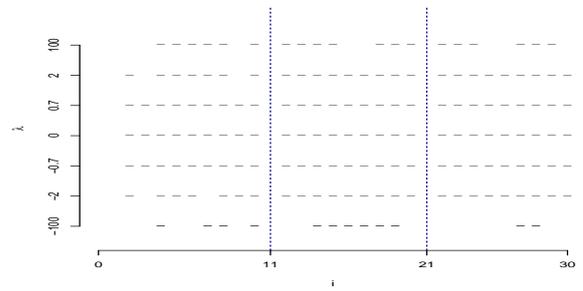
(a) Partição mais provável - Dados  $\lambda = -100$



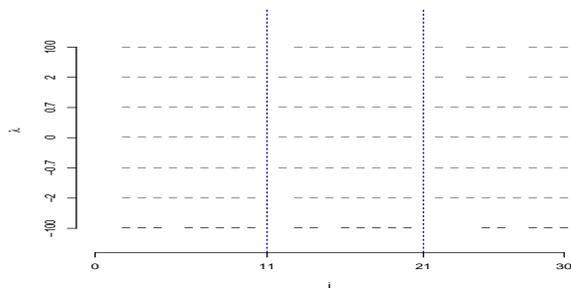
(b) Partição mais provável - Dados  $\lambda = -2$



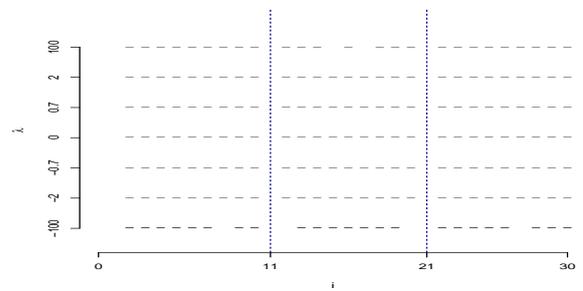
(c) Partição mais provável - Dados  $\lambda = -0.7$



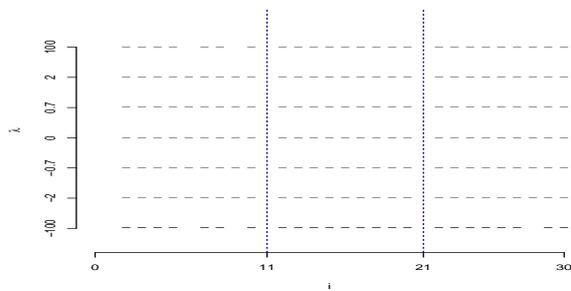
(d) Partição mais provável - Dados  $\lambda = 0$



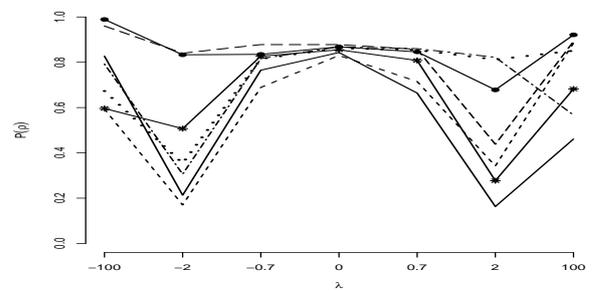
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

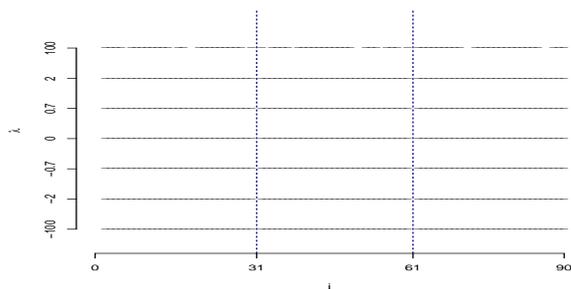


(g) Partição mais provável - Dados  $\lambda = 100$

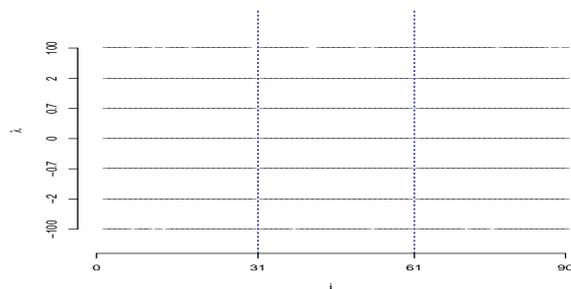


(h)  $P(\rho = \rho')$

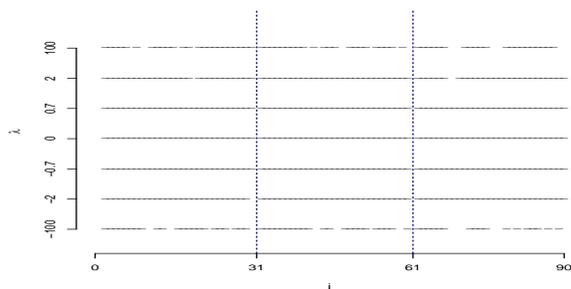
Figura 3.72: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário degrau não simétrico,  $n = 30$ .



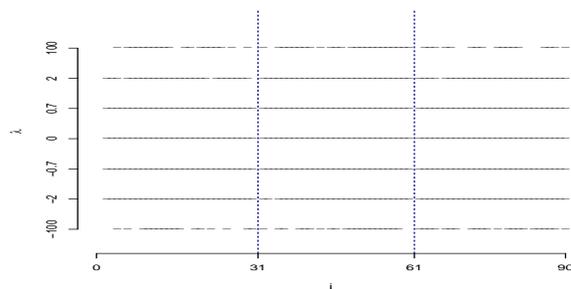
(a) Partição mais provável - Dados  $\lambda = -100$



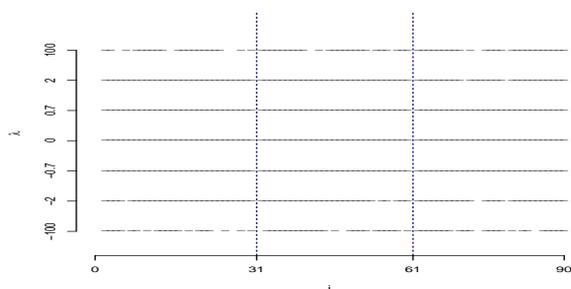
(b) Partição mais provável - Dados  $\lambda = -2$



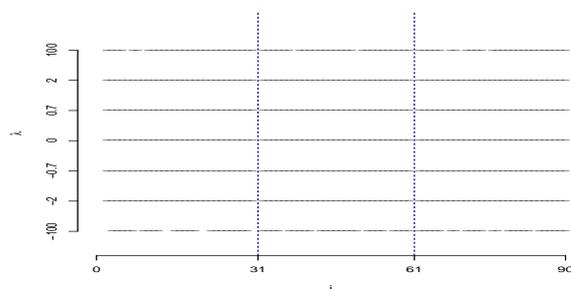
(c) Partição mais provável - Dados  $\lambda = -0.7$



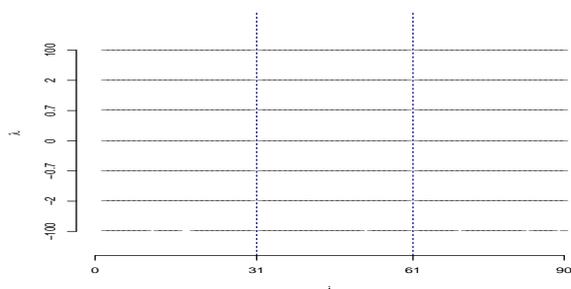
(d) Partição mais provável - Dados  $\lambda = 0$



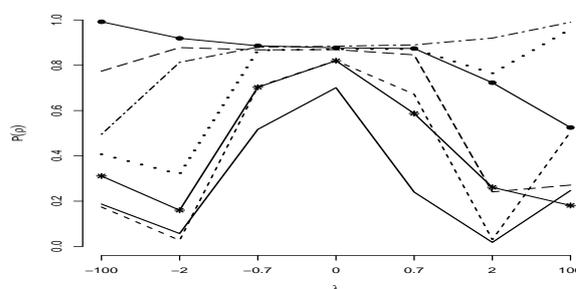
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$



(g) Partição mais provável - Dados  $\lambda = 100$



(h)  $P(\rho = \rho')$

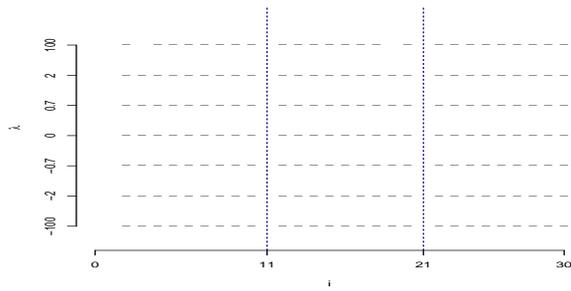
Figura 3.73: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário degrau não simétrico,  $n = 90$ .

As Figuras 3.74 e 3.75, a seguir, mostram os resultados em relação a partição mais provável para o cenário degrau positivo.

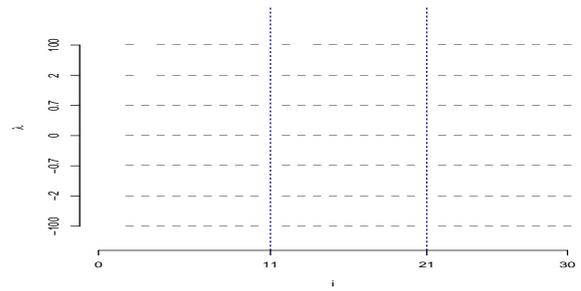
Nota-se nas Figuras 3.74(a), (g) e (h) e 3.75(a), (g) e (h), considerando que os dados foram gerados de  $\lambda = \pm 100$ , que apenas para  $n = 90$  e dados gerados e ajustados de  $\lambda = 100$  resultou em uma partição diferente da correta com alta probabilidade. Em relação a  $\lambda = 0$ , quando os dados foram gerados e ajustados assumindo  $\lambda = 0$  para  $n = 30$ , Figura 3.74(d), a partição mais provável foi próxima a gerada, mas não idêntica, porém para  $n = 90$  a partição foi identificada corretamente, Figura 3.75(d).

O comportamento do cenário epsilon é observado nas Figuras 3.78 e 3.79 e nota-se que a partição estimada como mais provável pelo Modelo II em nenhum momento foi igual a partição correta, independente do valor de  $\lambda$  e do tamanho da amostra, mas neste cenário a diferença entre os blocos é pequena o que dificulta a estimação do momento exato da mudança.

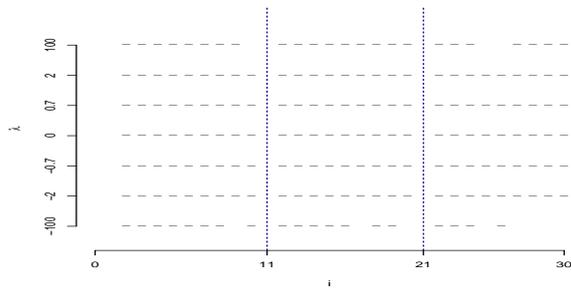
No cenário tendência cujo os resultados são apresentados nas Figuras 3.80 e 3.81, observa-se que na maior parte das combinações entre os valores de  $\lambda$  gerados e estimados, a partição correta é indicada como sendo sem mudança, no entanto, o cenário tendência possui 3 pontos de mudança. Nota-se também, que para  $n = 120$  com exceção de  $\lambda = \pm 100$  as partições são indicadas com baixa probabilidade.



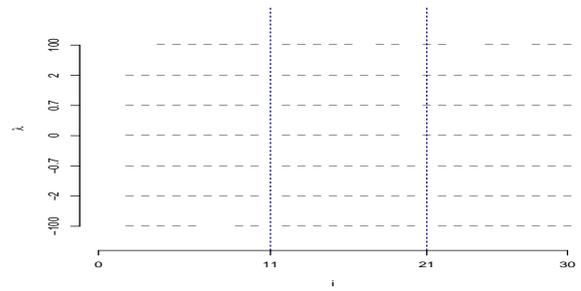
(a) Partição mais provável - Dados  $\lambda = -100$



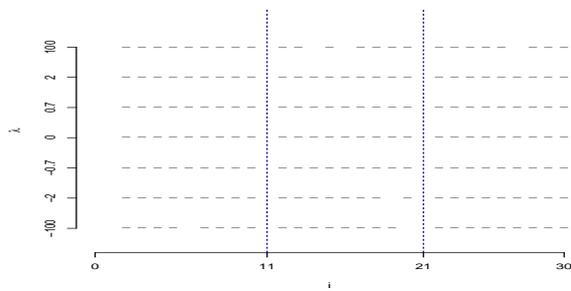
(b) Partição mais provável - Dados  $\lambda = -2$



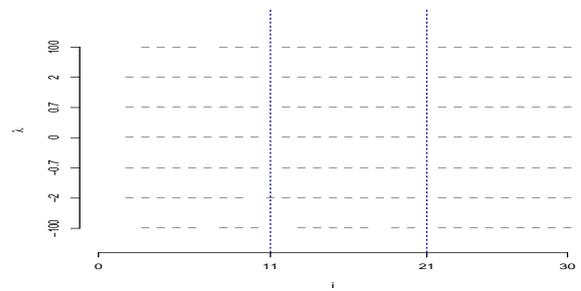
(c) Partição mais provável - Dados  $\lambda = -0.7$



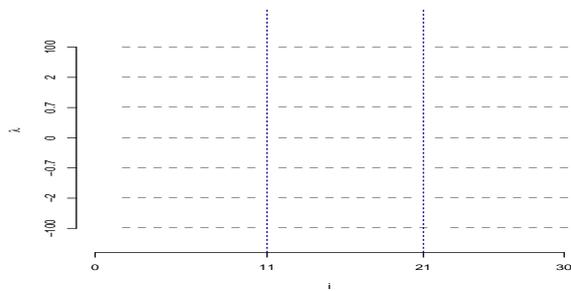
(d) Partição mais provável - Dados  $\lambda = 0$



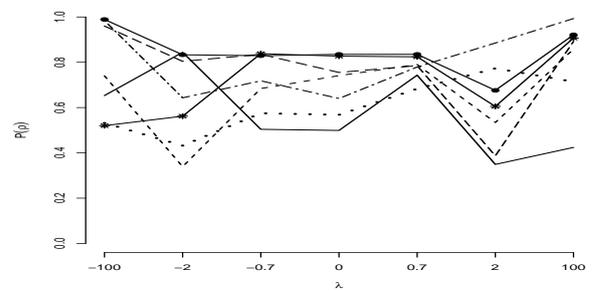
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

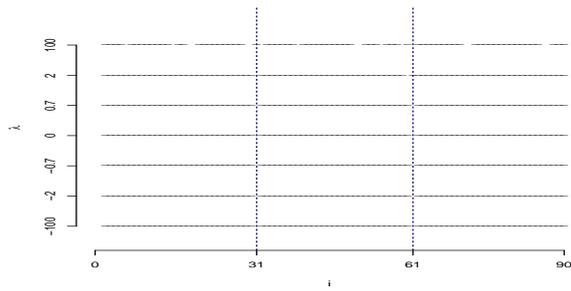


(g) Partição mais provável - Dados  $\lambda = 100$

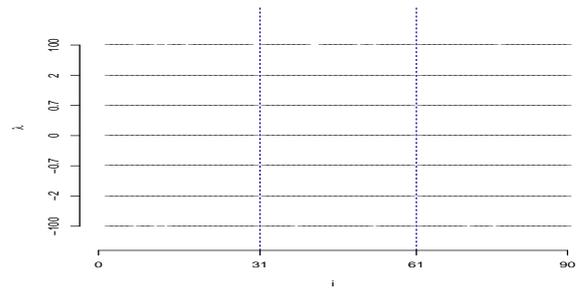


(h)  $P(\rho = \rho')$

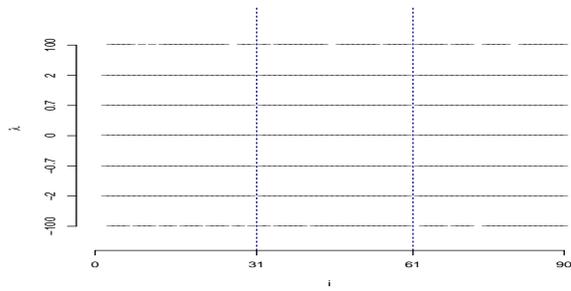
Figura 3.74: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário degrau positivo,  $n = 30$ .



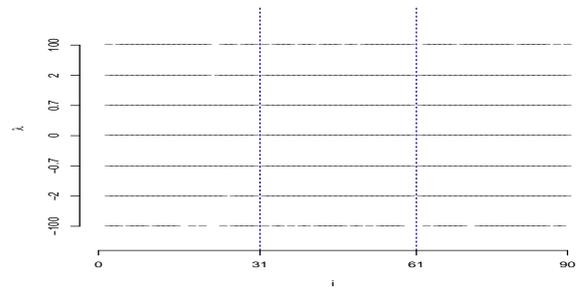
(a) Partição mais provável - Dados  $\lambda = -100$



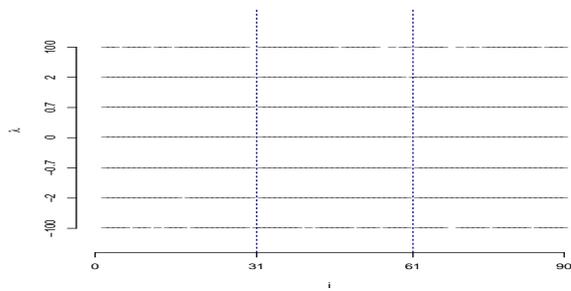
(b) Partição mais provável - Dados  $\lambda = -2$



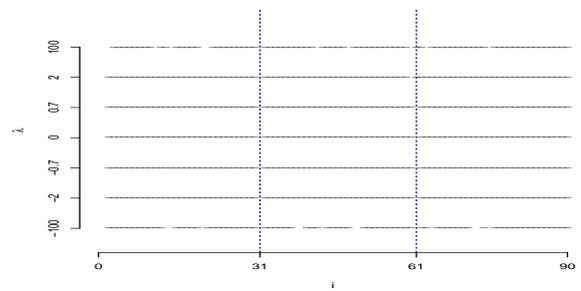
(c) Partição mais provável - Dados  $\lambda = -0.7$



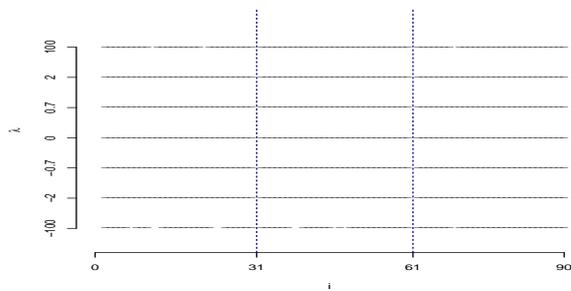
(d) Partição mais provável - Dados  $\lambda = 0$



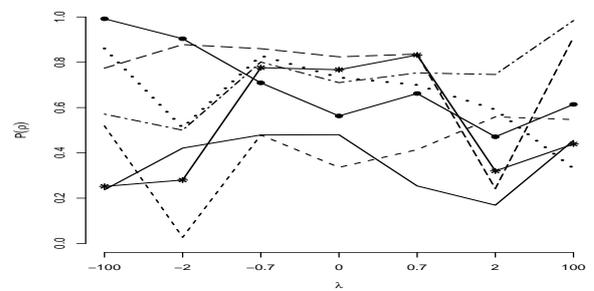
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

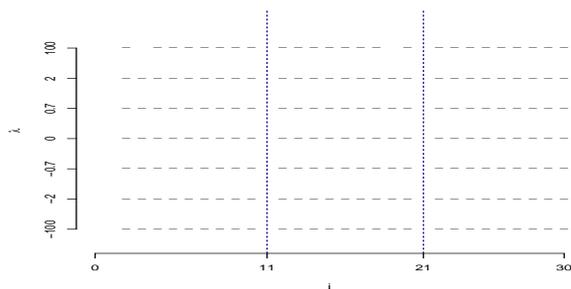


(g) Partição mais provável - Dados  $\lambda = 100$

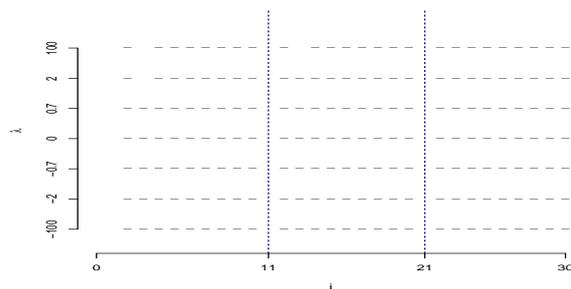


(h)  $P(\rho = \rho')$

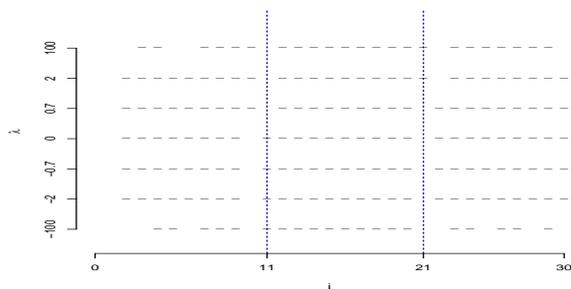
Figura 3.75: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário degrau positivo,  $n = 90$ .



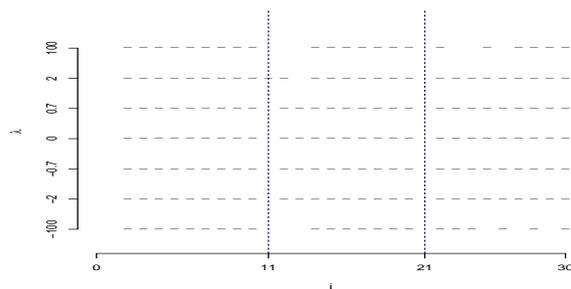
(a) Partição mais provável - Dados  $\lambda = -100$



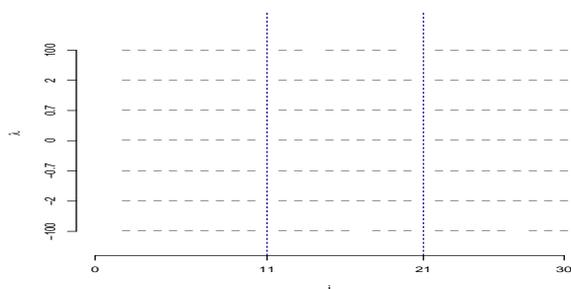
(b) Partição mais provável - Dados  $\lambda = -2$



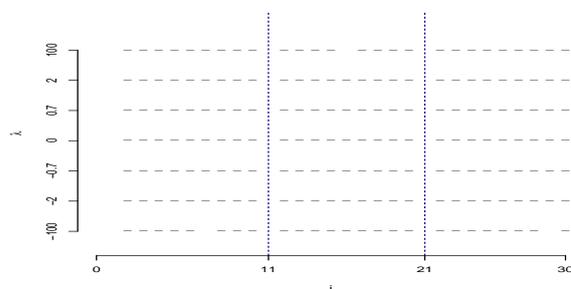
(c) Partição mais provável - Dados  $\lambda = -0.7$



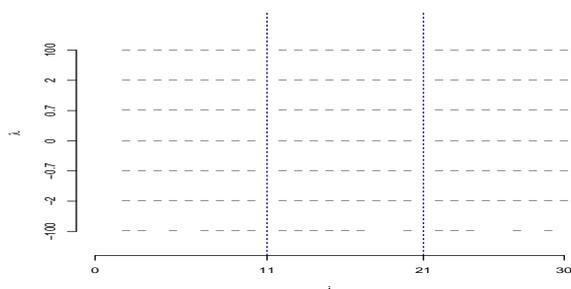
(d) Partição mais provável - Dados  $\lambda = 0$



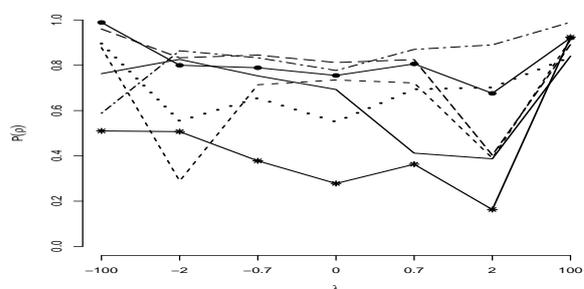
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

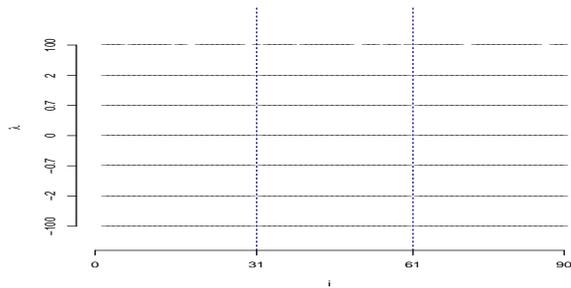


(g) Partição mais provável - Dados  $\lambda = 100$

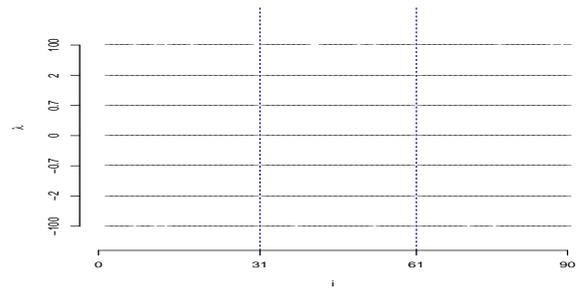


(h)  $P(\rho = \rho')$

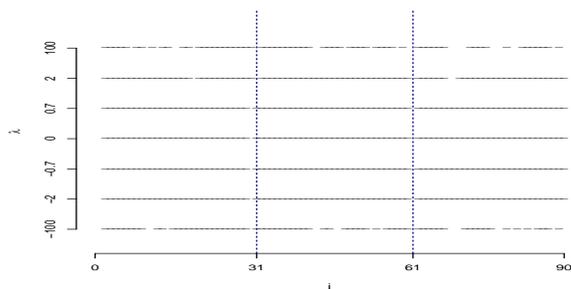
Figura 3.76: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário escada crescente,  $n = 30$ .



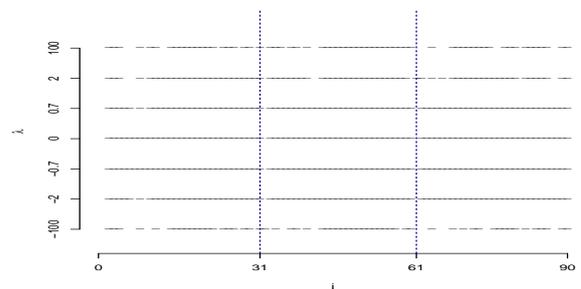
(a) Partição mais provável - Dados  $\lambda = -100$



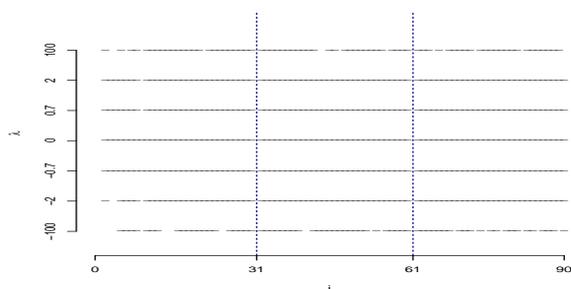
(b) Partição mais provável - Dados  $\lambda = -2$



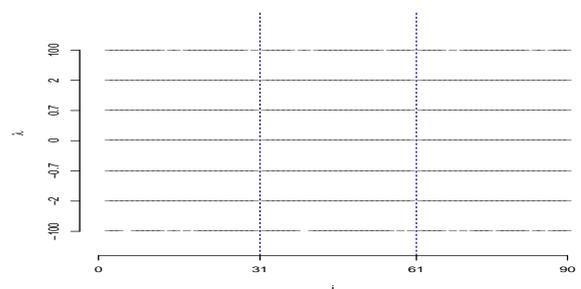
(c) Partição mais provável - Dados  $\lambda = -0.7$



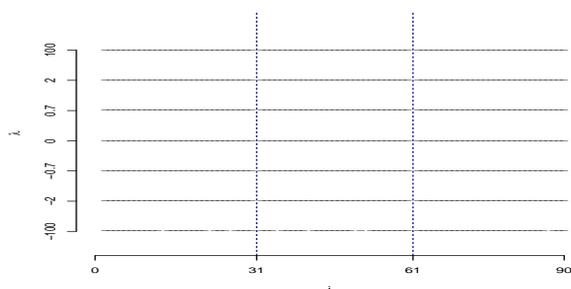
(d) Partição mais provável - Dados  $\lambda = 0$



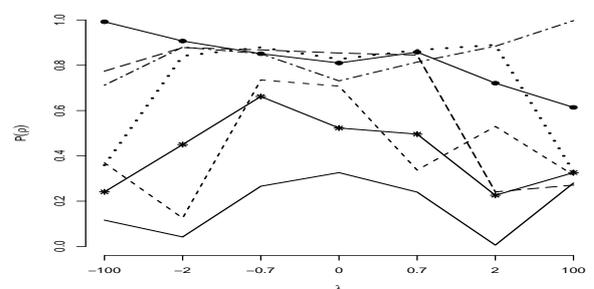
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

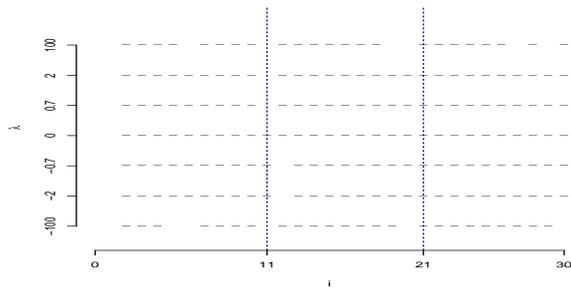


(g) Partição mais provável - Dados  $\lambda = 100$

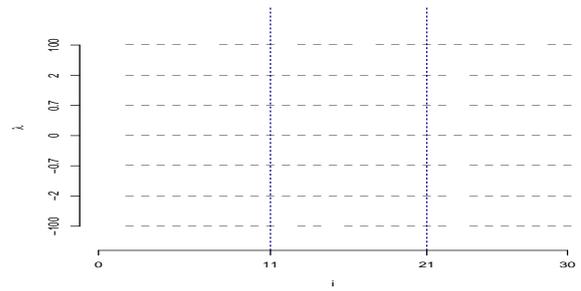


(h)  $P(\rho = \rho')$

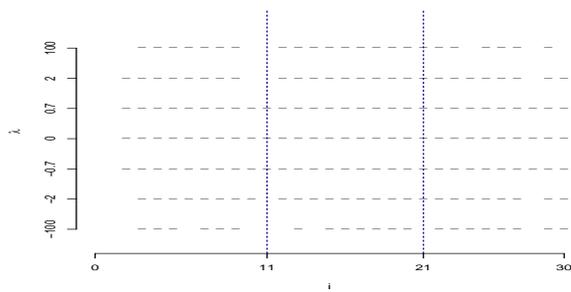
Figura 3.77: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário escada crescente,  $n = 90$ .



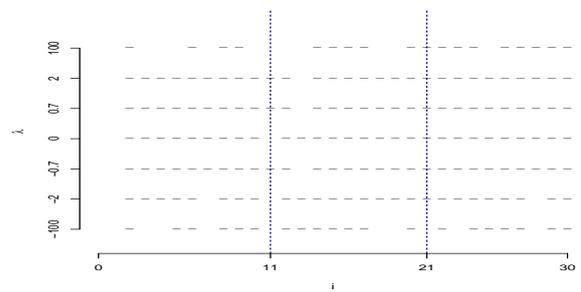
(a) Partição mais provável - Dados  $\lambda = -100$



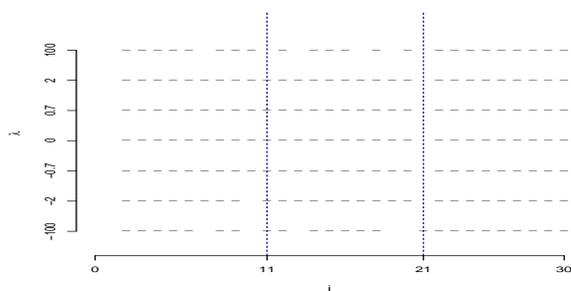
(b) Partição mais provável - Dados  $\lambda = -2$



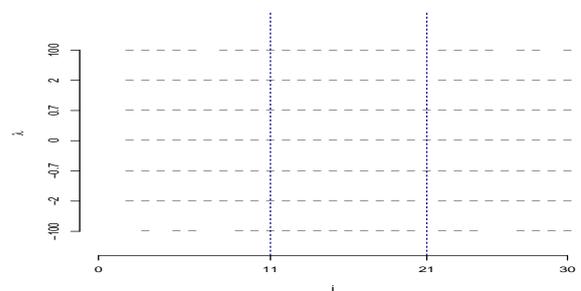
(c) Partição mais provável - Dados  $\lambda = -0.7$



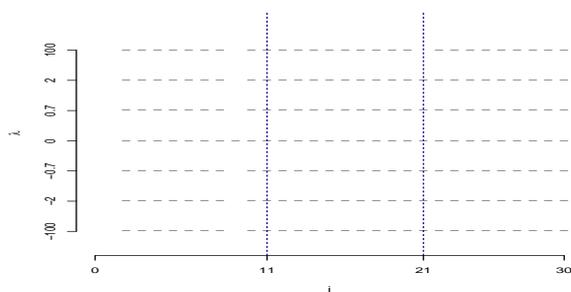
(d) Partição mais provável - Dados  $\lambda = 0$



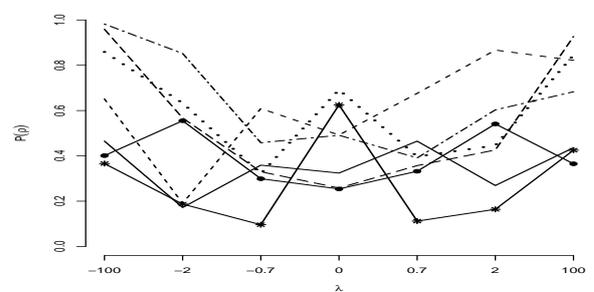
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

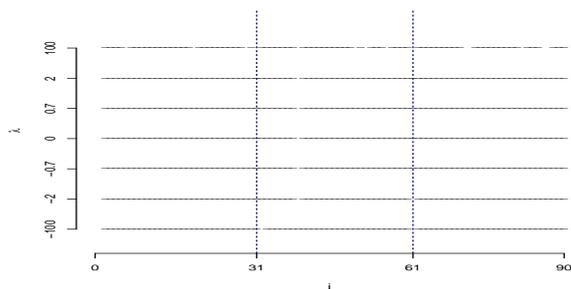


(g) Partição mais provável - Dados  $\lambda = 100$

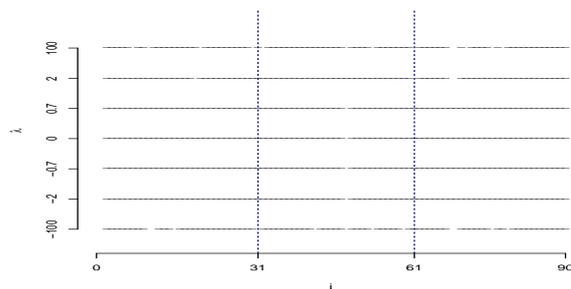


(h)  $P(\rho = \rho')$

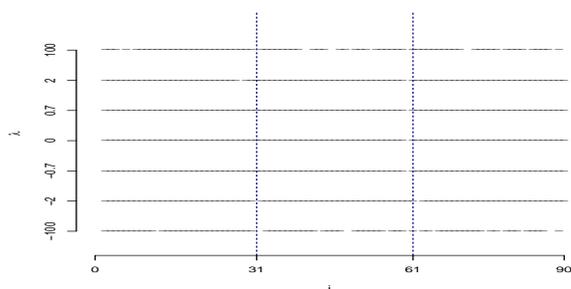
Figura 3.78: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário epsilon,  $n = 30$ .



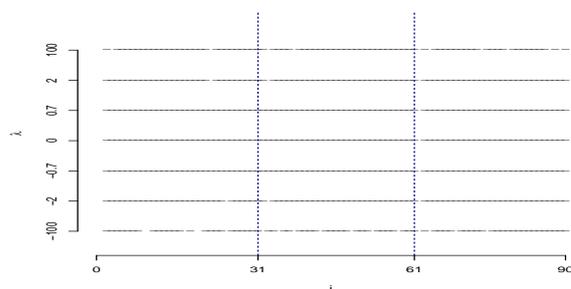
(a) Partição mais provável - Dados  $\lambda = -100$



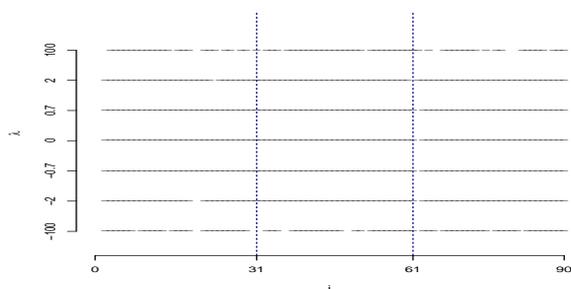
(b) Partição mais provável - Dados  $\lambda = -2$



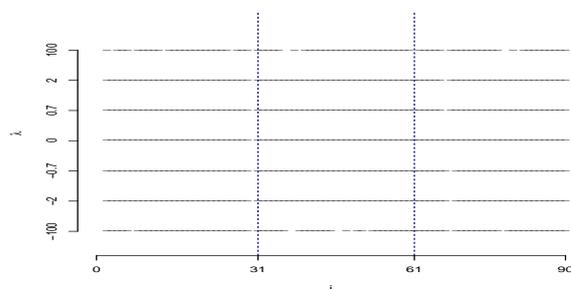
(c) Partição mais provável - Dados  $\lambda = -0.7$



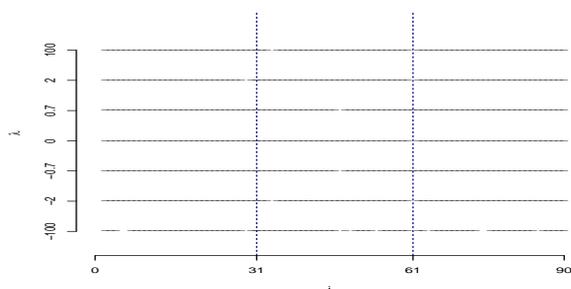
(d) Partição mais provável - Dados  $\lambda = 0$



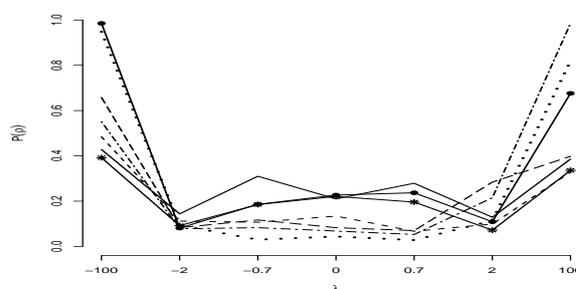
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

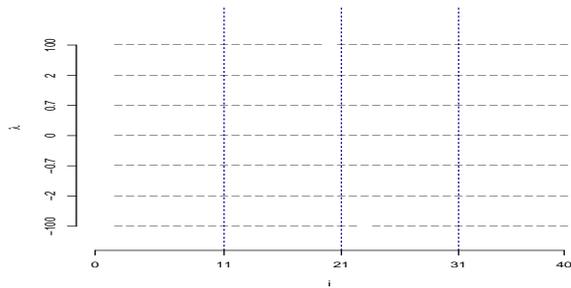


(g) Partição mais provável - Dados  $\lambda = 100$

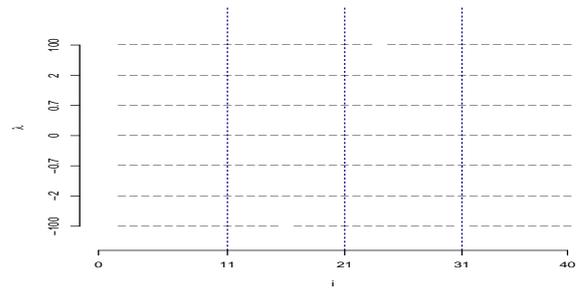


(h)  $P(\rho = \rho')$

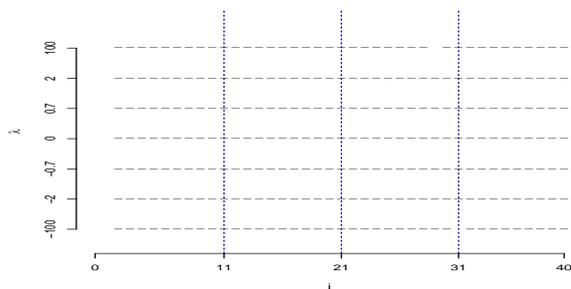
Figura 3.79: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário epsilon,  $n = 90$ .



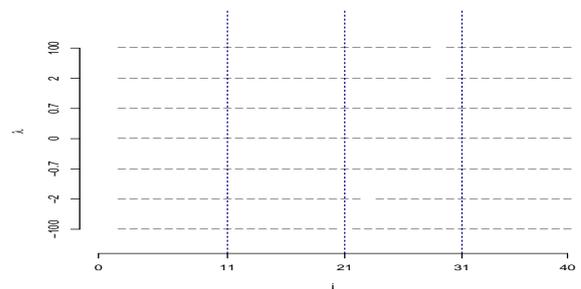
(a) Partição mais provável - Dados  $\lambda = -100$



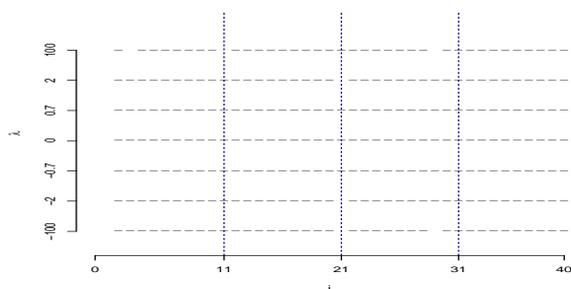
(b) Partição mais provável - Dados  $\lambda = -2$



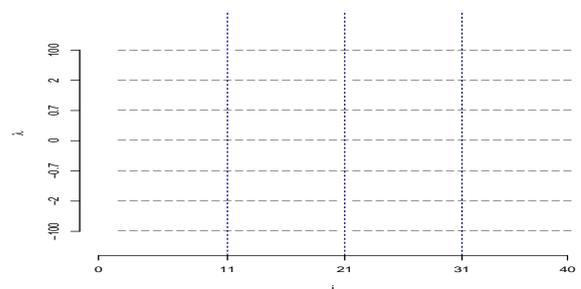
(c) Partição mais provável - Dados  $\lambda = -0.7$



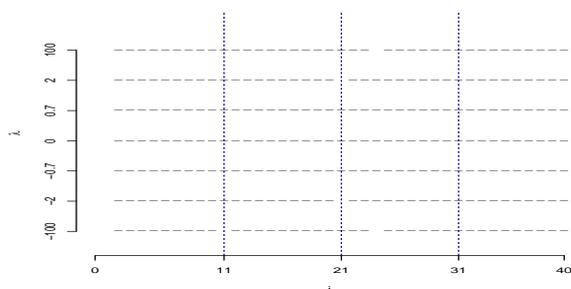
(d) Partição mais provável - Dados  $\lambda = 0$



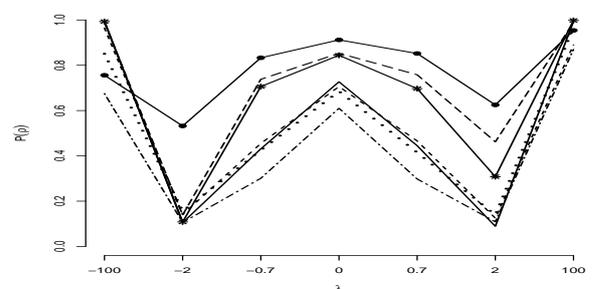
(e) Partição mais provável - Dados  $\lambda = 0.7$



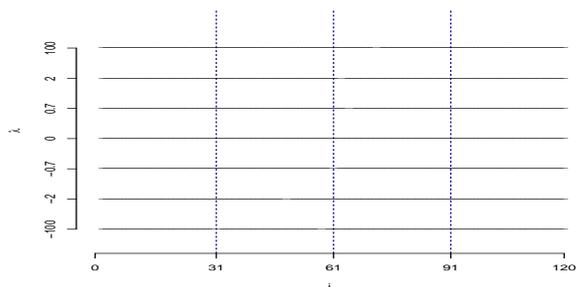
(f) Partição mais provável - Dados  $\lambda = 2$



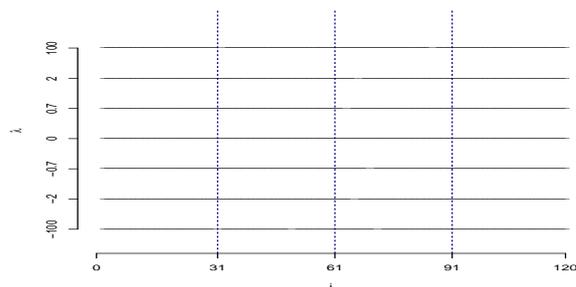
(g) Partição mais provável - Dados  $\lambda = 100$



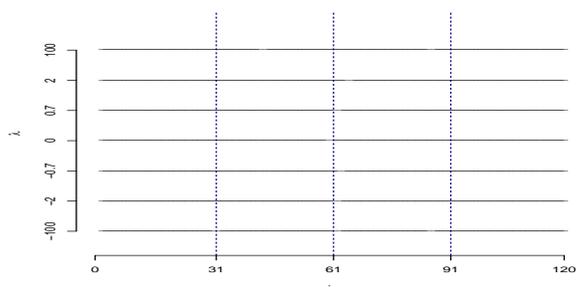
(h)  $P(\rho = \rho')$



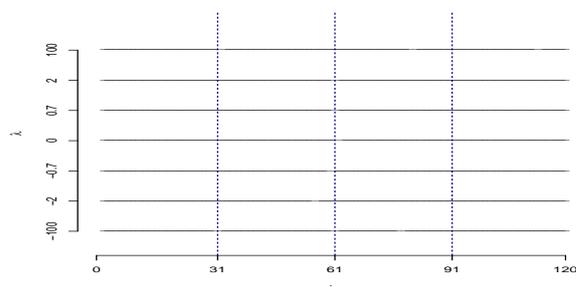
(a) Partição mais provável - Dados  $\lambda = -100$



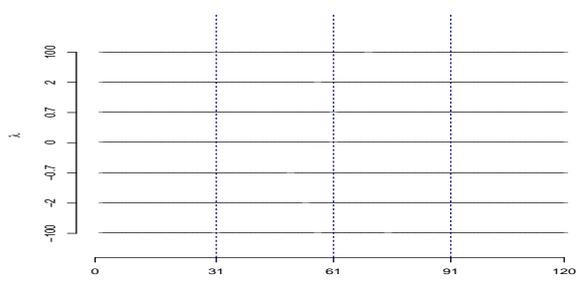
(b) Partição mais provável - Dados  $\lambda = -2$



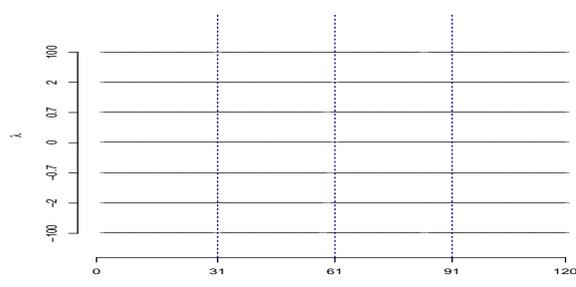
(c) Partição mais provável - Dados  $\lambda = -0.7$



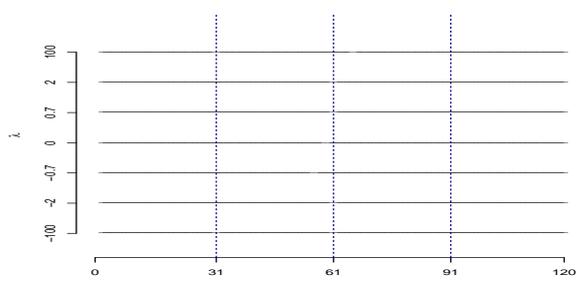
(d) Partição mais provável - Dados  $\lambda = 0$



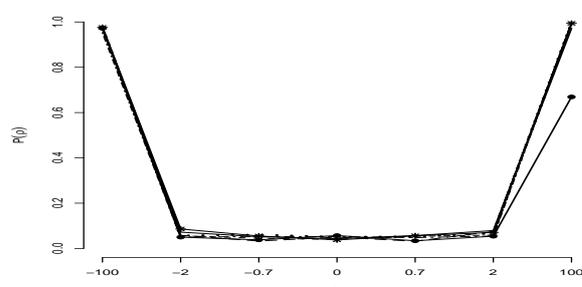
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$



(g) Partição mais provável - Dados  $\lambda = 100$



(h)  $P(\rho = \rho')$

Figura 3.81: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário tendência,  $n = 90$ .

De forma geral, o número de blocos é superestimado se o modelo é gerado com  $\lambda = 0$  e ajustado com  $\lambda = \pm 100$ .

Ao avaliar os resultados para a partição mais provável e o número de blocos, conclui-se que o número de blocos é bem estimado mesmo quando se assume uma assimetria muito forte. Com exceção do cenário escada, epsilon e tendência que tendem a subestimar o número de blocos para dados gerados e ajustados com  $\lambda = \pm 100$ .

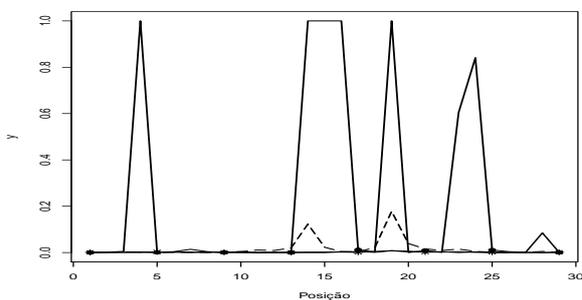
### 3.2.3 A influência de $\lambda$ e $n$ nas inferências sobre $P(A_t)$

As figuras ilustradas nessa seção referem-se a probabilidade de mudança em cada instante. Cada gráfico representa o resultado para os dados gerados com  $\lambda = (-100, -2, -0.7, 0, 0.7, 2, 100)$  e ajustados com todos os valores de  $\lambda$  avaliados nesse estudo de sensibilidade. Desta forma, a Figura 3.82 (a) e (b) mostra a probabilidade de cada instante ser um ponto de mudança quando os dados foram gerados assumindo  $\lambda = 0$  e ajustados com  $\lambda = (-100, -2, -0.7, 0)$  na Figura 3.82(a) e ajustados com  $\lambda = (0, 0.7, 2, 100)$  na Figura 3.82(b).

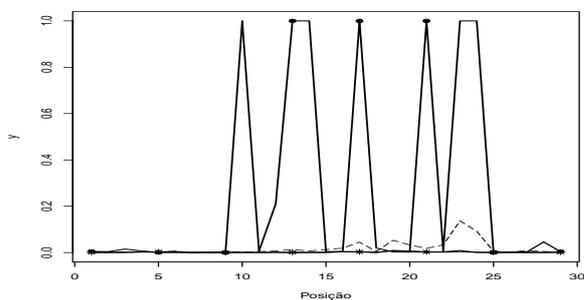
As Figuras 3.82, 3.83 apresentam os resultados para o cenário sem mudança e tamanho de amostra 30 e 90.

Observa-se nas Figuras 3.82(a) e (b) e 3.83(a) e (b) que quando os dados são gerados com  $\lambda = 0$  e ajustados com  $\lambda = \pm 100$  cerca de 5 à 7 pontos de mudança são indicados com alta probabilidade. Quando o tamanho de amostra é maior,  $n = 90$ , cerca de 20 pontos de mudança são estimados para  $\lambda = -100$  e cerca de 30 pontos para  $\lambda = 100$ .

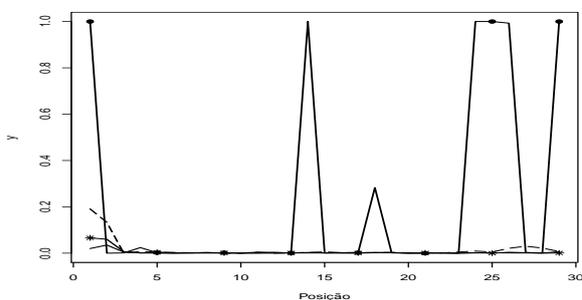
As Figuras 3.82, 3.83 mostram que para todos os valores de  $\lambda$  considerados no estudo de sensibilidade informal, a probabilidade de mudança em cada instante foi alterada indicando que neste cenário sem mudança a escolha equivocada do valor de  $\lambda$  distante do valor gerado influencia nas estimativas da probabilidade de mudança em cada instante.



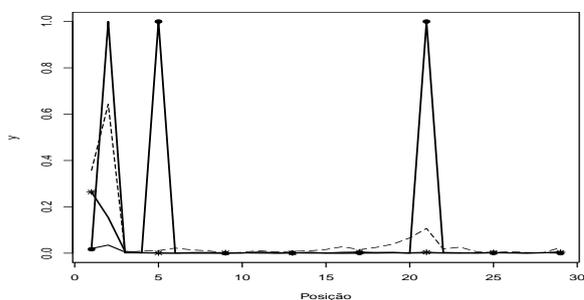
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



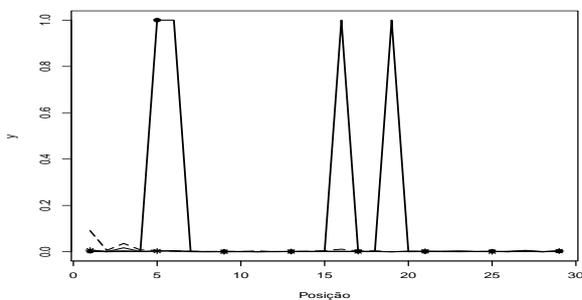
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



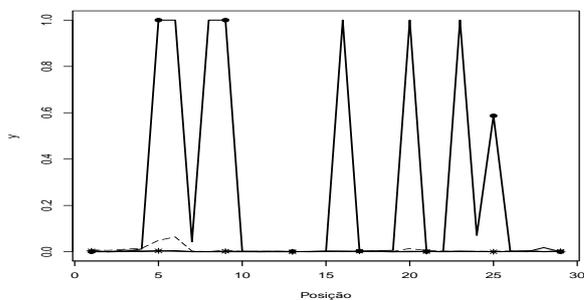
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



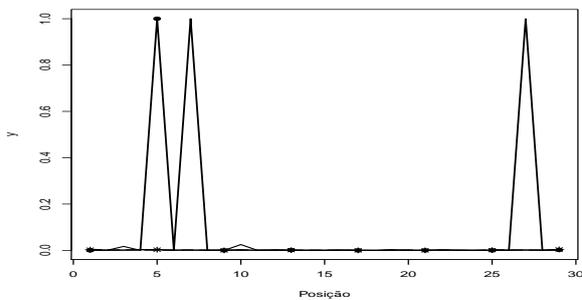
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



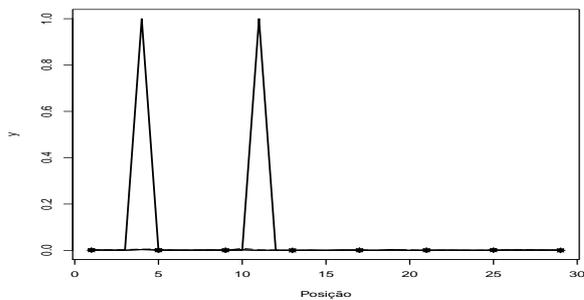
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

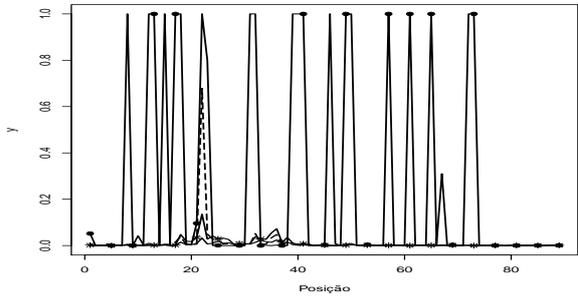


(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

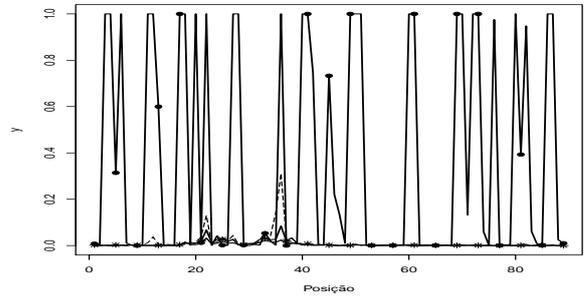


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

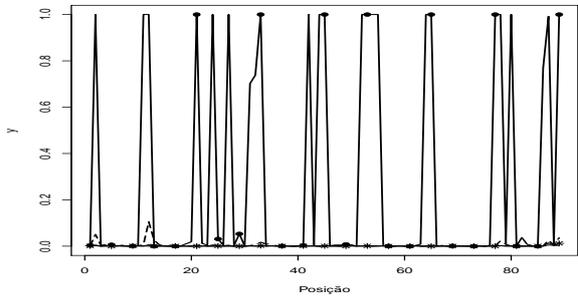
Figura 3.82: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 30$ .



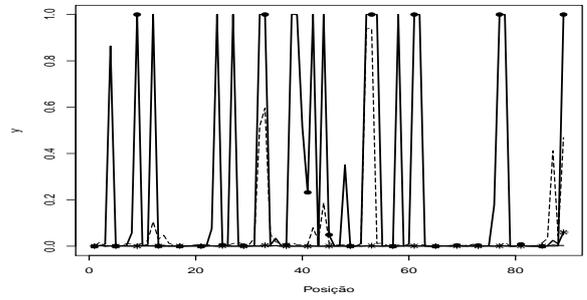
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



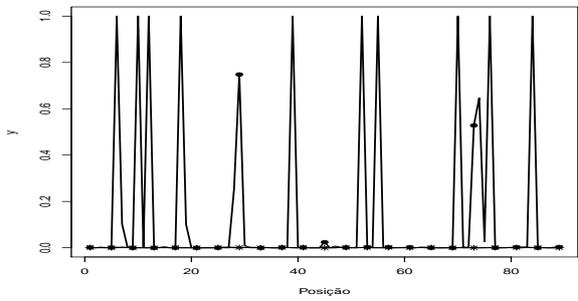
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



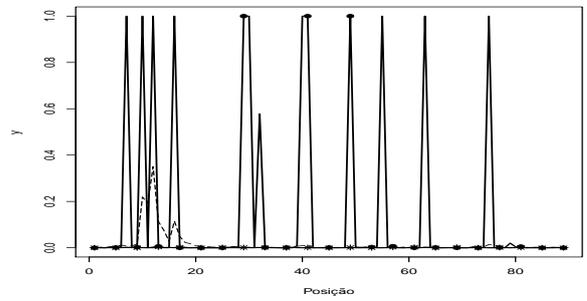
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



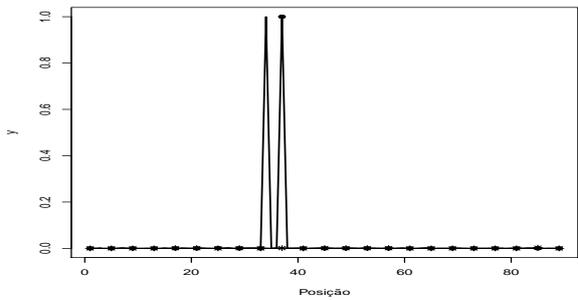
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



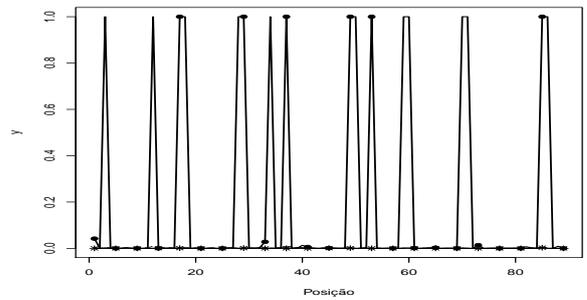
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$



(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



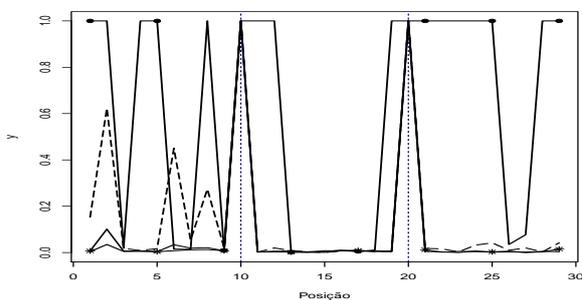
(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

Figura 3.83: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 90$ .

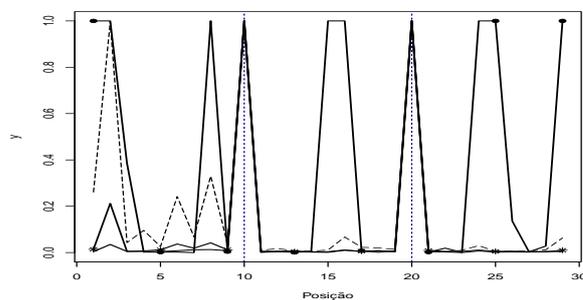
As Figuras 3.84 e 3.85 ilustram os resultados para o cenário degrau não simétrico, nesse cenário o tamanho do salto entre os blocos é maior e a série possui dados negativos e positivos.

A Figura 3.84(h) ilustra que para  $n = 30$ , o ajuste do modelo assumindo  $\lambda = (0, -0.7, -2)$  foi melhor que assumindo  $\lambda = -100$ , mesmo os dados sendo gerados com  $\lambda = -100$ , mas se o tamanho da amostra é de 90, Figuras 3.85(h), e os dados são gerados de  $\lambda = -100$  e ajustados corretamente com  $\lambda = -100$  as estimativas são tão boas quanto assumir  $\lambda = (0, -0.7, -2)$ .

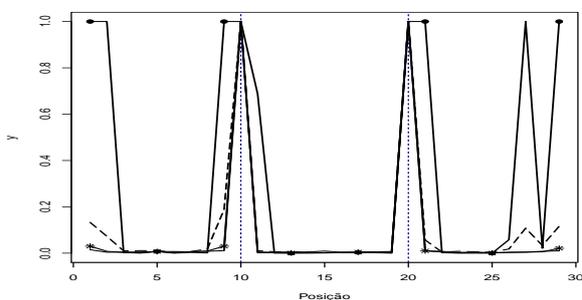
Nota-se, no cenário degrau não simétrico, uma superestimação dos instantes de mudança quando os dados são gerados assumindo  $\lambda = (0, \pm 0.7)$ , e ajustados com  $\lambda = -100$ , como ocorreu no cenários sem mudança.



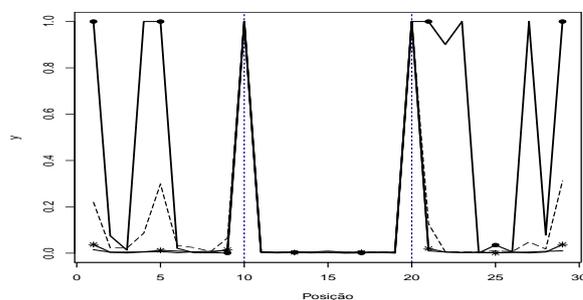
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



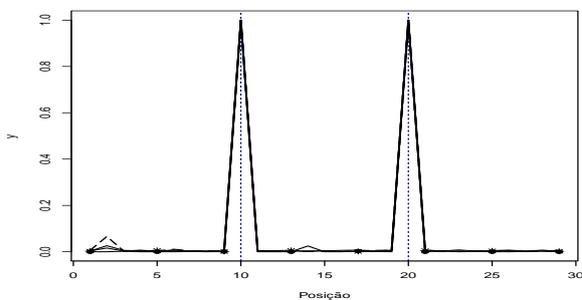
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



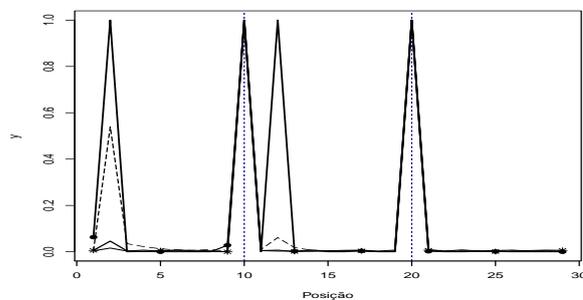
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



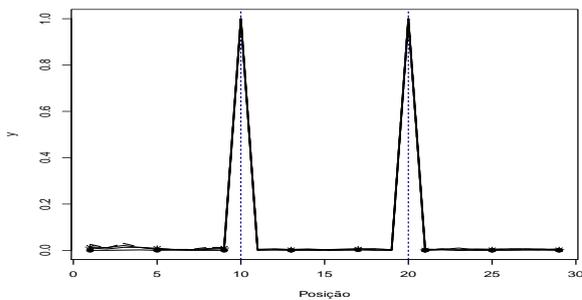
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



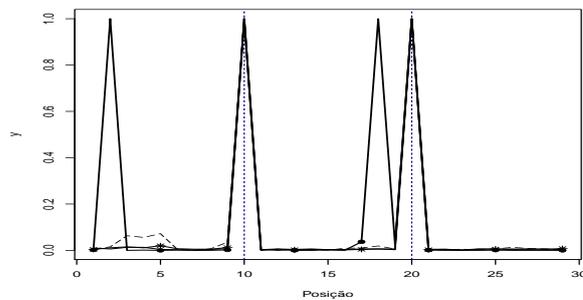
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

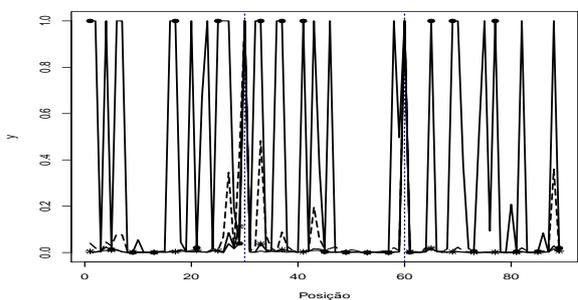


(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

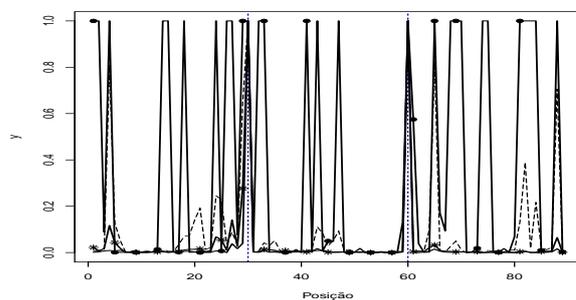


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

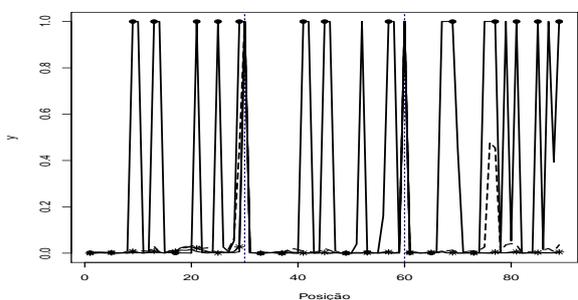
Figura 3.84: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 30$ .



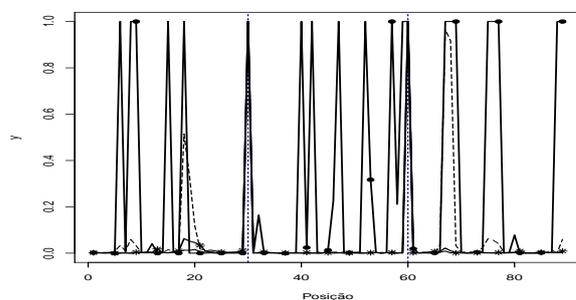
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



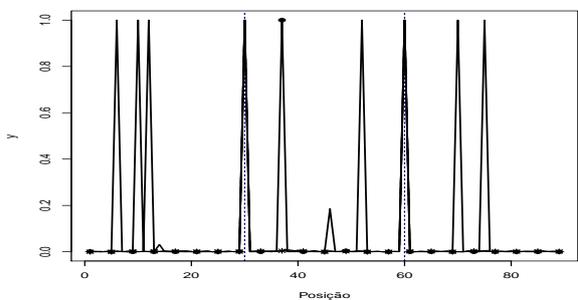
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



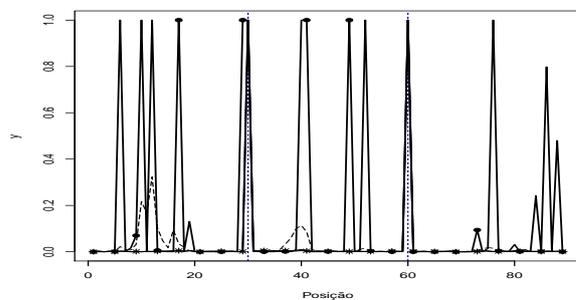
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



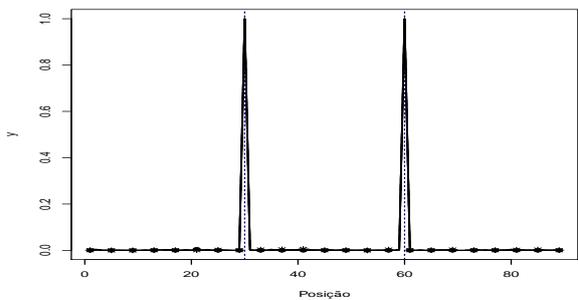
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



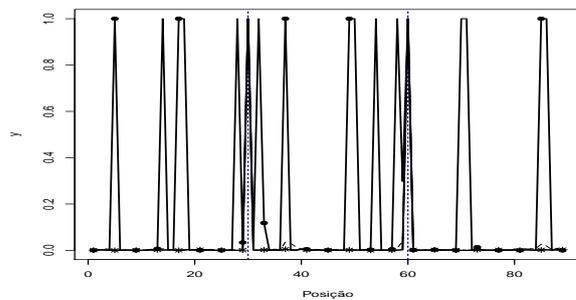
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$



(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



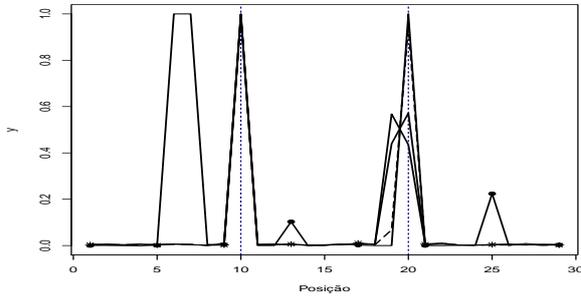
(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

Figura 3.85: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 90$ .

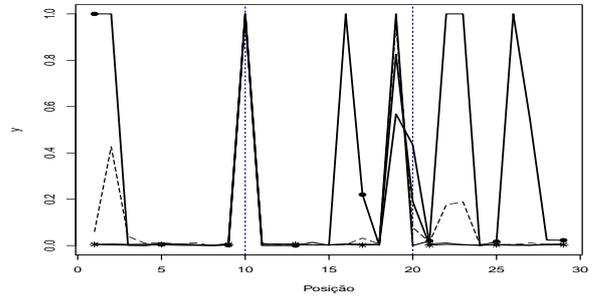
As Figuras 3.86 e 3.87 mostram os resultados para o cenário degrau positivo, sendo que a Figura 3.86 apresentam os resultados para  $n = 30$  e os resultados para  $n = 90$  estão na Figura 3.87.

No cenário degrau positivo, Figura 3.86 (b), nota-se que nenhum modelo identificou corretamente a mudança no instante 20. Além disso, existe uma superestimação dos pontos de mudança nesse caso. Em contra partida, quando os dados foram gerados de  $-100$  e  $-2$ , Figuras 3.86 (a) e (c) e ajustados com  $\lambda \leq 0$  os instantes de mudança foram corretamente identificados. No entanto, o modelo não é assertivo se os mesmos dados gerados de  $\lambda = -100, -2$  são ajustados assumindo  $\lambda = 2, 100$ .

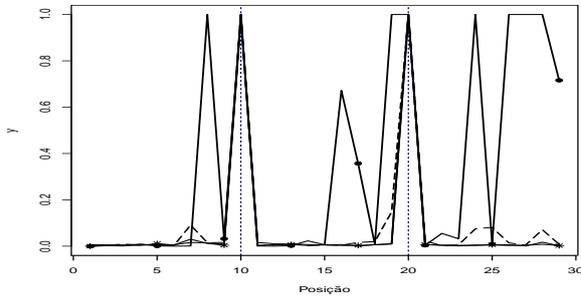
Ao analisar a Figura 3.87 (e) e (f), nota-se que para os dados gerados com  $\lambda = -2$  e ajustados com  $\lambda = 0, \pm 0.7, -2$  os instantes de mudança foram bem estimados, mas ao ajustar  $\lambda = 2$  alguns pontos de mudança são indicados erroneamente.



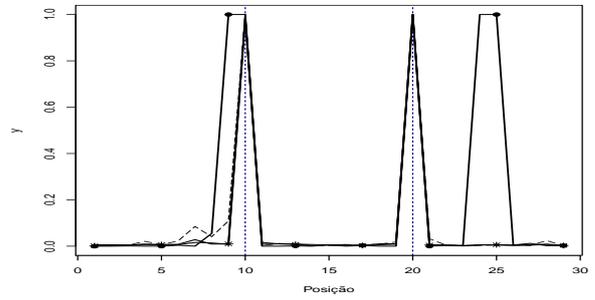
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



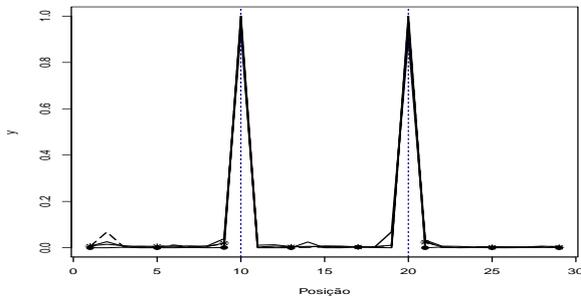
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



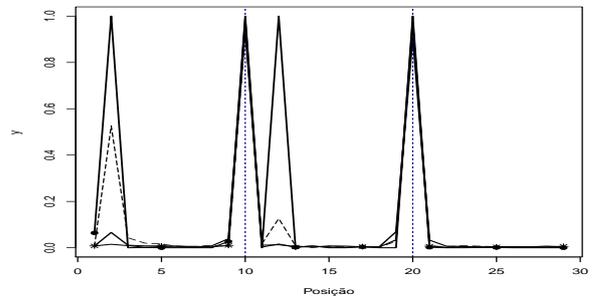
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



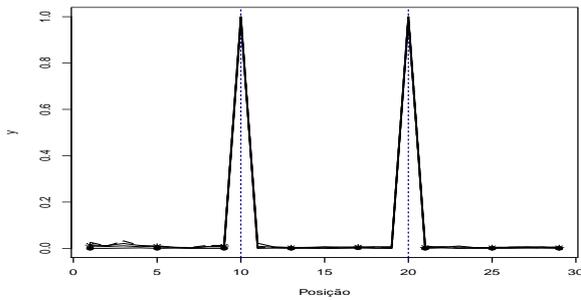
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



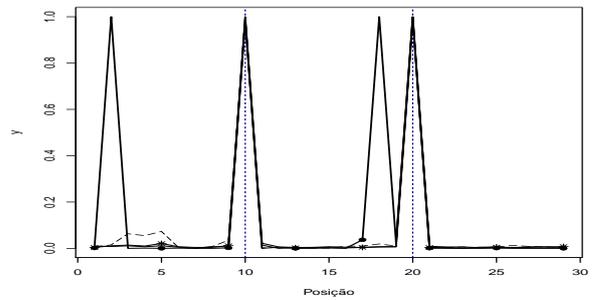
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

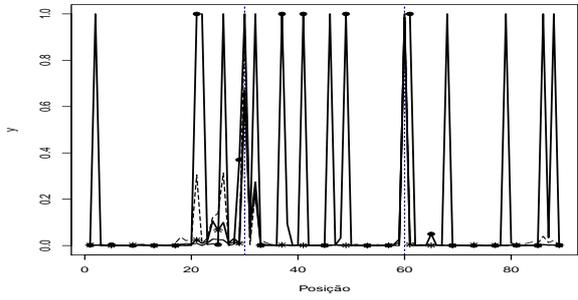


(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

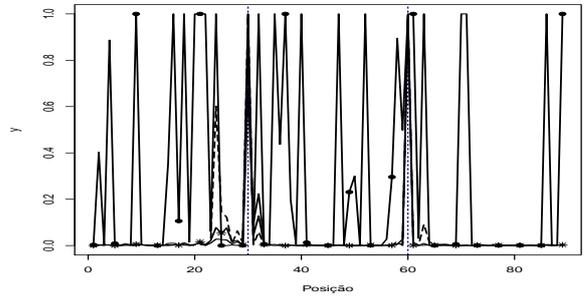


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

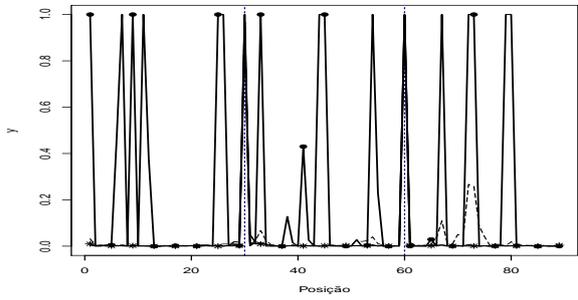
Figura 3.86: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 30$ .



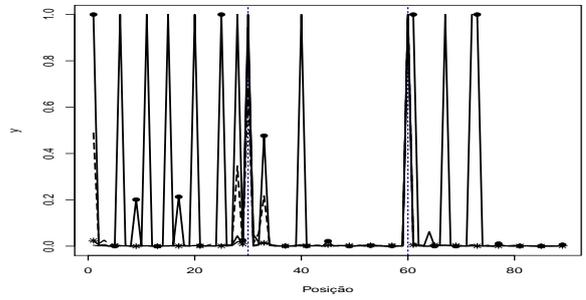
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



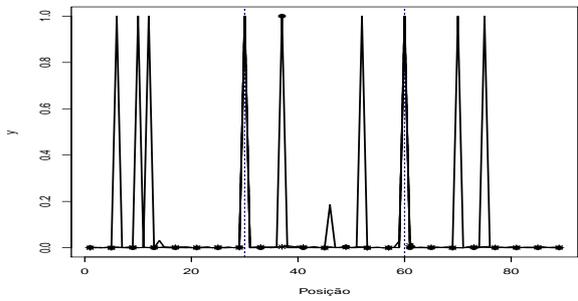
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



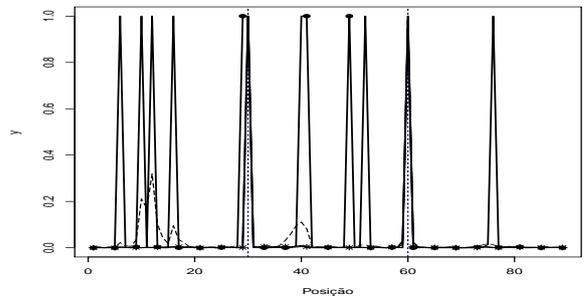
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



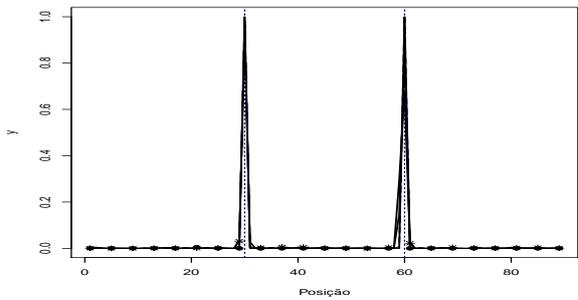
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



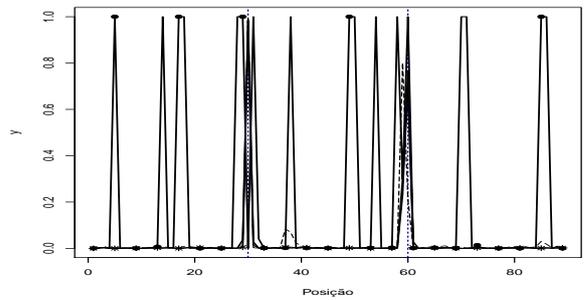
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$



(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

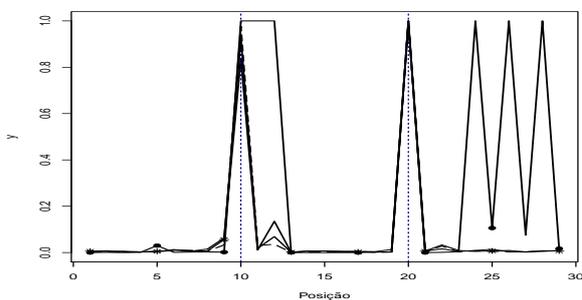
Figura 3.87: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 90$ .

As Figuras 3.88, 3.89 apresentam os resultados para o cenário escada crescente.

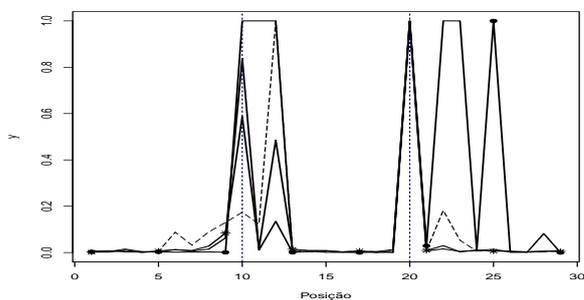
No cenário escada crescente, Figuras 3.88(h) e 3.89(h), observa-se que ao contrário do que aconteceu nos outros cenários, tanto para  $n = 30$  quanto para  $n = 90$  quando os dados são gerados de  $\lambda = -100$  e ajustados com  $\lambda = -100$ , o modelo identifica corretamente os instantes de mudança.

No caso dos dados gerados assumindo  $\lambda = 0, \pm 0.7$  e ajustados com  $\lambda = \pm 2, \pm 100$  nota-se uma superestimação do número de instantes de mudança Figuras 3.88, 3.89 (a), (b), (c) e (d).

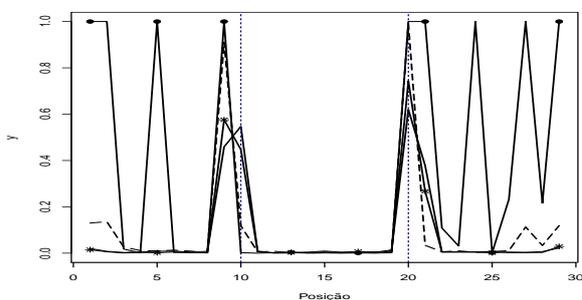
Os resultados para o cenário escada decrescente foram similares aos do cenário escada crescente e podem ser encontrados no Apêndice D.



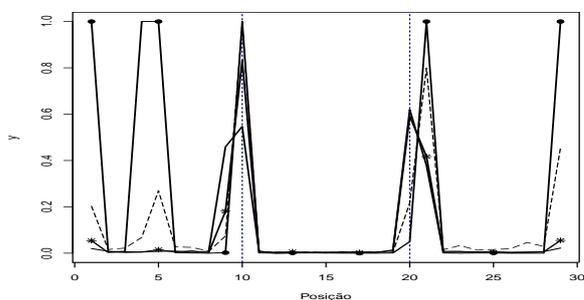
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



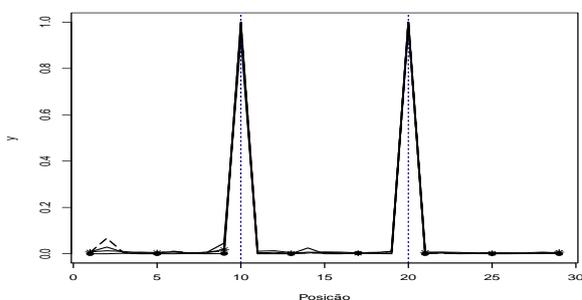
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



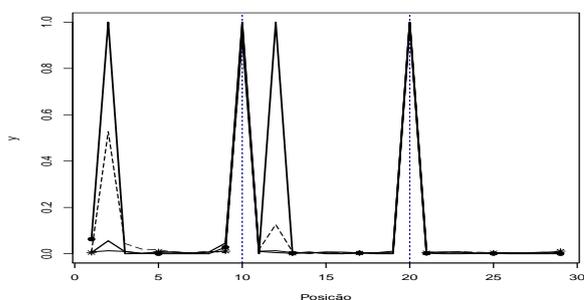
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



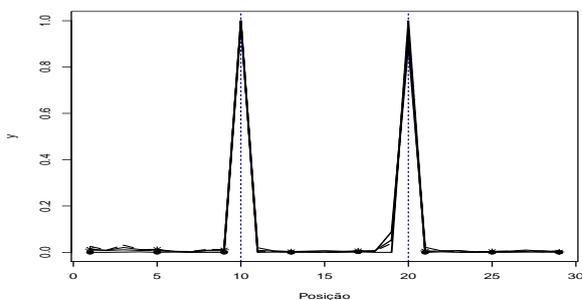
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



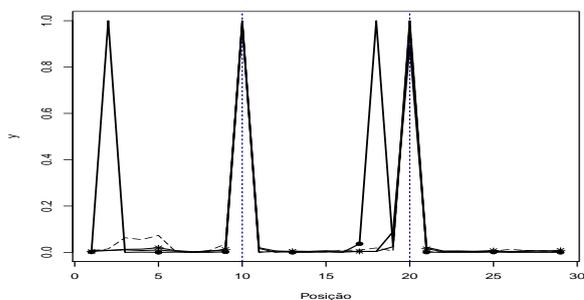
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

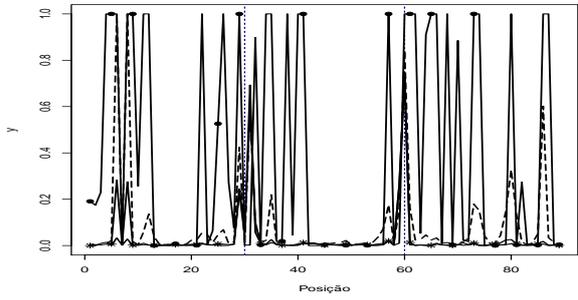


(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

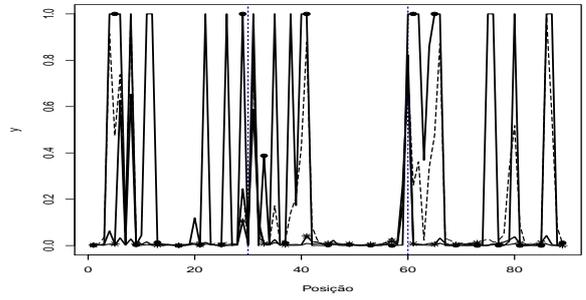


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

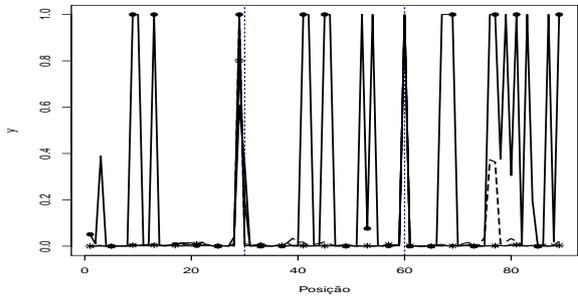
Figura 3.88: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 30$ .



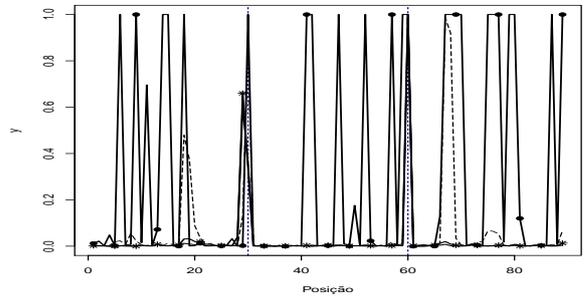
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



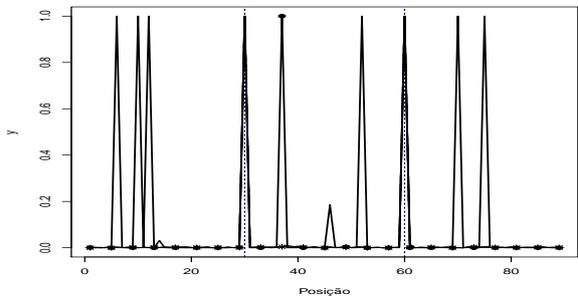
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



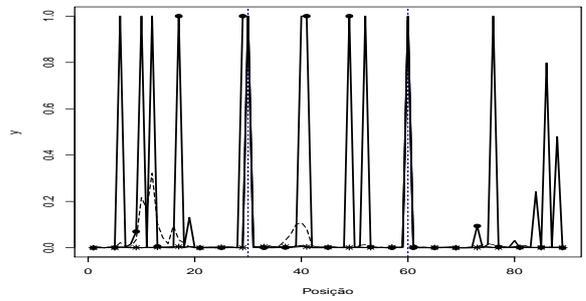
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



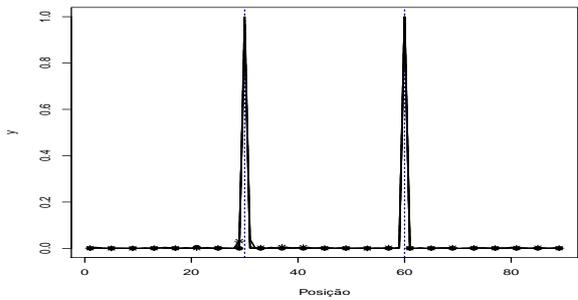
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



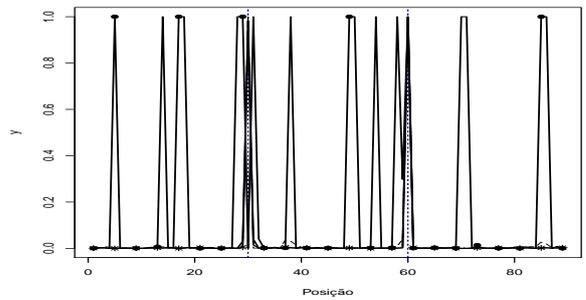
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$



(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



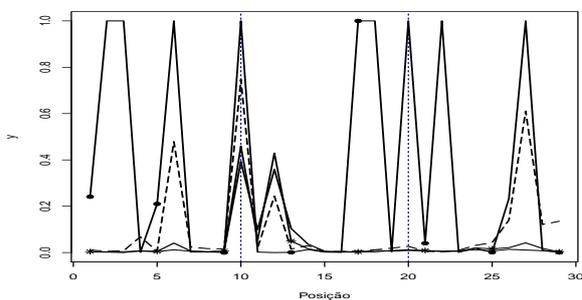
(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

Figura 3.89: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 90$ .

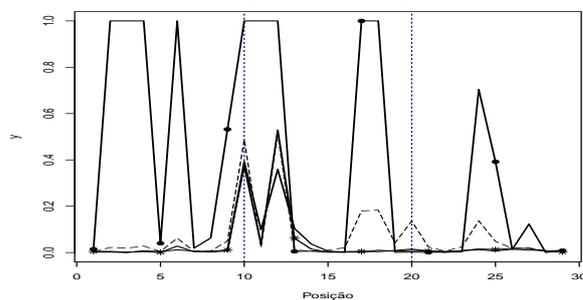
O cenário epsilon é ilustrado nas Figuras 3.90 e 3.91, observa-se nesse cenário uma tendência do modelo indicar mais pontos de mudança que o verdadeiro número de pontos de mudança, que nesse cenário é 2. A indicação de excessivos pontos de mudança diminui para o caso em que o tamanho de amostra é  $n = 90$  e  $\lambda$  é gerado e estimado com  $\lambda = -100$ .

As Figuras 3.92 e 3.93 apresentam os resultados para o cenário tendência, e ao contrário do cenário epsilon que indica excessivos pontos de mudança, o cenário tendência tende a subestimar o número de pontos de mudança independente do valor de  $\lambda$  avaliado se o tamanho da amostra é  $n = 40$ .

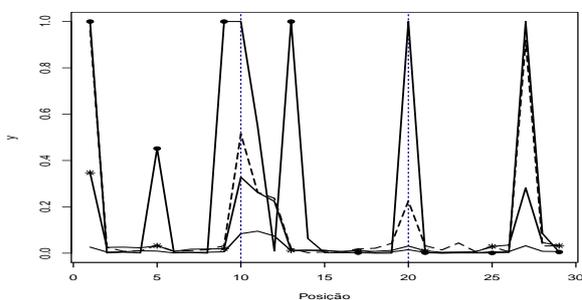
Se o tamanho da amostra aumenta para  $n = 120$  o número de pontos de mudança é melhor estimado, no entanto, a posição indicada pelo modelo do instante da mudança não é assertiva.



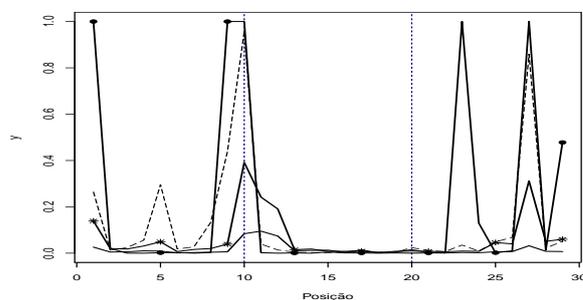
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



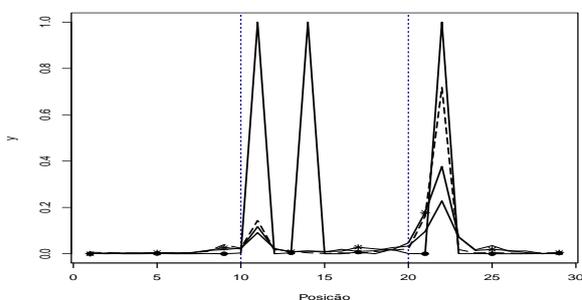
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



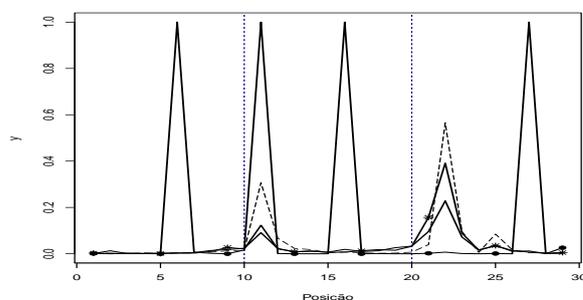
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



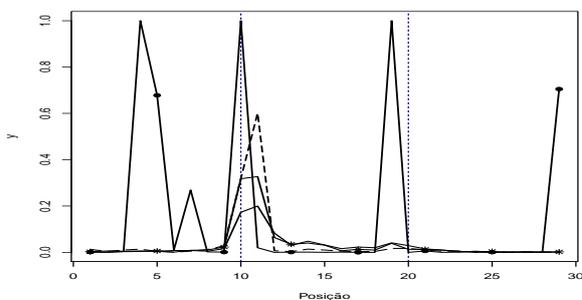
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



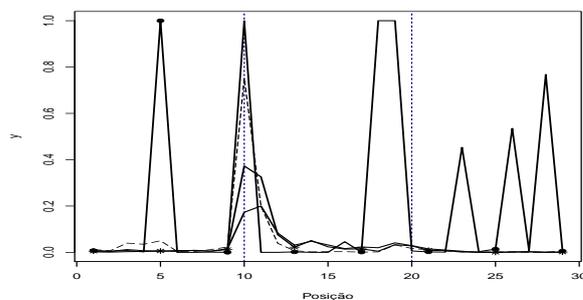
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

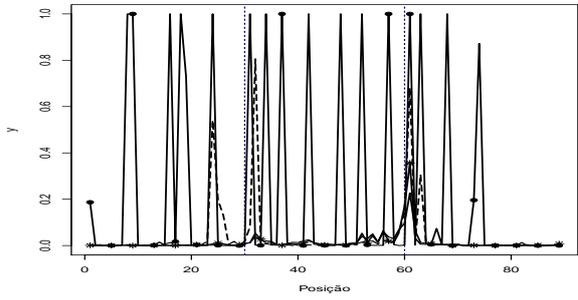


(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

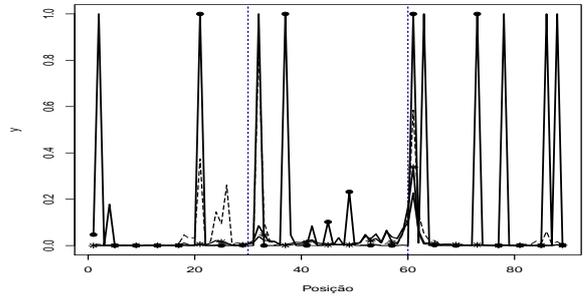


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

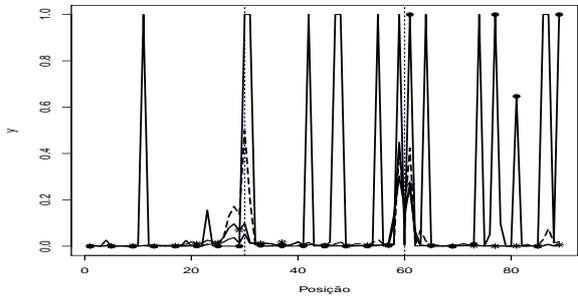
Figura 3.90: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 30$ .



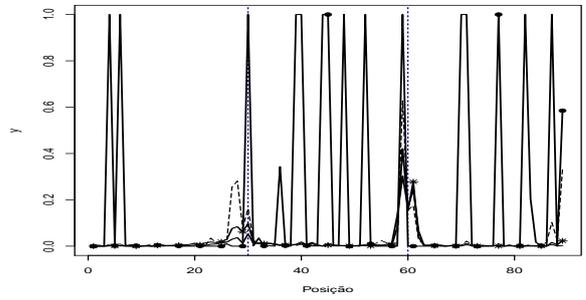
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



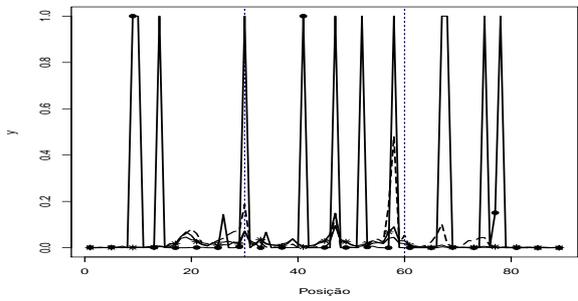
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



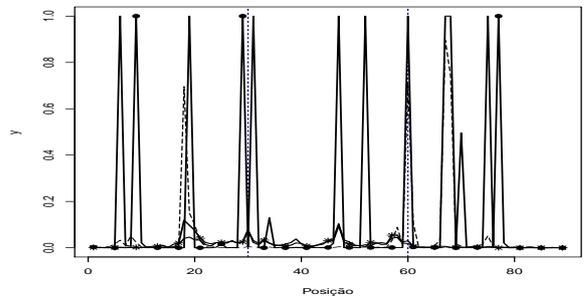
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



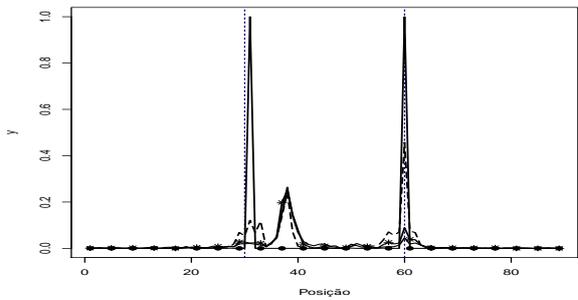
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



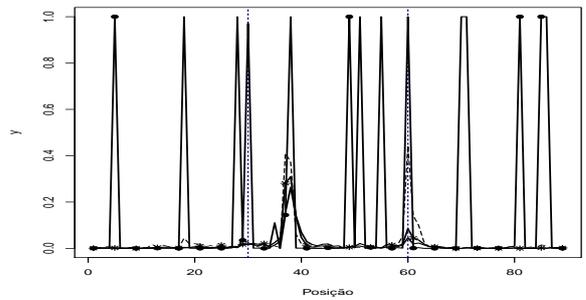
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

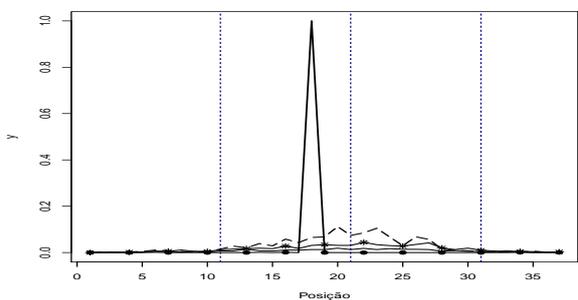


(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

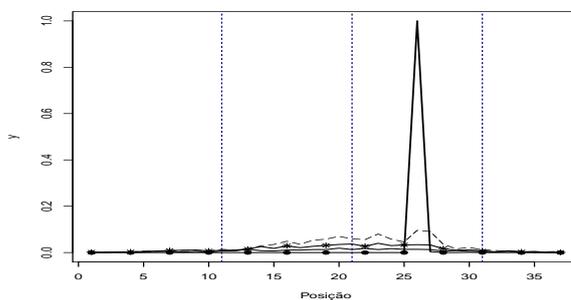


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

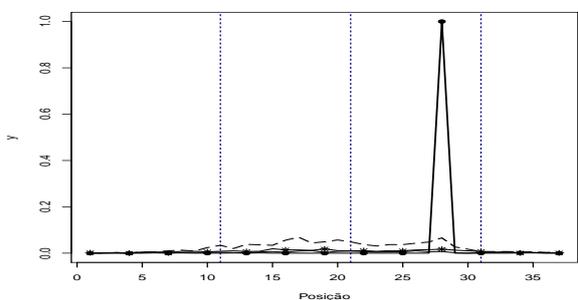
Figura 3.91: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 90$ .



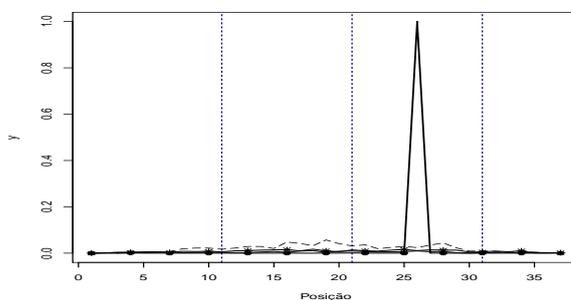
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



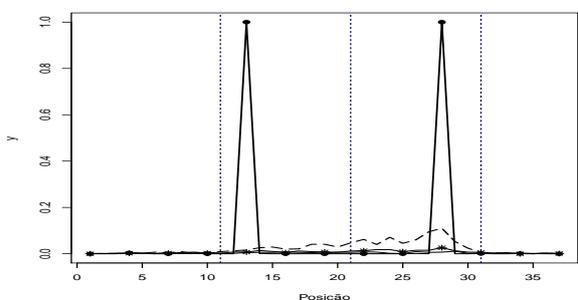
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



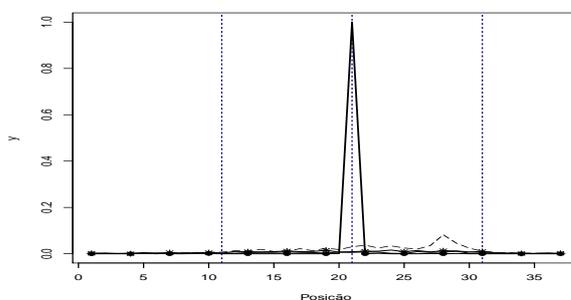
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



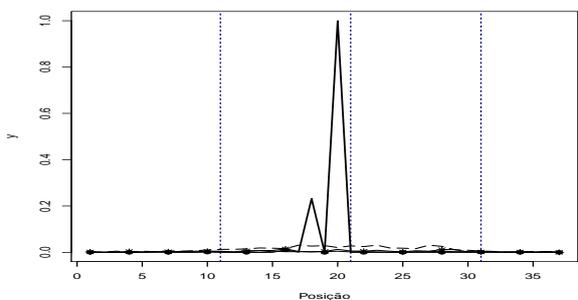
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



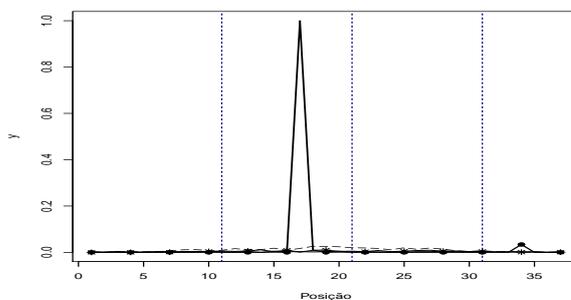
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

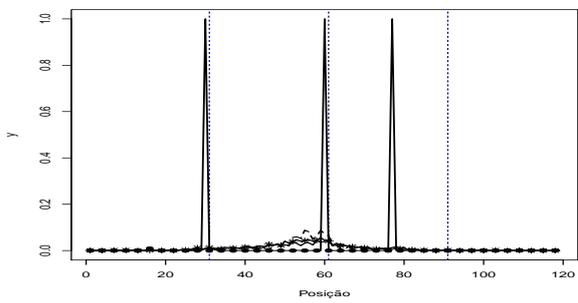


(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

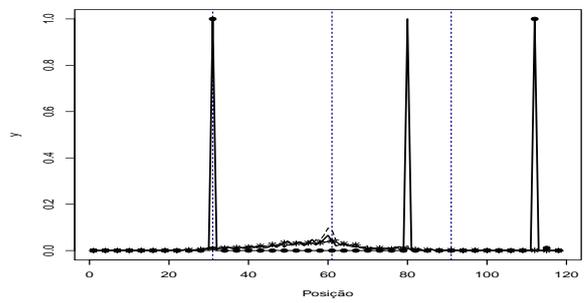


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

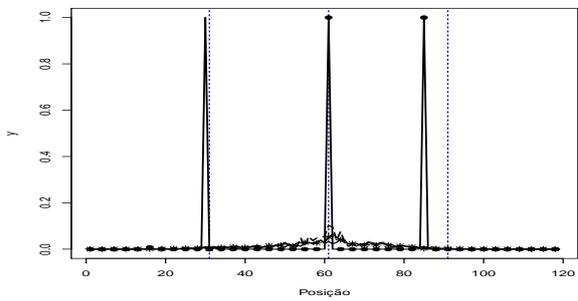
Figura 3.92: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 30$ .



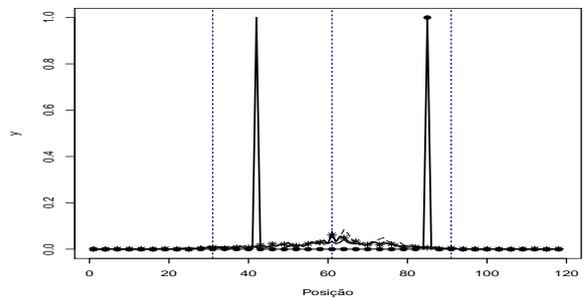
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



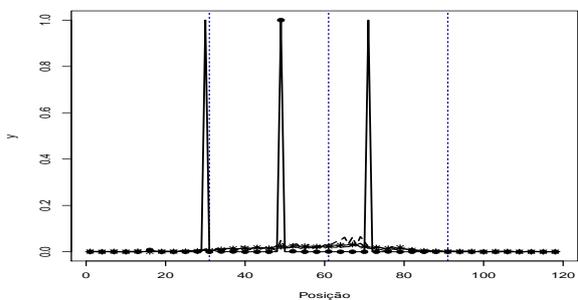
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



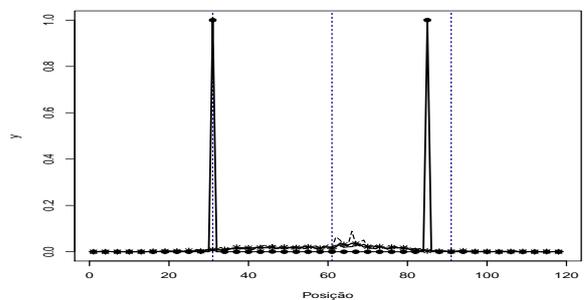
(c) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



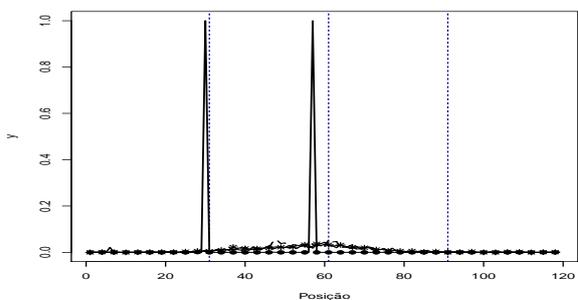
(d) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



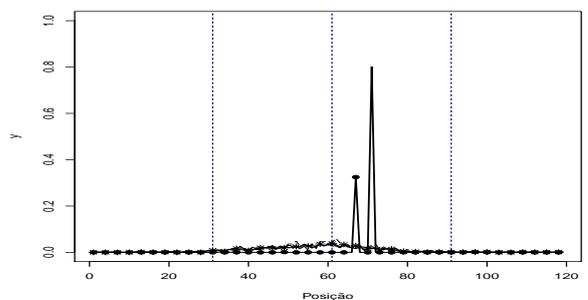
(e) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$



(g) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

Figura 3.93: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 90$ .

De modo geral, a probabilidade de mudança em cada instante apresenta um comportamento similar ao número de blocos e partição mais provável. Além disso, se os dados são gerados assumindo  $\lambda = (0, \pm 0.7)$  e ajustados com  $\lambda = \pm 100$ , em todos os cenários avaliados, a probabilidade de mudança em cada instante indicou com alta probabilidade mais instantes que o correto.

Além disso, na maioria dos casos, mesmo quando os dados foram gerados com parâmetro de assimetria alto, o modelo assumindo  $\lambda = 0$  foi capaz de identificar corretamente os instantes de mudança.

### 3.3 Análise Crítica do Estudo de Sensibilidade

#### 3.3.1 Estudo de sensibilidade - Modelo I

O estudo de sensibilidade contempla os dois modelos discutidos nas Seções 2.4 e 2.5 e em cada um desses modelos o comportamento dos parâmetros do modelo partição produto se comportam de forma distinta, além disso, o tipo de cenário combinado com o grau de assimetria também influencia nas estimativas.

No caso em que assume-se que a distribuição *a priori* é normal assimétrica, no cenário sem mudança existe uma leve subestimação se  $\lambda = -100$  e uma leve superestimação das estimativas produto se  $\lambda = 100$ , no entanto o número de blocos ( $B = 1$ ) é bem estimado por todos os valores de  $\lambda$  avaliados, mas para  $\lambda = \pm 100$  um único bloco é estimado com maior probabilidade em relação aos demais valores de  $\lambda$ , o mesmo comportamento é observado para a partição mais provável.

O comportamento observado para o cenário sem mudança no caso da distribuição *a priori* não se aplica para os demais cenários avaliados, principalmente ao cenário escada, pois nesse caso  $\lambda = \pm 100$  subestima o número de blocos e conseqüentemente produz estimativas produto ruins, mesmo quando o tamanho da amostra aumenta, os demais valores de  $\lambda$  produzem estimativas produto próximas. Em relação ao número de blocos, o modelo indica corretamente ( $B = 3$ ) para os valores de  $\lambda \neq \pm 100$  e nota-se que para o cenário escada decrescente utilizar  $\lambda = -2$  gerou um ganho em torno de 3% em relação a utilizar  $\lambda = 0$  o mesmo ocorre para  $\lambda = 2$  quando o cenário é

escada crescente.

Nos cenários degrais, mesmo com o tamanho do salto maior entre os blocos como no degrau não simétrico, as estimativas produto ficam comprometidas para  $\lambda = \pm 100$  para os demais valores de  $\lambda$  as estimativas são muito próximas para ambos os tamanhos de amostras considerados no estudo, em relação ao número de blocos, neste caso, todos os valores de  $\lambda$  indicam corretamente  $B = 3$  e  $\lambda = 100$  indica ( $B = 3$ ) com maior probabilidade no cenário degrau não simétrico, mas lembrando que as estimativas produto ficam comprometidas para  $\lambda = 100$ . O cenário degrau positivo indica que o número de blocos é estimado corretamente com uma probabilidade maior se  $\lambda = 100$ , mas a partição mais provável é estimada com menor probabilidade quando comparado aos resultados de  $\lambda = 0$ . No cenário negativo quando  $\lambda = -100$  para  $n = 30$  a partição e o número de blocos foram indicados corretamente e praticamente com a mesma probabilidade que  $\lambda = 0$ , mas para  $n = 90$  o valor de  $\lambda = -100$  indica corretamente a partição e o número de blocos com maior probabilidade que  $\lambda = 0$ , no entanto, vale ressaltar que as estimativas produto para  $\lambda = -100$  são insatisfatórias.

No cenário epsilon e tendência para um tamanho de amostra menor as estimativas produto ficam comprometidas para todos os valores de  $\lambda$  avaliados, mas se o tamanho da mostra aumenta as estimativas tendem a melhorar para  $\lambda = \pm 2, \pm 0.7, 0$  no cenário epsilon. No cenário tendência  $\lambda = \pm 2, \pm 0.7, 0$  produz estimativas melhores que  $\lambda = \pm 100$ , mas essas estimativas ainda estão distantes da verdadeira média dos blocos.

O número de blocos no cenário tendência e epsilon tendem a ser subestimados, exceto no caso do cenário epsilon para  $\lambda = \pm 2, \pm 0.7, 0$  e tamanho de amostra 90. Além disso, a probabilidade de indicar corretamente o ponto de mudança é baixa, nesse caso o cenário tendência é o mais afetado e a probabilidade de indicar corretamente o instante da mudança é próxima de zero. Consequentemente, as partições indicadas como mais prováveis são distintas das partições verdadeiras.

Em suma, para o Modelo I, o número de blocos tende a ser melhor estimado com valores de  $\lambda = \pm 100$ , com exceção do cenário escada, epsilon e tendência, onde o número de blocos tende a ser subestimado. Em relação as estimativas produto,  $\lambda = 0, \pm 0.7, \pm 2$  geram estimativas próximas,

mas valores de  $\lambda = \pm 100$  produzem estimativas muito distorcidas, com exceção do cenário sem mudança.

### 3.3.2 Estudo de sensibilidade - Modelo II

O estudo de sensibilidade do Modelo II é um pouco diferente do realizado no Modelo I, pois no caso em que os dados seguem a distribuição normal assimétrica as observações são geradas com um determinado parâmetro de assimetria e ajusta-se o modelo para o conjunto de dados gerados assumindo todos os valores de  $\lambda$  utilizados nesse estudo de sensibilidade.

Desta forma, nota-se que para todos os cenários as estimativas produto ficam comprometidas quando os dados são gerados com um parâmetro de assimetria baixo e ajustados com  $\lambda = \pm 100$ , principalmente quando o tamanho de amostra aumenta, nesses casos as estimativas produto tendem a ser menos suaves contornando a maioria das observações, ao invés de acompanhar a média do bloco.

Os cenários degrais e escada apresentam comportamento similar em relação as estimativas produto, ou seja, se os dados são gerados assumindo  $\lambda = \pm 100$  e ajusta-se o modelo com  $\lambda = 0$  as estimativas produto tendem a ser próximas da média, mas são subestimadas no caso de dados gerados de  $\lambda = -100$  e superestimados no caso em que os dados são gerados assumindo  $\lambda = 100$ . Por outro lado, se os dados são gerados assumindo  $\lambda = 0, \pm 0.7$  e ajusta-se o modelo com  $\lambda = \pm 100$  as estimativas produtos são ruins e conseqüentemente o número de pontos de mudança é superestimado.

Em relação ao número de blocos, no cenário sem mudança ajustando  $\lambda = 0, \pm 0.7$ , independente do valor de  $\lambda$  gerado o modelo indica corretamente um bloco, mas se o tamanho da amostra aumenta, utilizar  $\lambda = 0$  quando os dados possuem assimetria faz com que o número de blocos seja estimado corretamente, mas com uma probabilidade menor do que quando o modelo assume o verdadeiro valor do parâmetro de assimetria.

Ao ajustar o modelo assumindo  $\lambda = \pm 100$ , o número de blocos pode ser subestimado ou superestimado dependendo do cenário, pois os resultados não apresentaram um padrão bem definido para

$\lambda = \pm 100$ . Por exemplo, somente nos cenários degrau não simétrico e  $n = 90$ , degrau positivo e  $n = 30$  e escada crescente e  $n = 30$  ou  $90$  ajustar o modelo com o parâmetro de assimetria correto produziu estimativas corretas para o número de blocos e com alta probabilidade.

Nota-se também que a medida que o tamanho da amostra aumenta, utilizar o valor de  $\lambda$  correto no ajuste do modelo produz estimativas corretas para o número de blocos e com probabilidade mais alta que utilizar  $\lambda = 0$ . Esse comportamento se repete em relação a partição mais provável, pois quando a assimetria é levada em consideração as chances de estimar a partição corretamente aumentam. No entanto, para o cenário escada mesmo os dados sendo gerados e ajustados com  $\lambda = 0$  ou com o parâmetro correto de assimetria a partição mais provável ficou próxima da correta, mas não resultou na mesma partição gerada, mas essa é uma característica do modelo partição produto para o cenário escada.

Nos cenários epsilon e tendência o modelo apresentou maiores distorções em relação as estimativas dos parâmetros do modelo partição produto quando comparadas aos demais cenários. As estimativas produto, número de blocos e partição para esses cenários ficam comprometidas até mesmo para  $\lambda = 0$ , mas esses dois tipos de cenários normalmente são mal estimados pelo modelo partição produto.

No próximo capítulo será feita uma revisão dos conceitos de sensibilidade global e serão discutidas algumas medidas para realizar a análise de robustez no modelo partição produto para a identificação de mudanças nas médias quando a classe de distribuição normal assimétrica é considerada.

---

## Robustez Global no Modelo Partição Produto

---

Em modelos de identificação de pontos de mudança algumas das metas principais são a identificação do número de pontos de mudança e a identificação dos instantes onde tais mudanças ocorrem. O estudo de robustez que se considera neste trabalho inova, pois estuda como perturbações nas distribuições *a priori* das médias populacionais e perturbações na função de verossimilhança afetam as inferências *a posteriori* sobre o número de pontos de mudança ( $B$ ), a partição aleatória ( $\rho$ ), que indica os instantes das mudanças, e a probabilidade de cada instante ser um ponto de mudança ( $P(A_t)$ ). O estudo de robustez em modelos que consideram partições do espaço foi pela primeira vez considerado por Carota e Ruggeri (1994), onde estudam o efeito de contaminações na distribuição *a priori* das partições, mas não no contexto do MPP e nem para uma classe de distribuições contaminada de forma multiplicativa.

Neste capítulo, é discutido o problema de robustez bayesiana global e objetiva-se sugerir medidas para a avaliação da robustez no MPP. Dois contextos são considerados no estudo, a saber, a robustez quando há contaminação na distribuição *a priori* e a robustez quando há contaminação na distribuição amostral. As classes de distribuições consideradas são contaminadas de forma multiplicativa como em Godoi & Branco (2014) as quais foram apresentadas nas Seções 2.4 e 2.5.

A contribuição do presente capítulo é propor medidas para avaliar o efeito da assimetização da distribuição *a priori* para a média dos blocos e o efeito da escolha do modelo em uma classe de distribuições assimétricas com núcleo normal nas estimativas de todos os parâmetros envolvidos no MPP, conseqüentemente, objetiva-se estudar a influência de diferentes parâmetros de contaminação,  $\lambda$ , nas inferências *a posteriori* para o número de blocos, a média que indexa a distribuição das observações, a probabilidade de mudança em cada instante e a partição de interesse induzida pelos pontos de mudança.

## 4.1 Sensibilidade Global

Uma das motivações para se estudar a robustez bayesiana é a dificuldade inerente ao processo de elicitacão de distribuições *a priori* precisas para os parâmetros de interesse em diversas situações práticas. Em virtude dessa dificuldade, é importante avaliar como as estimativas *a posteriori* são afetadas pelas diferentes especificações *a priori*. Na impossibilidade de se fazer isso para todas as especificações *a priori* possíveis, os estudos de robustez se restringem a escolhas de distribuições *a priori* dentro de algumas classes de distribuições.

Entre as classes de distribuições frequentemente utilizadas em estudos de robustez estão as classes de contaminação aditivas e multiplicativas. No presente estudo considerar-se-á apenas classes de distribuições com contaminação multiplicativas, que são propostas no estudo de Van der Linden (2007) e também consideradas por Godoi & Branco (2014). Segundo Van der Linden (2007) uma classe de distribuições contaminada de forma multiplicativa é definida como segue.

**Definição 1.** Seja  $f_0(\theta)$  uma função densidade de probabilidade (f.d.p) basal definida sobre o espaço paramétrico  $\Theta$  e  $\omega(\theta)$  uma função não negativa tal que  $f(\theta) = f_0(\theta)\omega(\theta)$  seja uma função de densidade de probabilidade. A classe de distribuição de probabilidade sobre  $\Theta$  dada por

$$\mathcal{P}_M = \{f(\theta) = f_0(\theta)\omega(\theta) : \omega \in \mathcal{G}\} \quad (4.1)$$

é uma classe de distribuições de probabilidade contaminada multiplicativamente onde  $\mathcal{G}$  é uma

determinada subclasse de funções não negativas.

Godoi & Branco (2014) constatam que a distribuição normal assimétrica proposta por Azzalini (1985) pode ser vista como uma distribuição pertencente à classe de distribuições contaminada definida em (4.1). Como definida em (2.27), uma variável aleatória  $Y$  tem distribuição na família normal assimétrica introduzida por Azzalini (1985), denotada por  $Y \sim \mathcal{SN}(\mu, \sigma_y^2, \lambda)$ , se sua função de densidade de probabilidade é dada por:

$$f(y|\mu, \sigma_y) = \frac{2}{\sigma_y} \phi\left(\frac{y - \mu}{\sigma_y}\right) \Phi\left(\lambda \frac{(y - \mu)}{\sigma_y}\right), \quad (4.2)$$

onde  $\mu \in \mathbb{R}$  é o parâmetro de localização,  $\sigma_y^2 \in \mathbb{R}^+$  é o parâmetro de escala,  $\lambda \in \mathbb{R}$  é o parâmetro de assimetria,  $\phi(\cdot)$  e  $\Phi(\cdot)$  são, respectivamente, a função densidade de probabilidade (f.d.p) e a função de distribuição acumulada (F.D.A) da distribuição normal padrão.

No caso da distribuição normal assimétrica definida em (4.2), a distribuição basal  $f_0(\theta)$  é a f.d.p de uma distribuição normal e a função contaminante  $w(\theta)$  é a F.D.A da mesma distribuição normal multiplicada por 2. Note que se  $\lambda = 0$  tem-se a distribuição *a priori* basal  $f_0(\theta)$ , que neste caso é a distribuição normal. Mais alguns detalhes sobre a família dada em (4.2) podem ser encontrados em Azzalini (1985). Para maiores detalhes sobre desenvolvimentos nesta área de distribuições assimétricas ver Arnold & Beaver (2002), Genton (2004) e Azzalini (2005), por exemplo.

A modelagem Bayesiana, no entanto, segue a postura ou raciocínio subjetivista. Dentro desta ótica, tanto a distribuição *a priori* quanto a função de verossimilhança são consequências de julgamentos feitos sobre as quantidades observáveis. Sendo assim, a dificuldade inerente à eliciação da distribuição *a priori* também se faz presente na escolha da função de verossimilhança. Se a função de verossimilhança utilizada para descrever o comportamento dos dados é escolhida na classe contaminada dada em (4.2) pode-se avaliar o quanto as inferências *a posteriori* são afetadas ou perturbadas quando se assume modelos que se afastam da normalidade. Este tipo de análise é chamado de análise de robustez no modelo e as mesmas medidas usadas para fazer a avaliação da robustez global para perturbações *a priori* podem ser utilizadas para perturbações na distribuição amostral. Este tipo de estudo de robustez também será abordado neste trabalho.

Os principais tipos de robustez bayesiana são a robustez global e a robustez local as quais são discutidas, por exemplo, em Ruggeri et al. (2005), Berger et al. (1994) entre outros. Berger et al. (2000) afirmam que medidas de robustez local, em algumas situações complicadas, podem ser obtidas mais facilmente mas as interpretações nem sempre são claras. Por esse motivo, nesse estudo só será abordada a robustez global através do estudo de sensibilidade global, cuja definição é dada a seguir.

**Definição 2.** Seja  $\mathcal{P}$  uma classe de distribuições de probabilidade  $p(\theta)$  definida no espaço paramétrico  $\Theta$  do parâmetro de interesse  $\theta$ . Define-se como uma medida de sensibilidade global  $S_g$  a amplitude de variação do funcional  $T(p(\theta))$ , isto é

$$S_g(p(\theta)) = \sup_{p_\theta \in \mathcal{P}} T(p(\theta)) - \inf_{p_\theta \in \mathcal{P}} T(p(\theta)). \quad (4.3)$$

As inferências *a posteriori* são ditas robustas em relação a variações na classe de distribuições de probabilidade *a priori*  $\mathcal{P}$  se a medida de sensibilidade global  $S_g(\cdot)$  dada em (4.3) for pequena.

Em um estudo de sensibilidade global, algumas opções usualmente considerados para o funcional  $T$  são a média e a variância *a posteriori*. Se o funcional de interesse é a diferença entre a média da distribuição *a posteriori* contaminada  $\pi_\lambda(\mu|\mathbf{y})$  e a média *a posteriori* baseada na distribuição basal obtida assumindo-se que a distribuição *a priori* para  $\mu$  pertencente à classe de distribuição multiplicativa dada em (4.2), a medida de sensibilidade global é dada por

$$S_g(\pi_\lambda(\mu|\mathbf{y})) = \sup_{\mathcal{P}_M} |E_\lambda(\mu|\mathbf{y}) - E_0(\mu|\mathbf{y})| - \inf_{\mathcal{P}_M} |E_\lambda(\mu|\mathbf{y}) - E_0(\mu|\mathbf{y})|, \quad (4.4)$$

onde  $E_\lambda(\cdot)$  denota a média da distribuição *a posteriori* quando, *a priori*,  $\lambda \neq 0$  é especificado e  $E_0$  denota a média *a posteriori* quando a distribuição *a priori* basal para  $\theta$  é a distribuição  $\mathcal{N}(\mu, \sigma_\mu^2)$ , ou seja,  $E_\lambda(\cdot)$  denota a esperança *a posteriori* com respeito a classe contaminada que, neste caso, é a família de distribuição normal assimétrica. Essa medida foi construída, pois as estimativas produto são pautadas nas médias *a posteriori*.

Para o modelo partição produto avalia-se o comportamento desta medida ao longo do tempo, isto é, para todo  $\mu_i, i = 1, \dots, n$ . Este tipo de medida permitirá avaliar também se há modificações significativas nas inferências com perturbações em  $\lambda$  nas proximidades dos pontos de mudança para os modelos apresentados nas Seções 2.4 e 2.5,  $E_\lambda(\mu_i | \mathbf{y})$  é dada em (2.25) e (2.32) quando considera-se a contaminação na distribuição *a priori* e na função de verossimilhança, respectivamente. Já o valor esperado sob a classe normal, isto é,  $E_0(\mu_i | \mathbf{y})$  é dado em (2.18) na Seção 2.3. Se o objetivo é estudar a influência na variabilidade, a avaliação da robustez é feita similarmente e o funcional de interesse é  $|V_\lambda(\mu_i | \mathbf{y}) - V_0(\mu_i | \mathbf{y})|$ . No caso do MPP não se tem expressões fechadas para  $V_\lambda(\mu_i | \mathbf{y})$  e  $V_0(\mu_i | \mathbf{y})$  o que se têm são boas aproximações para tais variâncias.

Muitas vezes, não é possível calcular de forma analítica os limites em (4.3), como é o caso do MPP apresentado no Capítulo 2. Uma forma alternativa para encontrá-los consiste em avaliar numericamente o comportamento da quantidade  $S_g(\pi_\lambda(\mu|\mathbf{y}))$  como mostrar-se-á na Seção 4.2.

Uma outra maneira de se estudar robustez é através da função de concentração, como apresentado por Cifarelli & Regazzini (1987) e essa função é uma generalização da curva de Lorenz. Mas como no modelo partição produto existem  $n$  médias, o estudo de robustez considerando a função de concentração para  $\mu_i$  seria exaustivo, pois seria necessário uma função de concentração para cada  $\mu_i$ .

#### 4.1.1 Análise de Sensibilidade na probabilidade de mudança, na partição e no parâmetro $B$

Identificar corretamente o número de blocos e o instante das mudanças é de suma importância no modelo partição produto, por isso, nesta seção apresenta-se medidas para avaliar a sensibilidade no MPP em relação a probabilidade de mudança em cada instante  $P(A_l)$ , o número de blocos  $B$  e a partição de interesse  $\rho$ .

Obter-se estimativas *a posteriori* para o número blocos  $B$  por meio do vetor de  $U'$ s definido no Capítulo 2 é relativamente simples. Uma possível medida de perturbação na inferência sobre  $B$  pode ser obtida considerando como funcional em (4.3) a moda *a posteriori*. No entanto, as distâncias

entre as modas *a posteriori* não retrata da melhor forma a diferença da inferência sobre o número de blocos para cada  $\lambda$ . Por exemplo, se  $\lambda = 5$  retorna que o número de blocos mais provável é 3 com probabilidade 0.8 e  $\lambda = 0$  indica que o número de blocos é 2 com probabilidade 0.78, avaliar somente a diferença entre as probabilidades *a posteriori* do número de blocos mais prováveis levaria a diferença ( $|0.8 - 0.78| = 0.02$ ). Note que esta medida fornece uma evidência sobre a importância relativa das modas indicadas pelos modelos, mas não sobre a diferença das modas dificultando avaliar se há robustez ou não. O mesmo ocorreria se fosse feita a diferença entre as modas *a posteriori*. A diferença entre 9 e 10 seria apenas 1, mas se o verdadeiro número de blocos for 3 essa diferença entre as modas não retrataria a realidade no que diz respeito à qualidade sobre a inferência acerca do número de blocos detectados pelo modelo.

Desta forma, como medida para estudar a robustez do número de blocos propõem-se utilizar a maior distância entre as funções de distribuições acumuladas para o número de blocos. Assim, para cada valor  $b$  de  $B$ , determina-se a função de distribuição acumulada avaliada em cada um dos modelos definidos por  $\lambda$ , para cada  $b$  calcula-se a maior diferença entre a função de distribuição acumulada do modelo contaminado e do modelo basal da seguinte forma

$$S_g(\pi_\lambda(B|\mathbf{y})) = \text{Sup}_{\mathcal{P}_M} | F_\lambda(b|\mathbf{y}) - F_0(b|\mathbf{y}) | - \text{Inf}_{\mathcal{P}_M} | F_\lambda(b|\mathbf{y}) - F_0(b|\mathbf{y}) | . \quad (4.5)$$

onde  $F_\lambda(b | \mathbf{y})$  e  $F_0(b | \mathbf{y})$  denotam a função de distribuição acumulada para  $B$  avaliada em  $b$  quando o parâmetro de assimetria assumido é  $\lambda$  e zero, respectivamente. Considera-se que a inferência sobre o número de blocos é robusta se  $S_g(\pi_\lambda(B | \mathbf{y}))$  é próximo de zero para todo  $b$ , como nesse caso a probabilidade varia de 0 a 1, espera-se que próximo de zero seja uma diferença inferior a 0.05. Como visto na equação (2.10) no MPP a distribuição *a posteriori* para  $B$  depende das distribuições preditivas por bloco, logo o  $\text{Inf}_{\mathcal{P}_M} | F_\lambda(b|\mathbf{y}) - F_0(b|\mathbf{y}) | = 0$ , assim a medida definida em (4.5) pode ser reescrita como  $S_g(\pi_\lambda(B|\mathbf{y})) = \text{Sup}_{\mathcal{P}_M} | F_\lambda(b|\mathbf{y}) - F_0(b|\mathbf{y}) |$ .

Para se avaliar a robustez na probabilidade *a posteriori* de mudança em cada instante propõem-se considerar o maior valor da maior distância de  $P(A_l)$  sob o modelo que assume  $\lambda \neq 0$  em relação ao modelo que assume  $\lambda = 0$ , isto é, considera-se a seguinte medida

$$S_g(\pi_\lambda(P(A_l)|\mathbf{y})) = \text{Sup}_{\mathcal{P}_M} | P_\lambda[A_l|\mathbf{y}] - P_0[A_l|\mathbf{y}] | - \text{Inf}_{\mathcal{P}_M} | P_\lambda[A_l|\mathbf{y}] - P_0[A_l|\mathbf{y}] |. \quad (4.6)$$

A medida em (4.6) foi proposta, pois usualmente a probabilidade de mudança em cada instante é dado por  $P(A_l)$ , no entanto a medida  $P(A_l)$  não permite avaliar se o comportamento de  $P(A_l | \mathbf{y})$  sofre mais influencia a medida que  $\lambda$  se afasta de zero ou não, mas é possível verificar se as distâncias são grandes ou não. E nesse caso, como trata-se também de uma medida de probabilidade, espera-se que variações pequenas sejam inferiores a 0.05.

Se as medidas definidas para avaliar a robustez no número de blocos e na probabilidade de mudança em cada instante retornam valores próximos de zero, então isso significa que a escolha de distribuições pertencentes a classe contaminada  $\mathcal{P}_M$  levarão às mesmas inferências sobre o número de blocos e probabilidade de mudança em cada instante, indicando a robustez em relação ao parâmetro de assimetria.

O estudo de robustez na partição  $\rho$  é mais complexo, pois existem  $2^{n-1}$  partições possíveis. Portanto, o custo computacional para avaliar a diferença entre as probabilidades de cada partição possível calculado sob os modelos contaminados é muito alto. Uma forma alternativa encontrada para estudar a robustez na partição, consiste em avaliar o comportamento da diferença entre as probabilidades de ocorrência da partição de interesse calculadas sob os modelos contaminado e basal. Seja  $\rho'$  a partição de interesse, ou seja, a partição simulada que de antemão sabe-se onde ocorrem os pontos de mudança. Logo, a robustez na inferência *a posteriori* sobre a partição é feita considerando-se a seguinte medida

$$S_g(\pi_\lambda(\rho = \rho'|\mathbf{y})) = \text{Sup}_{\mathcal{P}_M} | P_\lambda(\rho = \rho'|\mathbf{y}) - P_0(\rho = \rho'|\mathbf{y}) | - \text{Inf}_{\mathcal{P}_M} | P_\lambda(\rho = \rho'|\mathbf{y}) - P_0(\rho = \rho'|\mathbf{y}) |. \quad (4.7)$$

Note que valores da medida proposta em (4.7) próximos a zero indicam robustez no MPP em relação a partição de interesse.

Na próxima seção serão apresentados os procedimentos computacionais utilizados para obter aproximações para as medidas de robustez aqui propostas.

## 4.2 Métodos computacionais para o estudo de robustez no MPP - Modelo I e Modelo II

Segundo Godoi (2011), para se avaliar a robustez na média e na variância *a posteriori* pode-se considerar, respectivamente, as medidas

$$d_M(\pi_\lambda(\mu_i|\mathbf{y})) = | E_\lambda [\mu_i|\mathbf{y}] - E_0 [\mu_i|\mathbf{y}] |$$

e

$$d_{var}(\pi_\lambda(\mu_i|\mathbf{y})) = | Var_\lambda [\mu_i|\mathbf{y}] - Var_0 [\mu_i|\mathbf{y}] | .$$

e para obter-se uma aproximação para o  $sup_\lambda\{d_M(\pi_\lambda(\mu_i|\mathbf{y}))\}$  e  $sup_\lambda\{d_{var}(\pi_\lambda(\mu_i|\mathbf{y}))\}$  utiliza-se a metodologia de Gauss-Hermite. Neste caso, considera-se uma partição finita e suficientemente grande do espaço paramétrico de  $\lambda$ , por exemplo,  $\lambda = 0.1 \times i, i = 0 \dots T$ , e então, para cada amostra fixada, o supremo é aproximado por  $max\{d_M(\pi_{\lambda_0}(\mu|\mathbf{y})), d_M(\pi_{\lambda_1}(\mu|\mathbf{y})), \dots, d_M(\pi_{\lambda_T}(\mu|\mathbf{y}))\}$ . Similarmemente, obtém-se uma aproximação para o ínfimo tomando-se  $min\{d_M(\pi_{\lambda_0}(\mu|\mathbf{y})), d_M(\pi_{\lambda_1}(\mu|\mathbf{y})), \dots, d_M(\pi_{\lambda_T}(\mu|\mathbf{y}))\}$ . Raciocínio similar pode ser usado para fazer o estudo de robustez no MPP. No entanto, no MPP o custo computacional é muito alto.

Para estimar o número de blocos no MPP utiliza-se a variável  $U$  definida em (2.1) e para cada valor da grade finita e suficientemente grande de  $\lambda$  constrói-se a distribuição acumulada empírica de  $B$  e aproxima-se o  $Sup_{\mathcal{P}_M} | F_\lambda(b | \mathbf{y}) - F_0(b | \mathbf{y}) |$  pelo  $Max_\lambda | F_\lambda(b | \mathbf{y}) - F_0(b | \mathbf{y}) |$  e o ínfimo pelo mínimo. Logo, no Capítulo 5 onde são apresentados os resultados para a robustez global, serão apresentados dois gráficos, a saber, um com a distribuição empírica com o  $Max | F_\lambda(B \leq b | \mathbf{y}) - F_0(B \leq b | \mathbf{y}) |$  e  $Min | F_\lambda(B \leq b | \mathbf{y}) - F_0(B \leq b | \mathbf{y}) |$  e o outro com a medida  $S_g(\pi_\lambda(B|\mathbf{y}))$  referente a diferença entre essas distribuições que será denotada por  $d_B$ . Como o número de blocos é estimado pela moda *a posteriori*, na análise de robustez será mostrado também, o número estimado de blocos *a posteriori* para cada  $\lambda$  e a probabilidade *a posteriori* da moda para cada  $\lambda$ , pois por meio desses gráficos é possível visualizar o comportamento da mudança

nas estimativas de  $B$  em função de  $\lambda$ .

A probabilidade de cada instante ser um ponto de mudança é calculada como descrito em (2.2) e para avaliar a robustez em  $P(A_l)$  foram calculadas as medidas  $\widehat{P(A_l)}$  para cada instante e para cada  $\lambda$  em seguida o ínfimo foi aproximado pelo mínimo e o supremo pelo máximo e a diferença foi calculada. A medida de robustez para a probabilidade de mudança em cada instante será denotada ao longo desse texto como  $d_{P(A_l)}$ .

A análise de robustez para a partição foi limitada a partição de interesse devido a complexidade computacional, ou seja, a partição “correta” foi fixada como a partição de interesse e foi calculada a probabilidade da partição correta ocorrer para cada valor de  $\lambda$ , em seguida os valores do supremo e do ínfimo foram aproximados pelo máximo e pelo mínimo como nas demais medidas propostas neste capítulo. A medida de robustez para a partição de interesse será denotada por  $d_\rho$ .

O próximo capítulo apresenta os resultados da robustez global para os modelos I e II através das medidas aqui apresentadas.

---

## Resultados da robustez global no MPP

---

Neste capítulo serão apresentados os resultados referentes a robustez global no MPP com base nas medidas de sensibilidade global apresentadas no Capítulo 4. A análise de robustez está dividida em duas seções com base nos modelos que foram discutidos no Capítulo 2, o Modelo I que assume que a distribuição *a priori* para a média segue uma distribuição normal assimétrica e o Modelo II que assume que a distribuição dos dados é normal assimétrica.

A apresentação dos resultados segue a mesma lógica da análise de sensibilidade, primeiro serão apresentados os resultados da robustez para as estimativas produto e variância *a posteriori*, depois os resultados para a partição e número de blocos, finalizando com a probabilidade de mudança em cada instante.

### 5.1 Robustez no Modelo I

Na análise de robustez global do Modelo I os dados foram considerados idênticos dentro de cada bloco, por exemplo, o bloco  $[1; 10]$  é formado pelo vetor de observações  $y_{[1;10]} = [5, 5, 5, 5, 5, 5, 5, 5, 5, 5]$ . Esta estratégia foi adotada para garantir que para cada  $\lambda$  avaliado as estatísticas de cada bloco seriam as mesmas para qualquer tamanho de amostra, ou seja, assim como no estudo de sensibil-

idade informal visto no Capítulo 3, o estudo de robustez proposto neste trabalho não contempla réplicas Monte Carlo, pois as réplicas poderiam minimizar \ suavizar o impacto das estimativas, logo não seria possível visualizar a magnitude do impacto causado pela escolha equivocada de  $\lambda$  nas estimativas produto, no número de blocos, na partição mais provável e na probabilidade de mudança em cada instante.

Para ilustrar como a assimetria modifica a distribuição dos dados, segue a Figura 5.1 com a distribuição  $\mathcal{SN}(0, 1, \lambda)$  para alguns valores de  $\lambda$ .

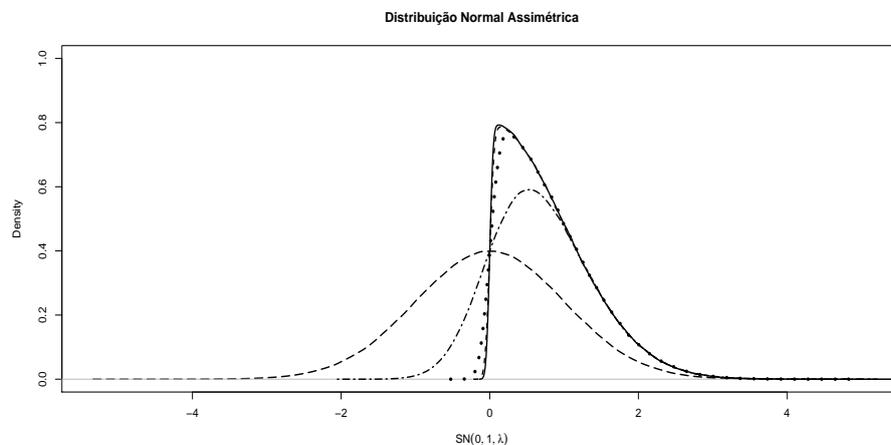


Figura 5.1: Distribuição  $\mathcal{SN}(0, 1, \lambda)$ . Linha cheia  $\lambda = 100$ , linha tracejado curto  $\lambda = 30$ , linha pontilhada  $\lambda = 10$ , Linha tracejada e pontilhada  $\lambda = 2$ , Linha tracejado longo  $\lambda = 0$ .

Note que para  $\lambda = 100$  e  $\lambda = 30$  as densidades praticamente se sobrepõem e as diferenças entre as densidades das distribuições que consideram  $\lambda = 100$ ,  $\lambda = 30$  e  $\lambda = 10$  são pequenas. Por esta razão, para avaliar a robustez global, foi fixada uma grade de  $\lambda$  no intervalo  $[-30, \dots, 30]$ , especificamente, considerar-se-a ( $\lambda = 0, \pm 0.1, \pm 0.2, \pm 0.3, \pm 0.4, \pm 0.5, \dots, \pm 29.8, \pm 29.9, \pm 30$ ) e foram considerados 2 tamanhos amostrais diferentes ( $n = 30, 90$ ). Para realizar o estudo MCMC, considerou-se uma cadeia de tamanho 60000 iterações e o período de *burn-in* foi de 50000 iterações com um lag de 10 observações, logo o tamanho da amostra *a posteriori* final foi de 1000 observações. Esse valor de *burn-in* foi escolhido, pois garantiu a convergência individual dos parâmetros do modelo.

Como na análise de sensibilidade informal, os resultados para os cenários escada crescente e escada

decrecente foram bem similares, bem como os resultados dos cenários degrau positivo e degrau negativo. Estes cenários onde não foram observadas mudanças substanciais, não foram considerados no estudo de robustez global. A Tabela 5.1 a seguir ilustra os cenários para esse estudo de sensibilidade global.

Tabela 5.1: Cenários avaliados

Cenário	Média no 1 <sup>o</sup> bloco	Média no 2 <sup>o</sup> bloco	Média no 3 <sup>o</sup> bloco	Média no 4 <sup>o</sup> bloco
Degrau simétrico	5	8	5	–
Degrau não simétrico	–3	7	0	–
Escada crescente	7	10	13	–
Epsilon	1.75	2.5	3.25	–
Tendência	0	$0.3 + 0.008 * i$	$0.6 + 0.018 * i$	1.1

Em todos os cenários propostos as séries possuem dois pontos de mudança, com exceção do cenário tendência que possui 4 blocos, mas todos os “clusters” possuem tamanhos iguais. Assim, a localização dos pontos de mudança variam de acordo com o número de observações na série, ou seja, na série com 30 observações os pontos de mudança são 11 e 21, o cenário tendência com 40 observações apresenta os pontos de mudança 11,21 e 31, na série com 90 observações os pontos de mudança estão localizados nas posições 31 e 61 e no cenário tendência com 120 observações os pontos de mudança são 31,61 e 91.

Além dos cenários apresentados na Tabela 5.1, outros cenários são considerados, de forma que as médias dos blocos descritos na Tabela 5.1 são acrescidas de uma quantidade  $\epsilon$ , tal que  $\epsilon$  assume os valores 2, 5 e 50, conseqüentemente o valor do hiperparâmetro  $\mu_0$  muda de acordo com o valor de  $\epsilon$ . A quantidade  $\epsilon$  é acrescida para garantir que os resultados não são específicos para as médias consideradas na Tabela 5.1. O MPP é então ajustado fixando valores ora positivos e ora negativos para  $\lambda$  e a meta é avaliar como as estimativas *a posteriori* se comportam em relação a distribuição basal assumindo  $\lambda > 0$  e  $\lambda < 0$ .

Os resultados apresentados a seguir foram calculados com base nas medidas de distâncias entre as estimativas *a posteriori* definidas no Capítulo 4 e as distribuições *a priori* foram as mesmas

utilizadas no Capítulo 3.

### 5.1.1 Robustez nas estimativas produto

Esta seção apresenta os resultados da robustez global para as estimativas produto e a variância destas estimativas. Os gráficos (a) e (c) mostram os resultados da medida  $d_M$  definida na Seção 4.1 para  $n = 30$  e  $n = 90$ , respectivamente e as Figuras (b) e (d) apresentam as diferenças  $d_{var}$  para  $n = 30$  e  $n = 90$ , respectivamente.

No modelo partição produto, como visto na Seção 2.1, não estima-se uma única média, mas sim  $n$  médias. Esta característica do MPP dificulta a apresentação dos resultados para a medida de robustez global da média. Logo, os gráficos não são capazes de mostrar em que sentido a magnitude das diferenças caminha. Neste caso, só é possível verificar o quão grande foi a diferença assumindo a grade de  $\lambda = (-30, -29.9, \dots, -0.1, 0)$  ou  $\lambda = (0, 0.1, \dots, 29.9, 30)$ .

As Figuras 5.2 e 5.3 apresentam os resultados para o cenário degrau simétrico positivo, quando os dados são ajustados com  $\lambda < 0$  e  $\lambda > 0$ .

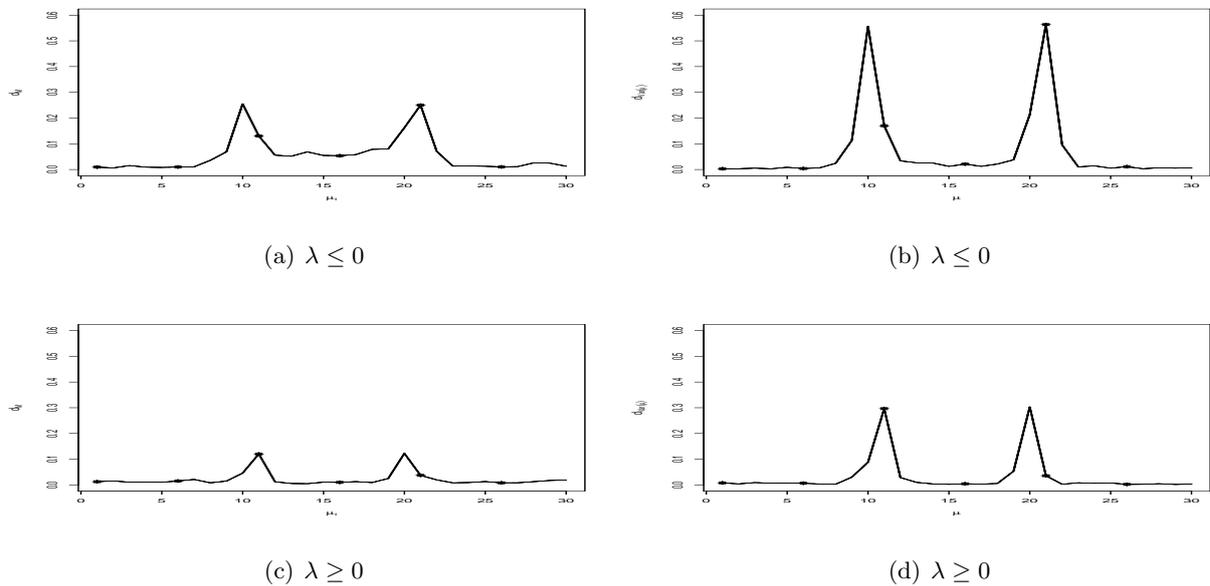


Figura 5.2: Cenário degrau positivo - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .

Percebe-se da Figura 5.2 que as estimativas produto assumindo que a distribuição *a priori* para  $\mu$  pertence a família normal assimétrica tendem a ser próximas independente do valor de  $\lambda$  assumido, uma vez que a magnitude das distâncias  $d_M$  são próximas de zero. No entanto, próximo aos pontos de mudança, as distâncias  $d_M$  são maiores indicando que as estimativas produto são mais afetadas pelas escolhas da distribuição *a priori* de  $\mu$ , principalmente, se o sinal de  $\lambda$  escolhido e da média dos blocos estão em desacordo. Nota-se, também, que as variâncias *a posteriori* de  $\mu_i$  longe dos pontos de mudança não são muito afetadas pela escolha de  $\lambda$ , mas isto não ocorre no entorno dos pontos de mudança.

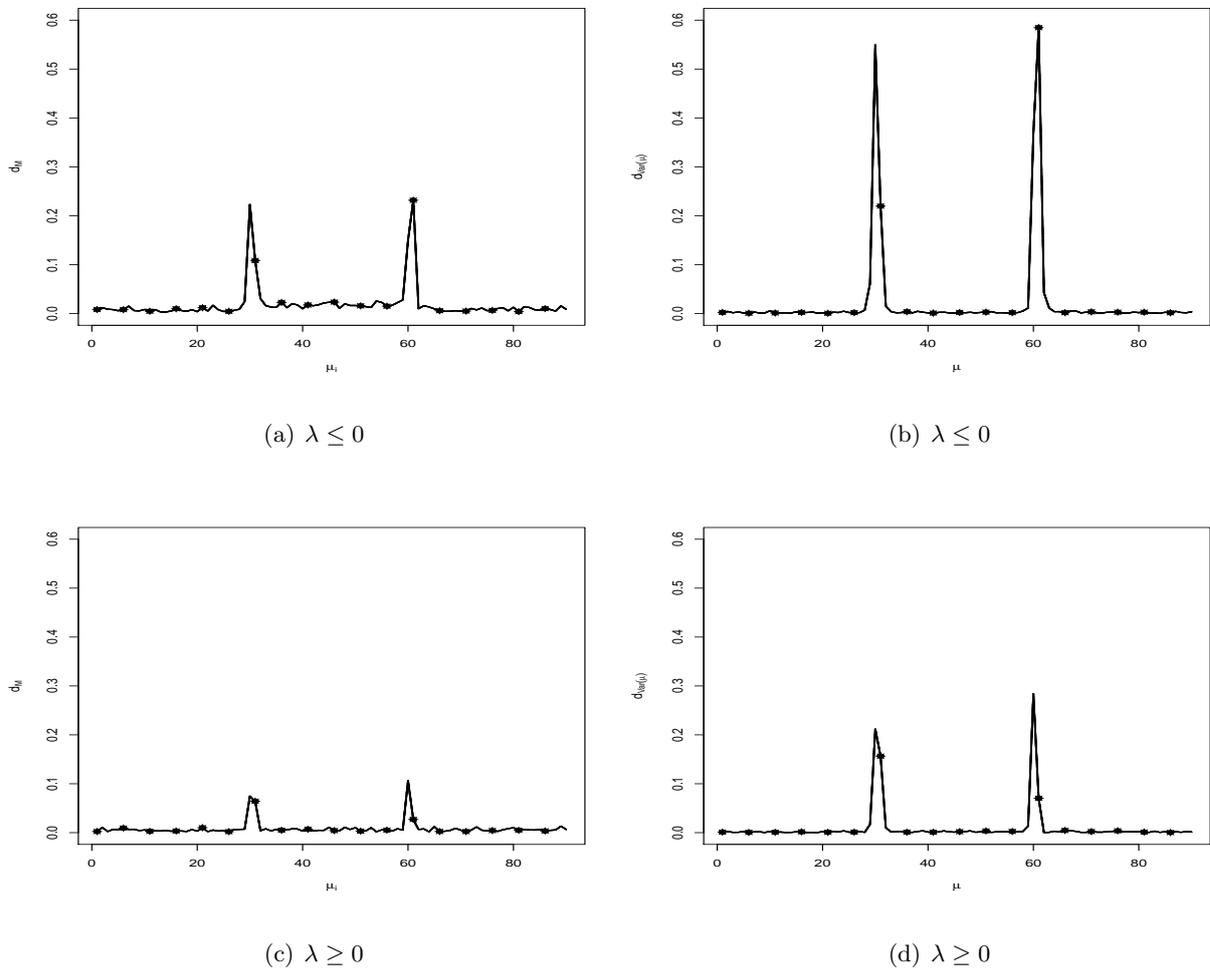


Figura 5.3: Cenário degrau positivo - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

A Figura 5.3 mostra que a medida  $d_{var}$  apresenta uma maior diferença na variabilidade *a posteriori* da distribuição de  $\mu_i$  próximo aos pontos de mudança, e a magnitude das diferenças foi próxima de 0.6 para ambos os tamanhos de amostra avaliados. Nota-se também, que para  $n = 90$  as diferenças  $d_M$  e  $d_{var}$  tendem a ser próximas das obtidas para  $n = 30$ . Cabe, também, mencionar que as diferentes informações amostrais conduzem às mesmas conclusões.

As Figuras 5.4 e 5.5 mostram os resultados para o degrau não simétrico, onde o tamanho do salto entre os blocos é maior que no cenário degrau positivo.

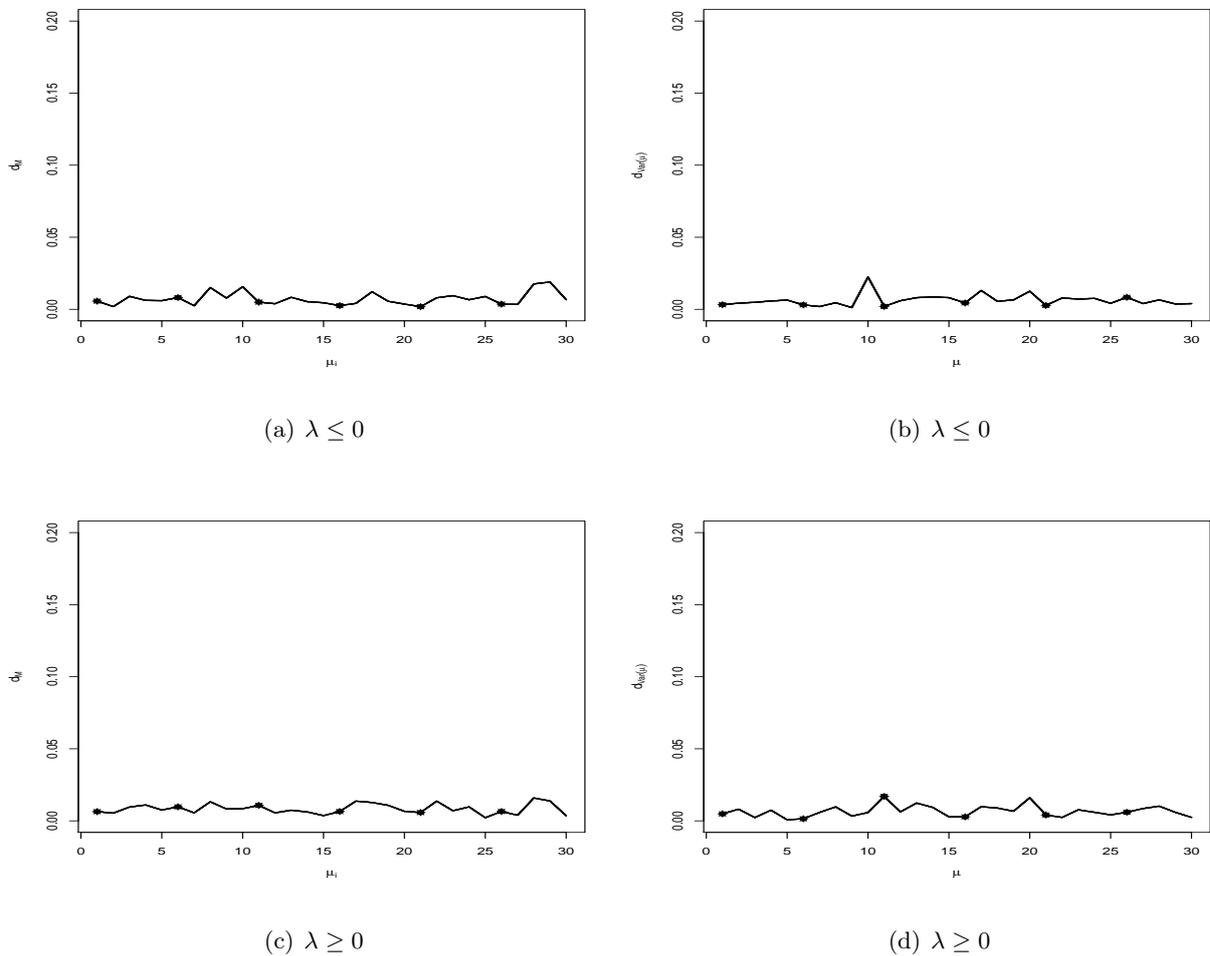


Figura 5.4: Cenário degrau não simétrico - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com \*  $\epsilon = 50$ ,  $n = 30$ .

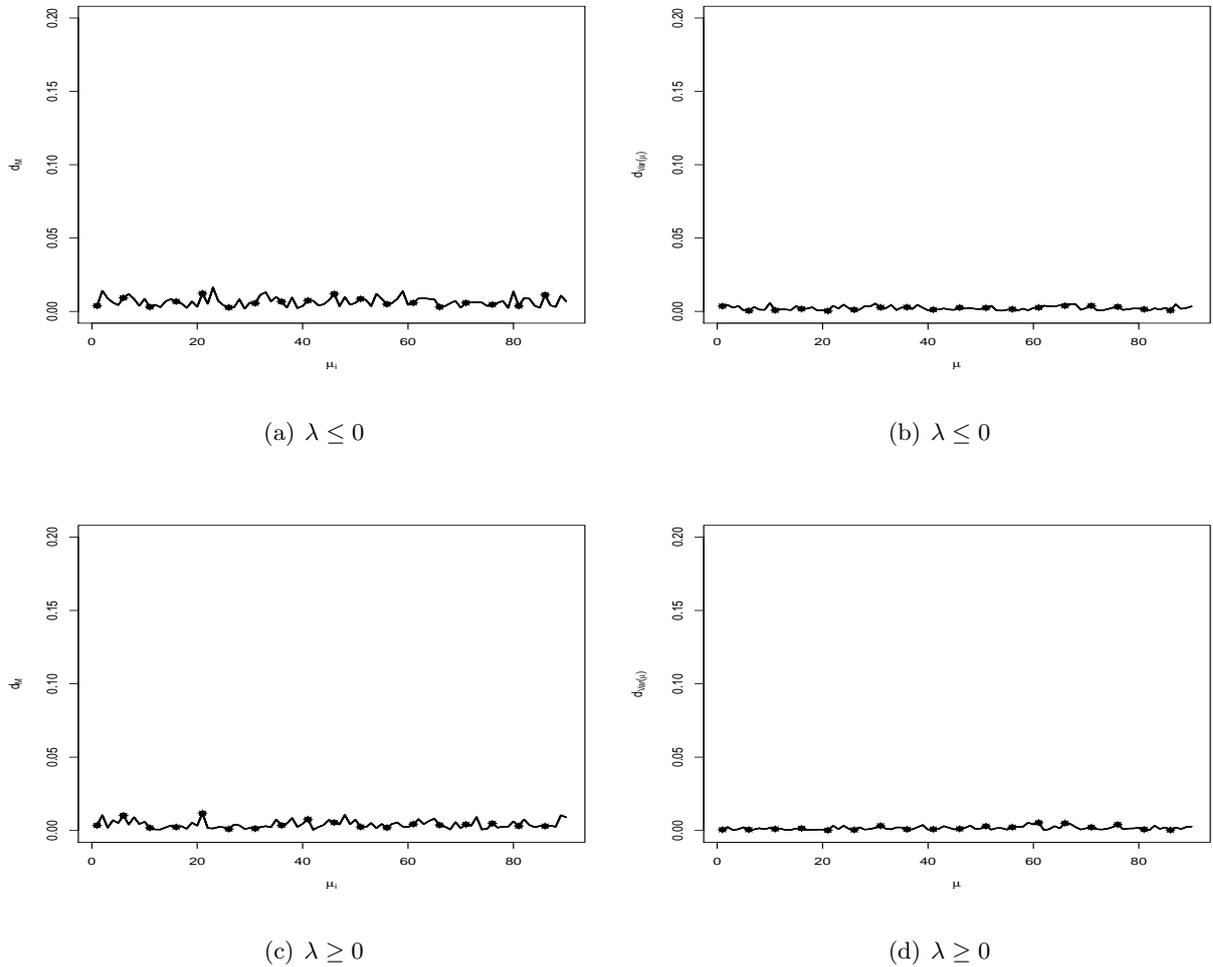


Figura 5.5: Cenário degrau não simétrico - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

No cenário degrau não simétrico observa-se que a magnitude da distância  $d_M$  não é maior próximo aos pontos de mudança, como observou-se no cenário degrau positivo. Talvez, isso ocorra porque a diferença da média de um bloco para outro é maior. A distância  $d_{var}$  também está próxima de zero para todos os pontos da série. Além disso, o valor das médias dos blocos não interfere no resultado, pois para todo  $\epsilon$  considerado as diferenças se mantiveram próximas a zero.

As Figuras 5.6 e 5.7 ilustram os resultados para o cenário escada crescente, assumindo  $n = 30$  e  $n = 90$ , respectivamente.

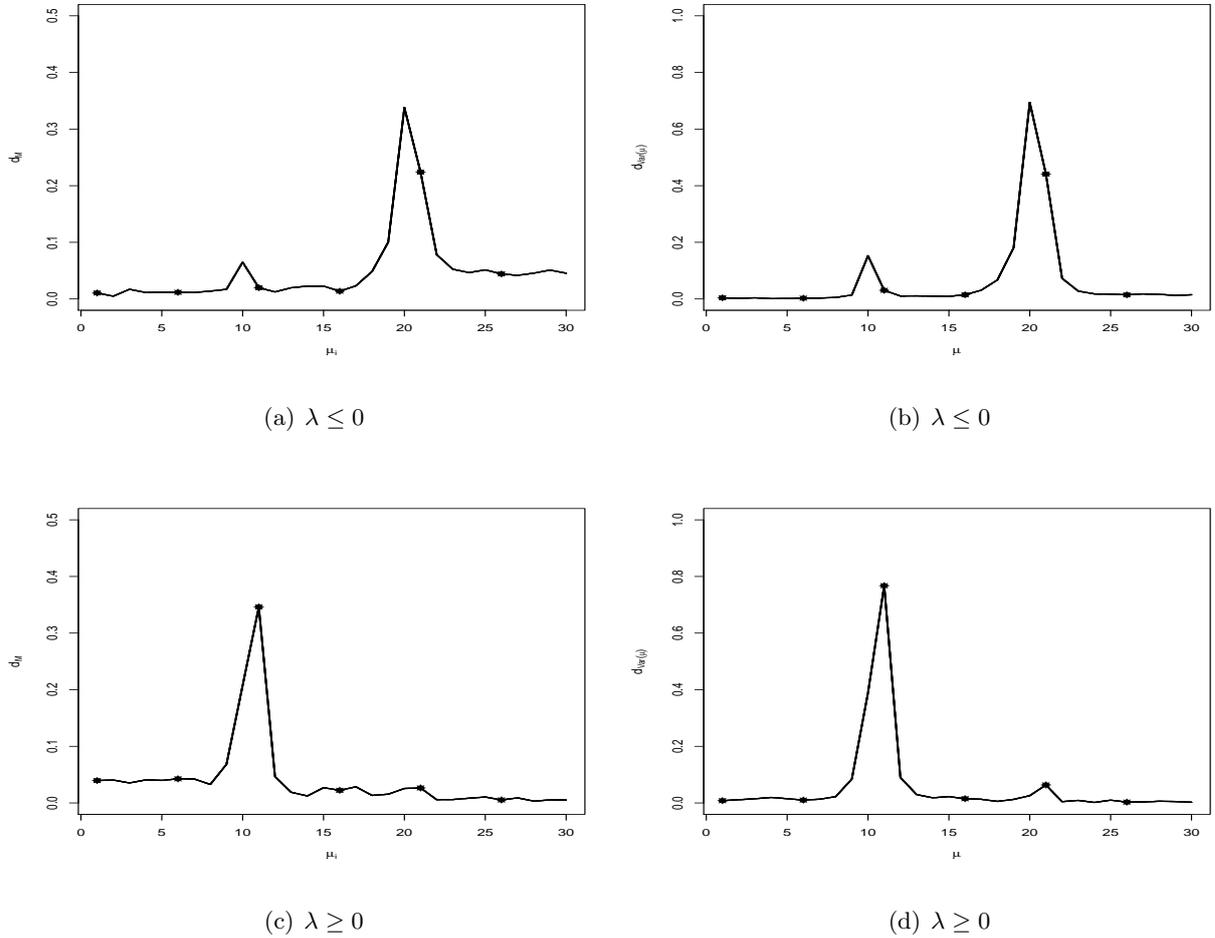


Figura 5.6: Cenário escada crescente - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com \*  $\epsilon = 50$ ,  $n = 30$ .

Observa-se nas Figuras 5.6(b) e (d) e 5.7(b) e (d) que as maiores diferenças entre  $d_{var}$  ocorrem em torno do instante do primeiro ponto de mudança quando  $\lambda > 0$ , mas se  $\lambda < 0$  as maiores diferenças ocorrem em torno do segundo ponto de mudança.

O mesmo comportamento de maior magnitude da medida  $d_{var}$  próximo aos pontos de mudança é observado para a medida  $d_M$ , ou seja, se  $\lambda < 0$  as maiores diferenças estão em torno do primeiro ponto de mudança. Comportamento similar foi observado no estudo de sensibilidade *a priori* onde as estimativas do último bloco ficaram comprometidas para  $\lambda = -100$  no cenário escada crescente.

Nota-se também, que as diferenças tanto para  $d_M$ , quanto para  $d_{var}$  são as mesmas independentemente dos valores de  $\epsilon$  avaliados.

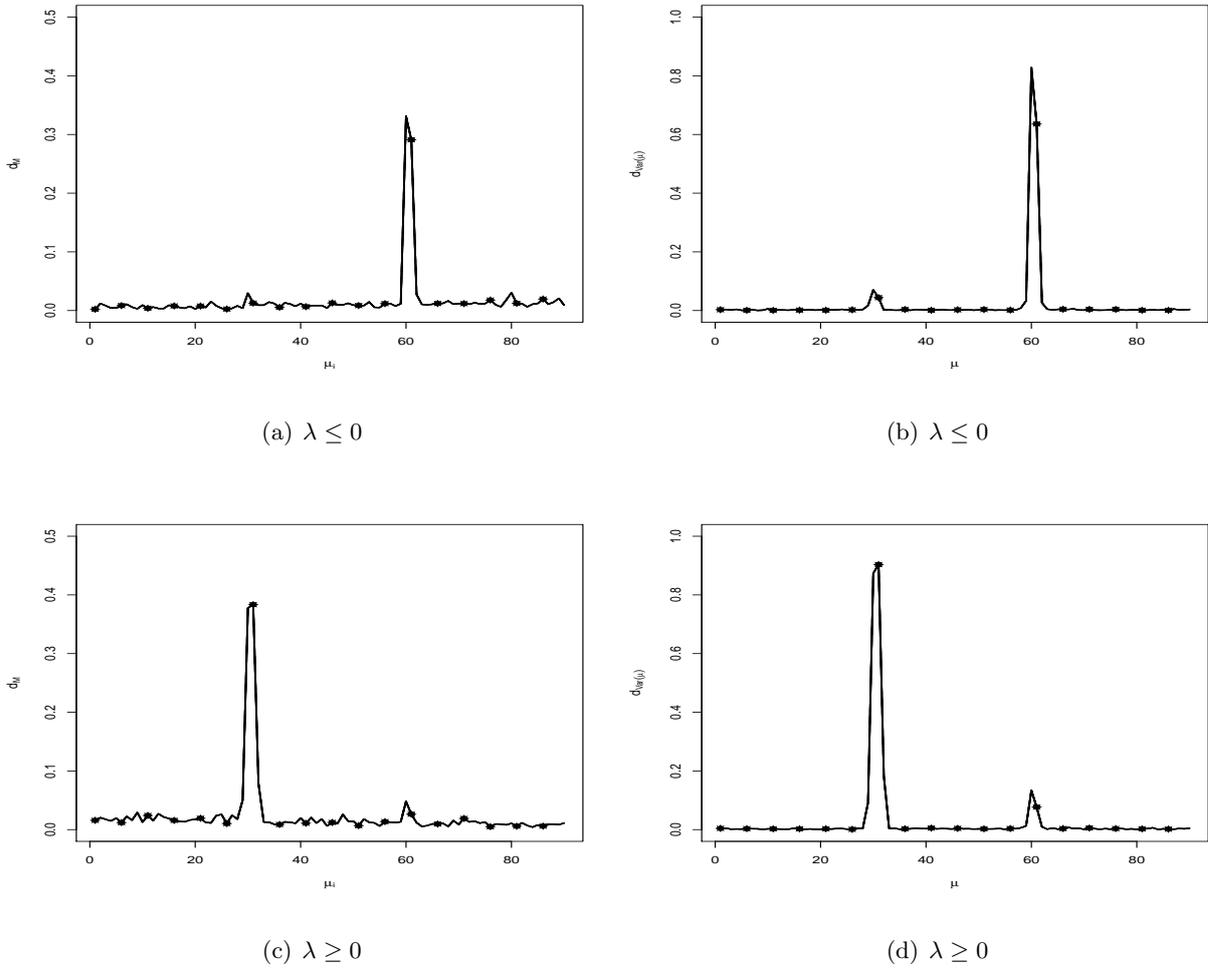
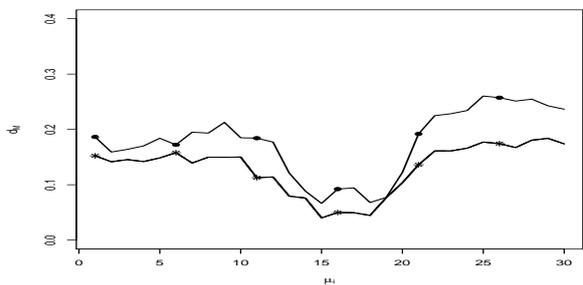
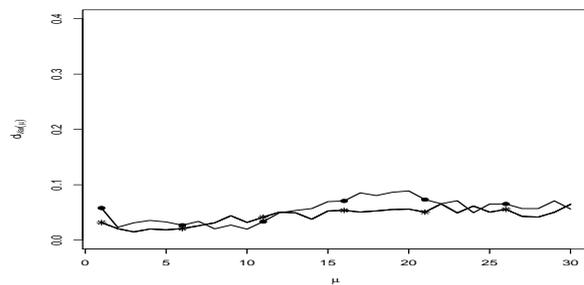


Figura 5.7: Cenário escada crescente - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

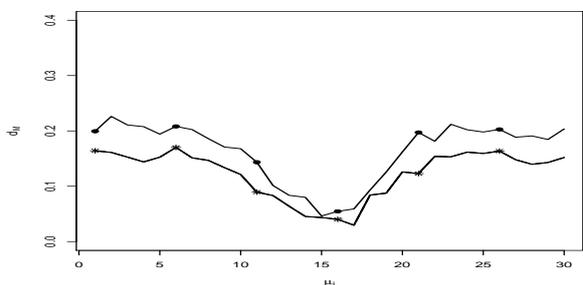
As Figuras 5.8 e 5.9 ilustram os resultados para o cenário Epsilon, nesse cenário as diferenças entre os blocos é de 0.75 e a série também possui o comportamento escada.



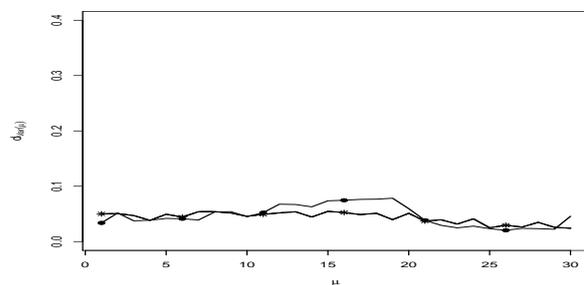
(a)  $\lambda \leq 0$



(b)  $\lambda \leq 0$

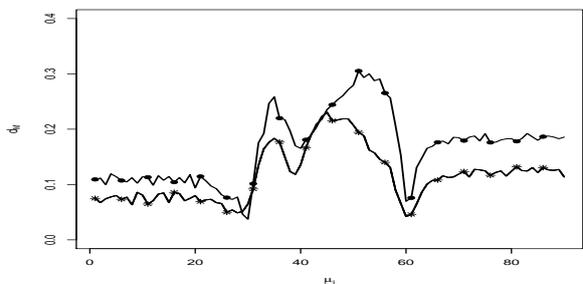


(c)  $\lambda \geq 0$

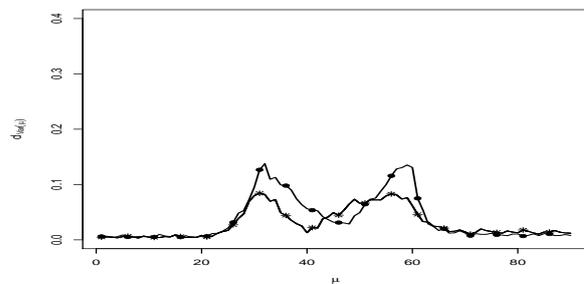


(d)  $\lambda \geq 0$

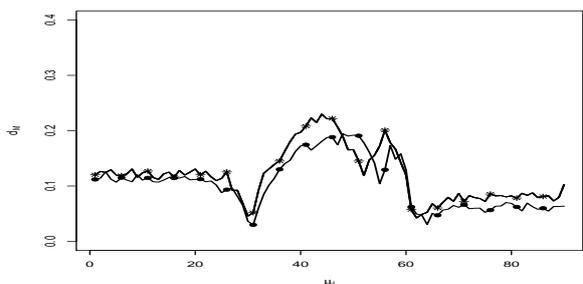
Figura 5.8: Cenário epsilon - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .



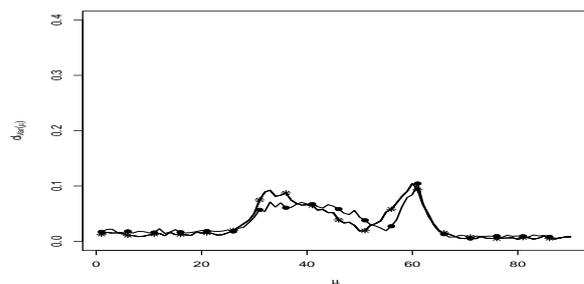
(a)  $\lambda \leq 0$



(b)  $\lambda \leq 0$



(c)  $\lambda \geq 0$



(d)  $\lambda \geq 0$

Figura 5.9: Cenário epsilon - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

Nota-se no cenário epsilon que a magnitude das medidas não excede 0.4, no entanto, vale ressaltar que na análise de sensibilidade foi possível perceber que o modelo não consegue captar corretamente um dos pontos de mudança para  $n = 30$ . Em relação a variância das estimativas de  $\mu_i$  há uma variação maior no segundo bloco para  $n = 90$ . A magnitude das diferenças também são maiores no segundo bloco para  $n = 90$ , mas para  $n = 30$  a grandeza da diferença é maior para o primeiro e o segundo bloco.

As Figuras 5.10 e 5.11 referem-se ao cenário tendência, nesse cenário as observações do primeiro bloco são iguais a 0, no segundo bloco e no terceiro bloco é incorporada uma tendência determinística as observações, por fim o quarto bloco volta a ser constituído em torno de uma média fixa, nesse caso 1.1.

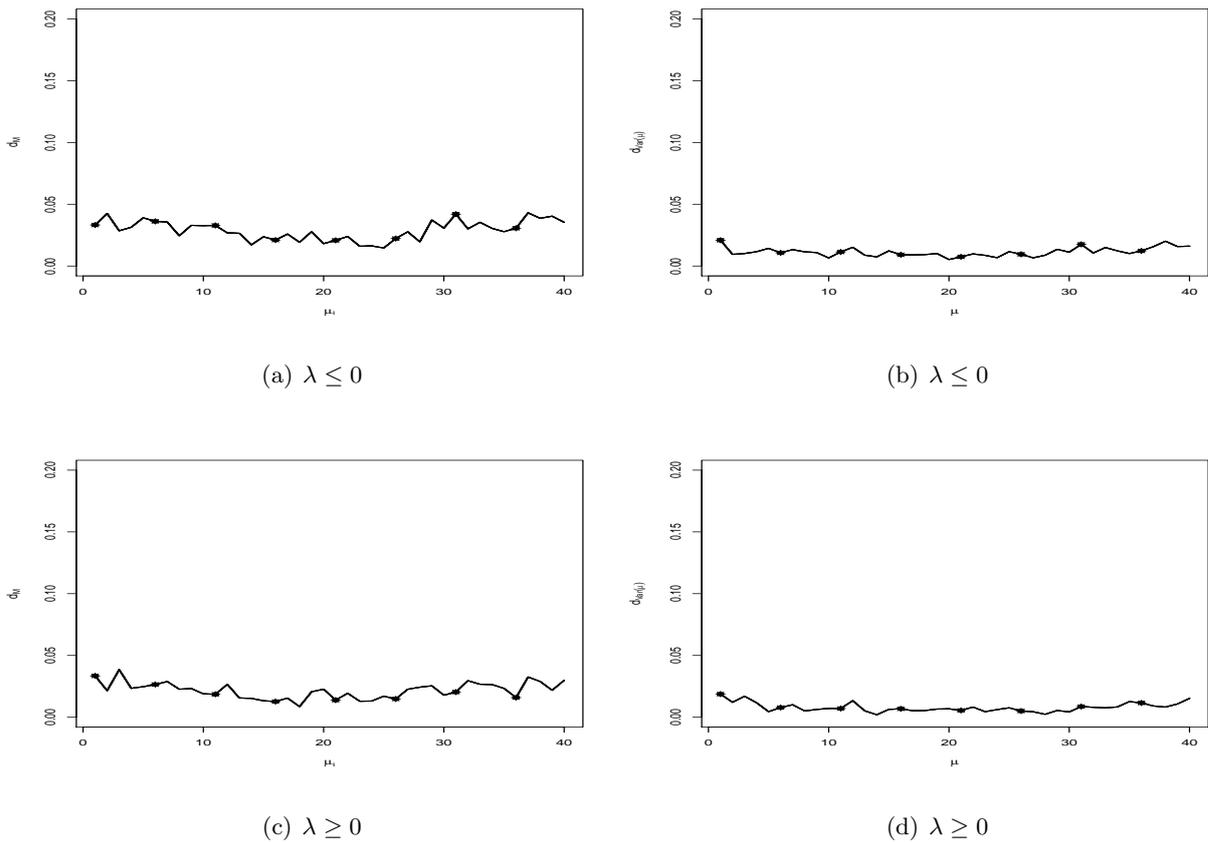


Figura 5.10: Cenário tendência - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com \*  $\epsilon = 50$ ,  $n = 40$ .

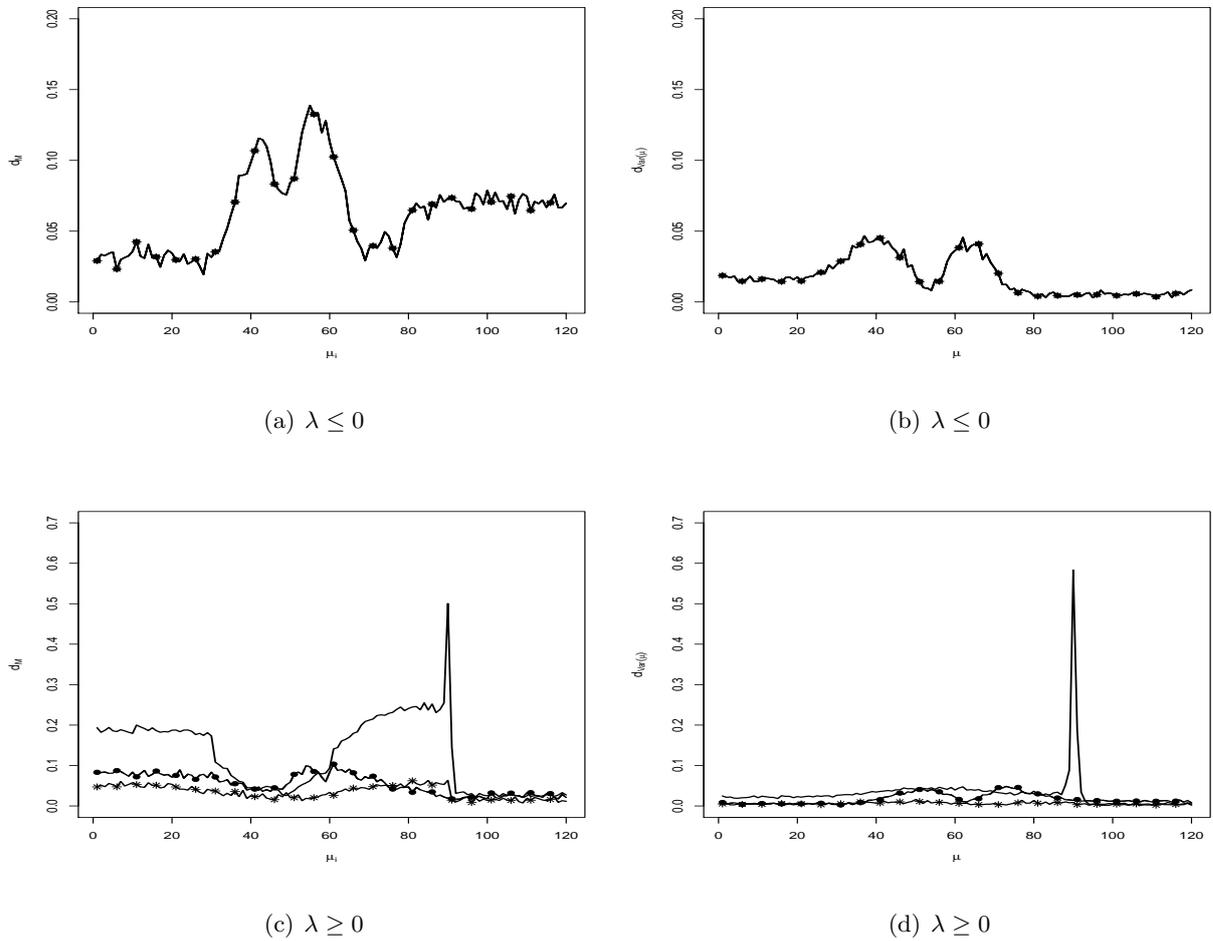


Figura 5.11: Cenário tendência - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 120$ .

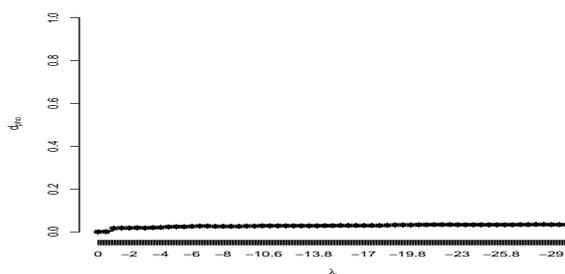
Observa-se que no cenário tendência que a magnitude da diferença para  $n = 40$  é menor do que para  $n = 120$ , mas na análise de sensibilidade o aumento no tamanho da amostra resultou em uma melhora nas estimativas produto. Nota-se que a maior variação na magnitude das estimativas ocorreu no ponto 91 e para  $\lambda > 0$ . As quantidades  $\epsilon$  adicionadas interferem nas estimativas produto, neste caso,  $\epsilon = 0$  retornou as maiores variações nas estimativas produto.

A Seção 5.1.2, a seguir, apresenta os resultados obtidos para a robustez global em relação ao número de blocos e identificação correta da partição de interesse.

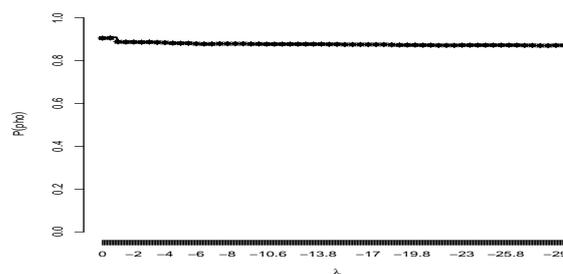
### 5.1.2 Robustez na partição e no número de blocos

As medidas apresentadas nos gráficos dessa seção são referentes as medidas definidas na Subseção 4.1.1. As figuras referem-se a robustez na partição e o comportamento do número de blocos na robustez global. Para facilitar a visualização, os gráficos (a) e (e) com a função de distribuição de probabilidade foram construídos ligando os pontos, pois utilizando os saltos não estava claro as diferença entre as funções.

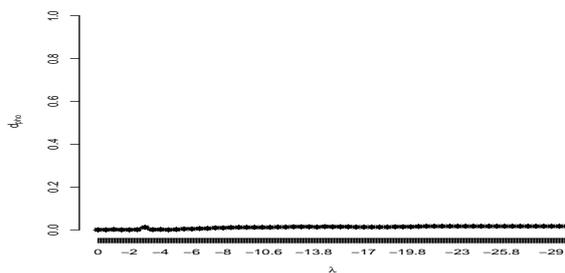
As Figuras 5.12 e 5.13 apresentam os resultados quando a partição de interesse escolhida é a partição real, ou seja, se  $n = 30$ , Gráficos (a) e (b), a partição de interesse que o modelo deve identificar é  $\rho = \{0, 10, 20\}$  e se  $n = 90$ , gráficos (c) e (d) a partição a ser identificada é  $\rho = \{0, 30, 60\}$ . Portanto, os gráficos (b) e (d) apresentam as probabilidades do modelo identificar a partição de interesse para cada  $\lambda$  e os gráficos (a) e (c) ilustram a medida  $d_\rho$  para cada  $\lambda$ .



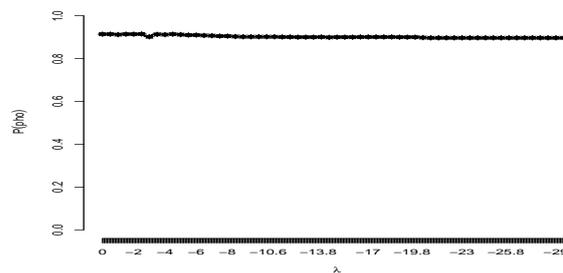
(a)  $d_\rho - n = 30$



(b)  $P(\rho = \rho') - n = 30$

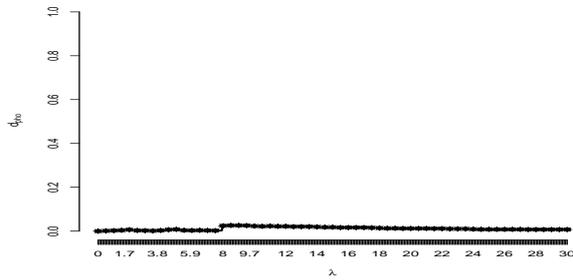


(c)  $d_\rho - n = 90$

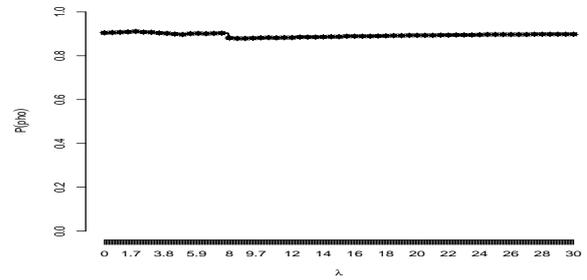


(d)  $P(\rho = \rho') - n = 90$

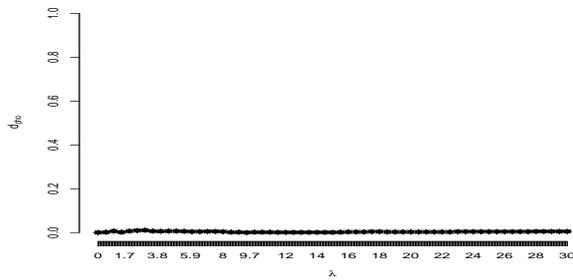
Figura 5.12: Cenário degrau não simétrico - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .



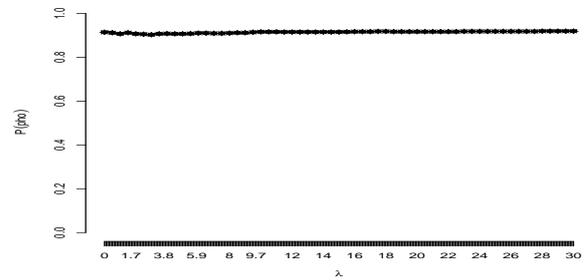
(a)  $d_\rho - n = 30$



(b)  $P(\rho = \rho') - n = 30$



(c)  $d_\rho - n = 90$

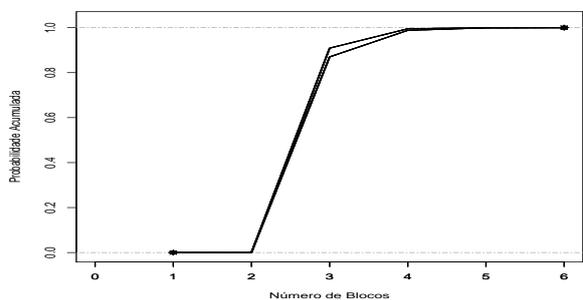


(d)  $P(\rho = \rho') - n = 90$

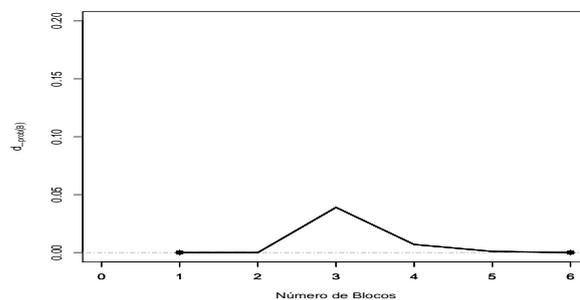
Figura 5.13:  $\lambda \geq 0$  - Cenário degrau não simétrico - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .

Nota-se no cenário degrau não simétrico que a medida  $d_\rho$  apresenta variações próximas a zero quando o tamanho da amostra é de 90 e  $\lambda \geq 0$ . Para  $\lambda \leq 0$  a medida  $d_\rho$  apresenta comportamento similar para ambos os tamanhos de amostra, indicando que para cada  $\lambda$  não há diferença entre as probabilidades *a posteriori* da partição real ser identificada corretamente. Além disso, percebe-se dos gráficos (b) e (d), que a partição real ocorre com grande probabilidade *a posteriori* para todos os valores de  $\lambda$ , mas para  $\lambda = 0$  observa-se a maior probabilidade da partição real ser estimada. Conclusões similares ocorrem para qualquer informação amostral considerada.

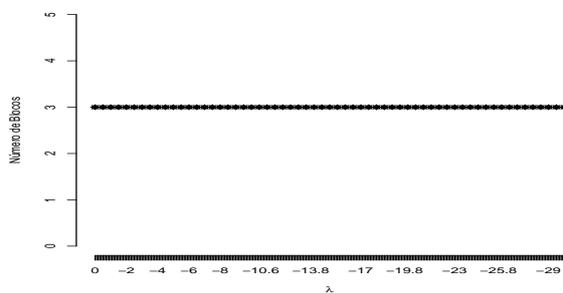
Os resultados do estudo de robustez global para o número de blocos no cenário degrau não simétrico encontram-se nas Figuras 5.14 e 5.15. Os gráficos (a), (b), (e) e (f) referem-se as medidas apresentadas na expressão (4.5) para a distribuição do número de blocos e os gráficos (c), (d), (g) e (h) ilustram a moda *a posteriori* para cada  $\lambda$ .



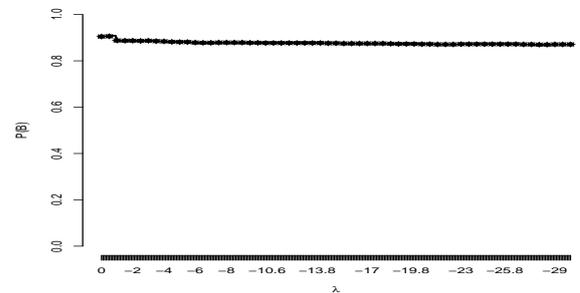
(a)  $P(B \leq b) - n = 30$



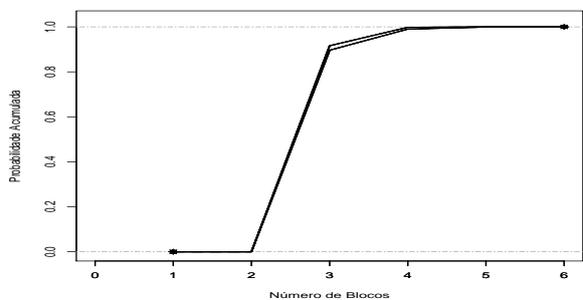
(b)  $|P_\lambda(B \leq b) - P_{\lambda=0}(B \leq b)| - n = 30$



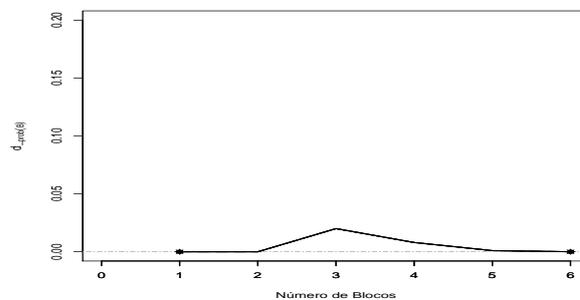
(c) Moda -  $n = 30$



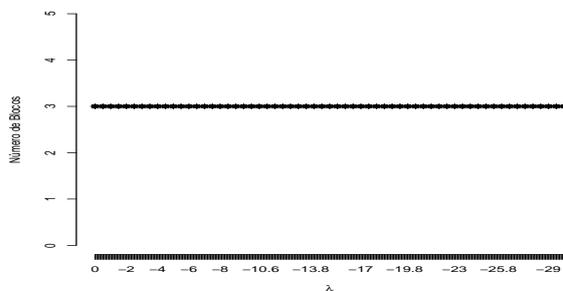
(d)  $P(B=b) - n = 30$



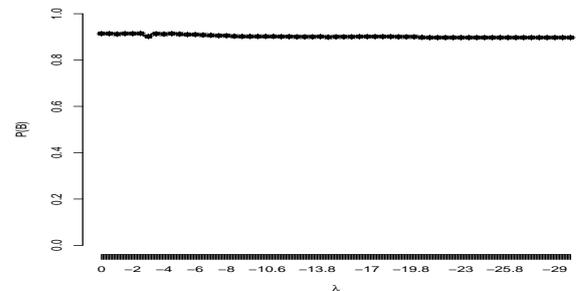
(e)  $P(B \leq b) - n = 90$



(f)  $|P_\lambda(B \leq b) - P_{\lambda=0}(B \leq b)| - n = 90$

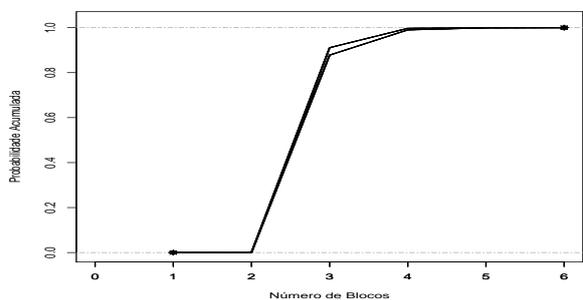


(g) Moda -  $n = 90$

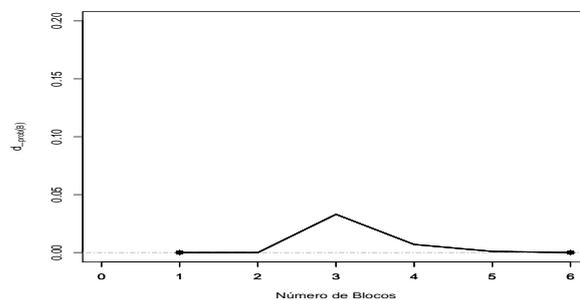


(h)  $P(B=b) - n = 90$

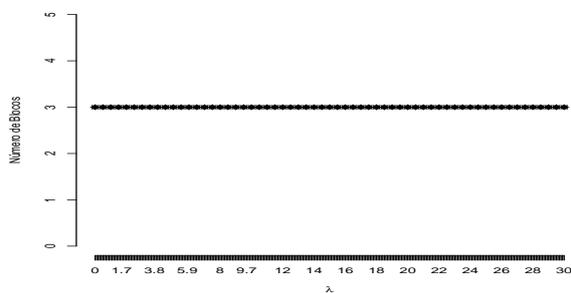
Figura 5.14:  $\lambda \leq 0$  - Cenário degrau não simétrico - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .



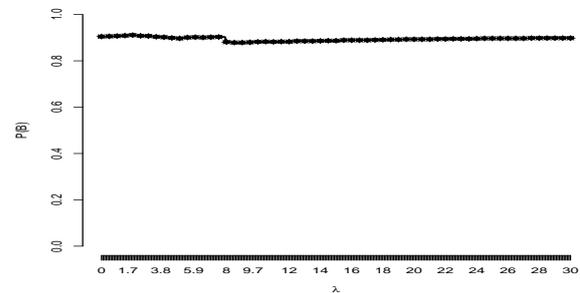
(a)  $P(B \leq b) - n = 30$



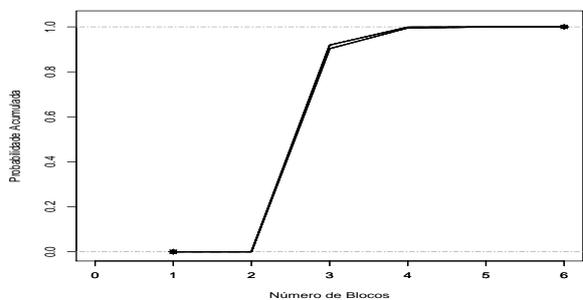
(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 30$



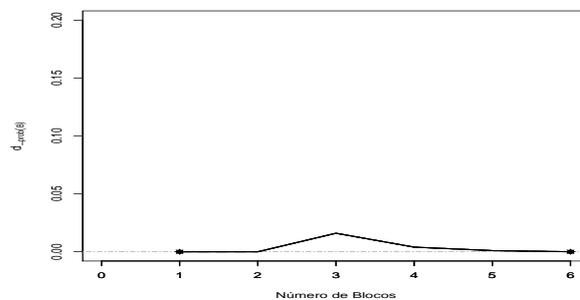
(c) Moda -  $n = 30$



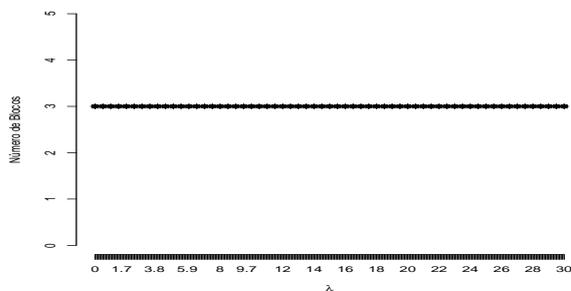
(d)  $P(B=b) - n = 30$



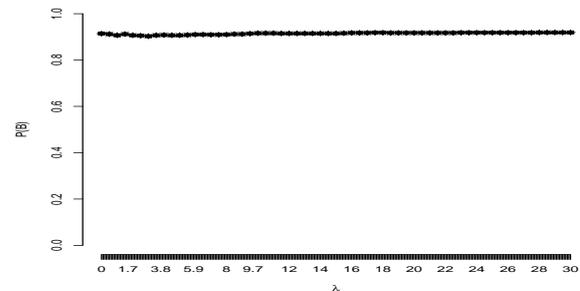
(e)  $P(B \leq b) - n = 90$



(f)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 90$



(g) Moda -  $n = 90$

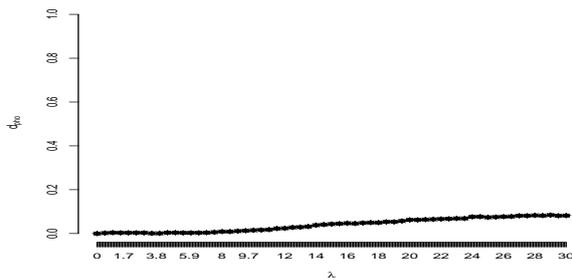


(h)  $P(B=b) - n = 90$

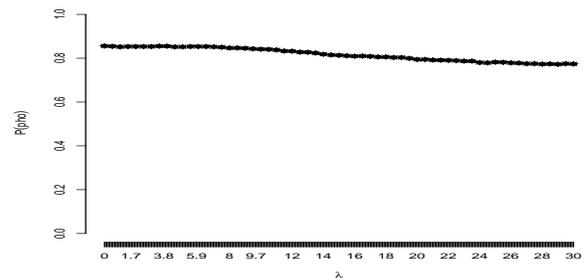
Figura 5.15:  $\lambda \geq 0$  - Cenário degrau não simétrico - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .

Em relação ao número de blocos, para todos os valores de  $\lambda$  o número de blocos é identificado corretamente ( $B = 3$ ). A diferença entre as distribuições de probabilidade são ligeiramente maiores para  $n = 30$  em relação a  $n = 90$ . Além disso, o número de blocos  $B = 3$ , é identificado com maior probabilidade quando  $\lambda = 0$ , mas para  $n = 90$  a diferença entre as probabilidades assumindo  $\lambda = 0$ ,  $\lambda \leq 0$  ou  $\lambda \geq 0$  são mínimas. Nota-se também que as medidas de robustez não são afetadas pelos valores da amostra, uma vez que as medidas foram as mesmas para os valores de  $\epsilon$  considerados.

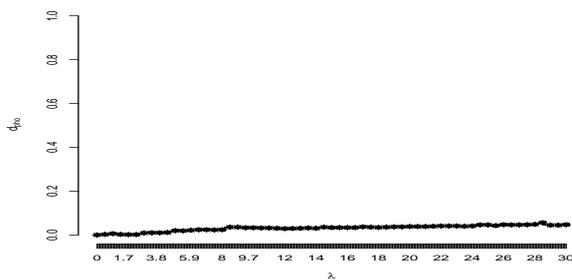
As Figuras 5.16 e 5.17 mostram os resultados para o cenário degrau não simétrico em relação a partição de interesse. Observa-se na Figura 5.17 que a distância  $d_\rho$  é maior se o modelo ajustado assume  $\lambda < 0$ . Se o modelo assume  $\lambda > 0$  a medida  $d_\rho$  apresenta menores variações para  $n = 90$  em relação a  $n = 30$ . A probabilidade do modelo identificar a partição de interesse é maior para  $\lambda = 0$ , principalmente nos casos em que o modelo é ajustado com  $\lambda < 0$ .



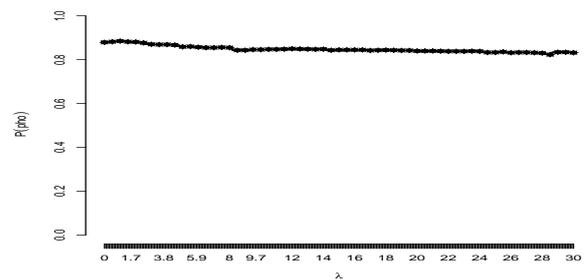
(a)  $d_\rho - n = 30$



(b)  $P(\rho = \rho') - n = 30$

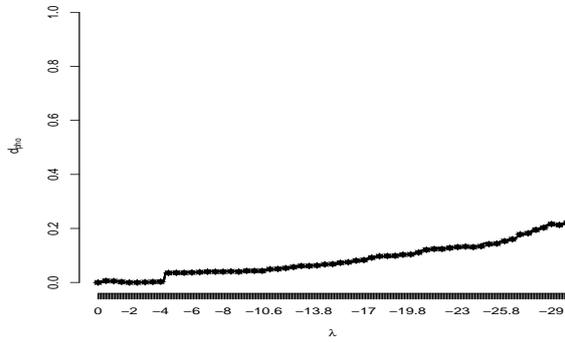


(c)  $d_\rho - n = 90$

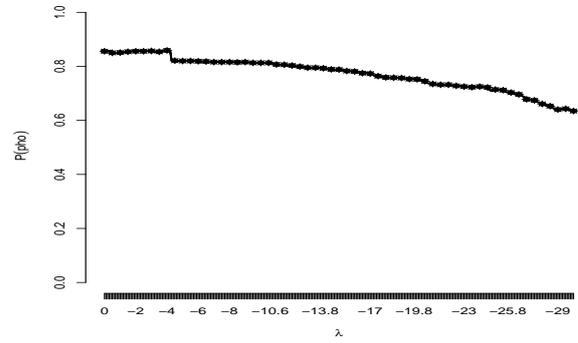


(d)  $P(\rho = \rho') - n = 90$

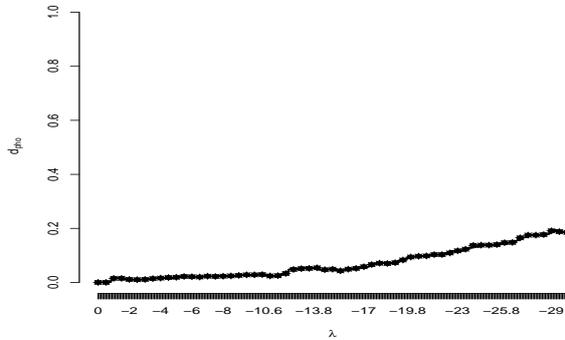
Figura 5.16:  $\lambda \geq 0$  - Cenário degrau positivo - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .



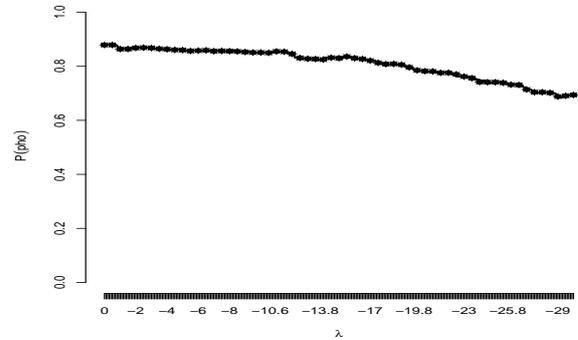
(a)  $d_\rho - n = 30$



(b)  $P(\rho = \rho') - n = 30$



(c)  $d_\rho - n = 90$

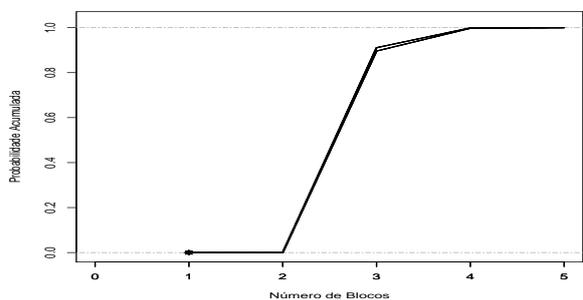


(d)  $P(\rho = \rho') - n = 90$

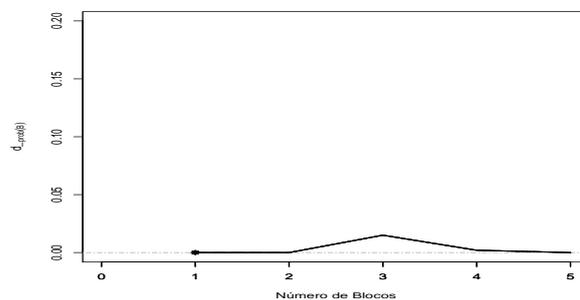
Figura 5.17:  $\lambda \geq 0$  - Cenário degrau positivo - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

As Figuras 5.18 e 5.19 mostram os resultados para o número de blocos no cenário degrau positivo. Nota-se que a diferença entre as funções de distribuição para  $\lambda > 0$  são pequenas. Essa diferença é um pouco maior se  $\lambda < 0$  for ajustado para o cenário degrau positivo. Observa-se também, que os resultados são os mesmos independente dos valores de  $\epsilon$ .

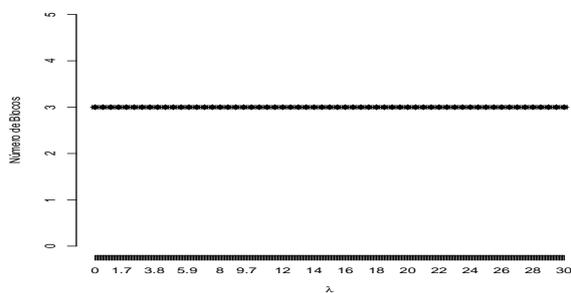
O número de blocos é identificado corretamente para todos os valores de  $\lambda$ , mas para  $\lambda > 0$  as probabilidades estão mais próximas, Figuras 5.18 (d) e (h).



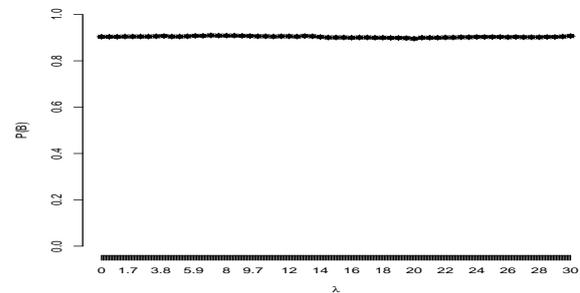
(a)  $P(B \leq b) - n = 30$



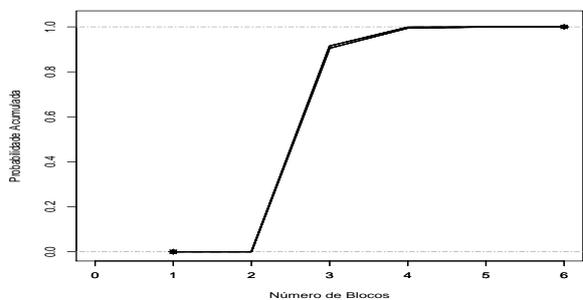
(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 30$



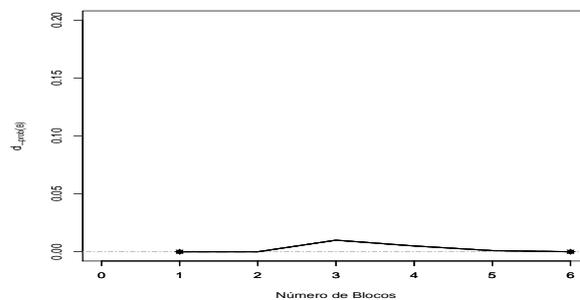
(c) Moda -  $n = 30$



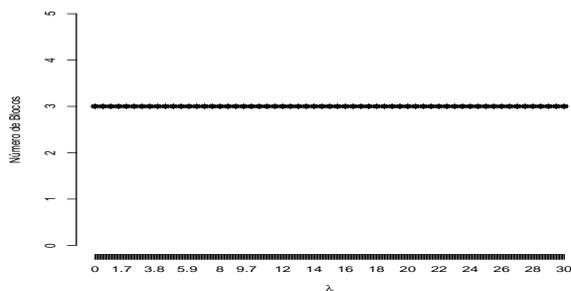
(d)  $P(B=b) - n = 30$



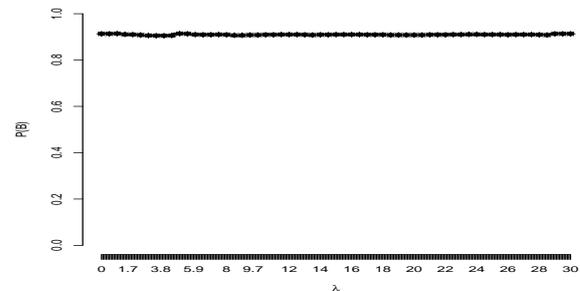
(e)  $P(B \leq b)$



(f)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 90$

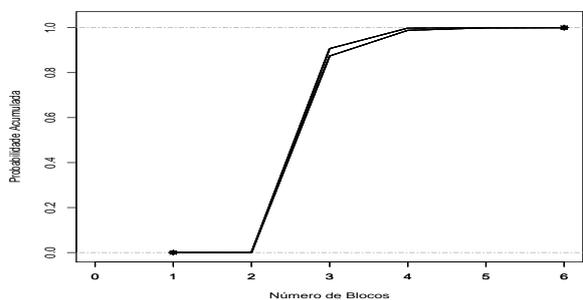


(g) Moda -  $n = 90$

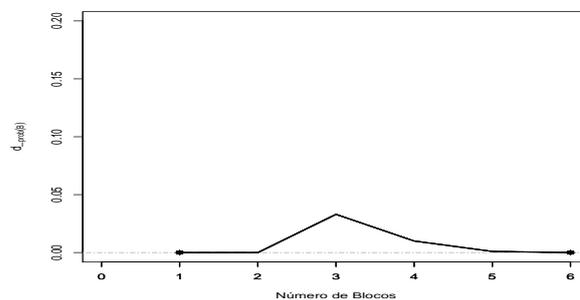


(h)  $P(B=b) - n = 90$

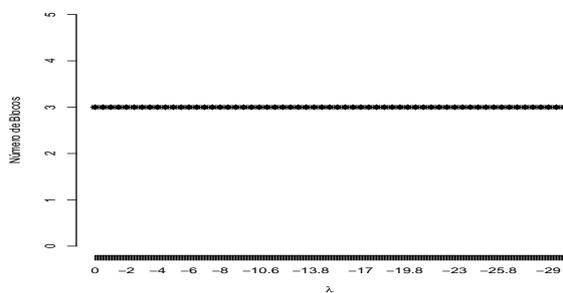
Figura 5.18:  $\lambda \geq 0$  - Cenário degrau positivo - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .



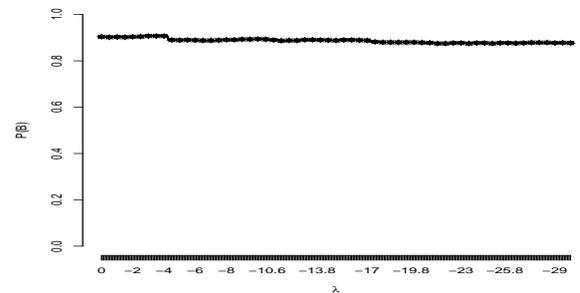
(a)  $P(B \leq b) - n = 30$



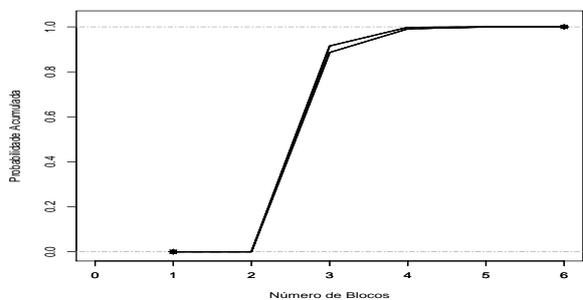
(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 30$



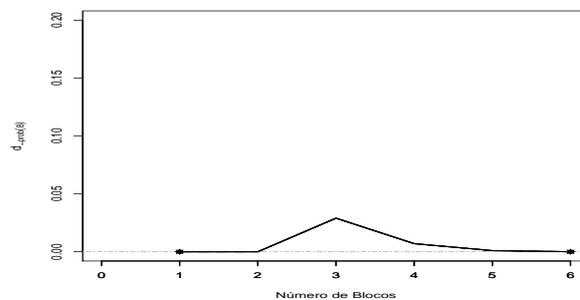
(c) Moda -  $n = 30$



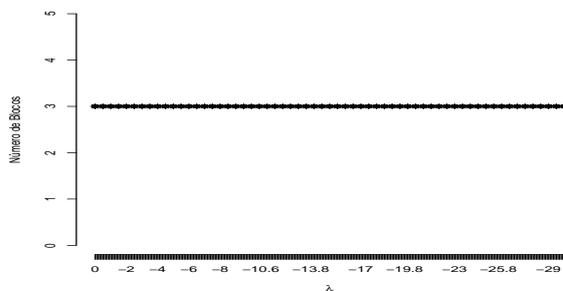
(d)  $P(B=b) - n = 30$



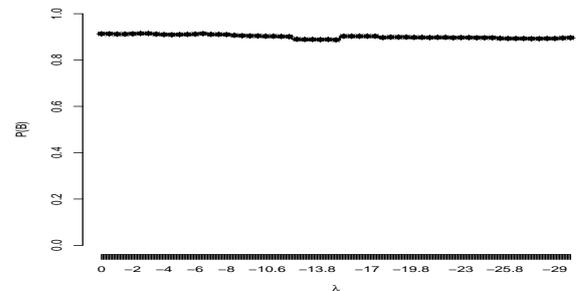
(e)  $P(B \leq b) - n = 90$



(f)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 90$



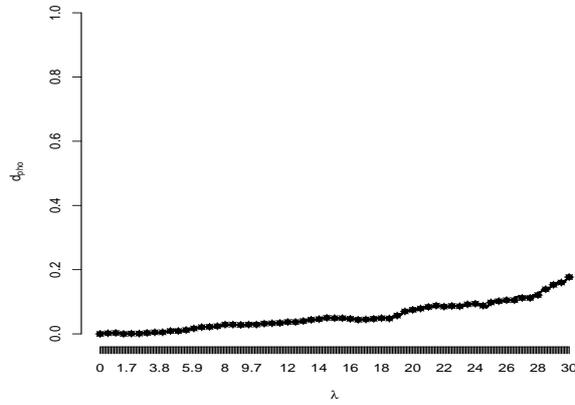
(g) Moda -  $n = 90$



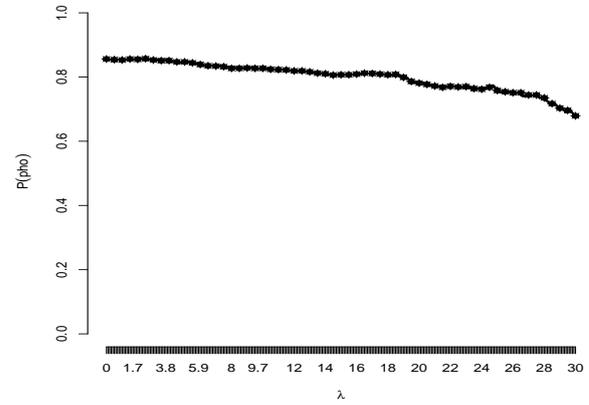
(h)  $P(B=b) - n = 90$

Figura 5.19:  $\lambda \leq 0$  - Cenário degrau positivo - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .

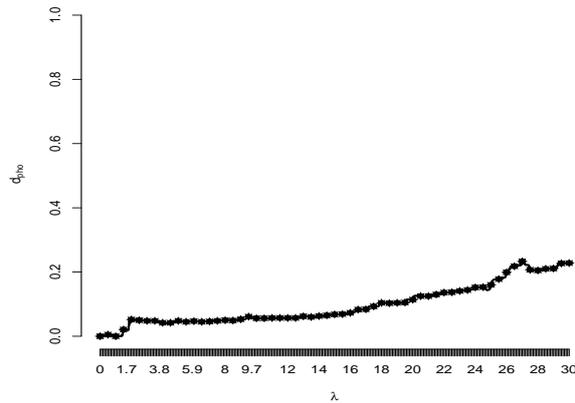
As Figura 5.24 e 5.25 referem-se a partição de interesse no cenário escada crescente.



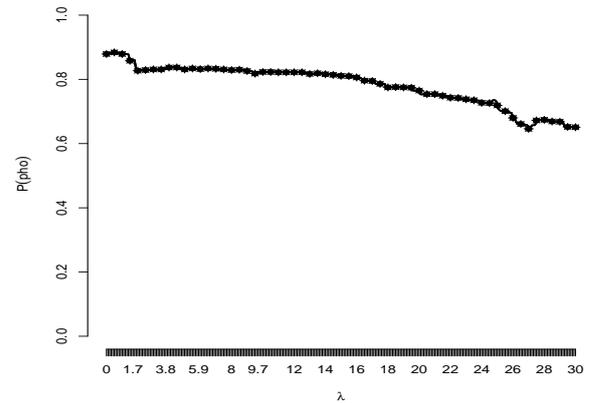
(a)  $d_\rho - n = 30$



(b)  $P(\rho = \rho') - n = 30$



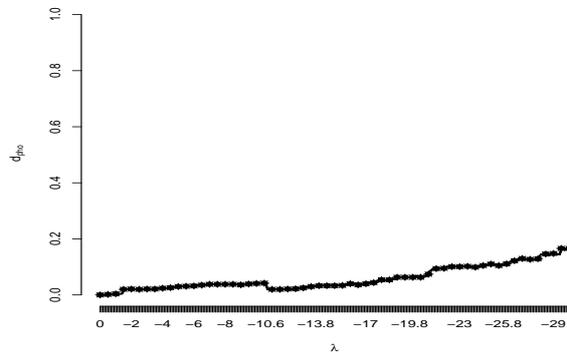
(c)  $d_\rho - n = 90$



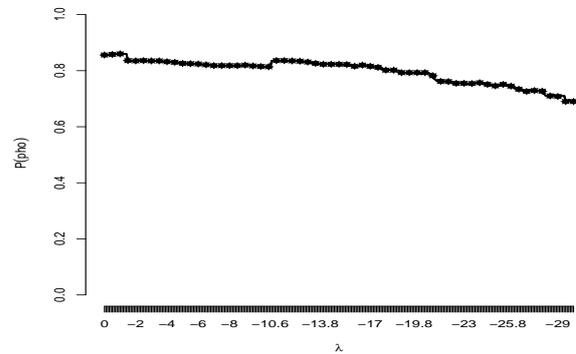
(d)  $P(\rho = \rho') - n = 90$

Figura 5.20:  $\lambda \leq 0$  - Cenário escada crescente - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

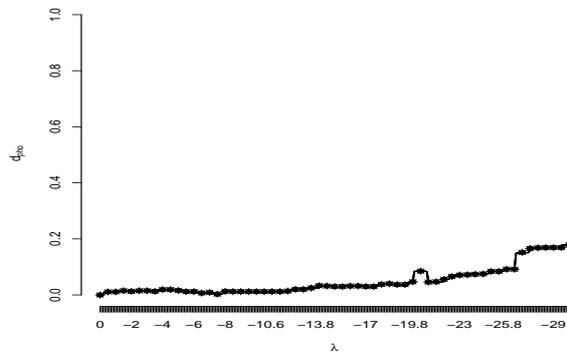
Observa-se que a partição de interesse é identificada com menor probabilidade a medida que  $\lambda$  se afasta de zero. Conseqüentemente, a medida  $d_\rho$  aumenta quando  $\lambda$  se afasta de zero, seja  $\lambda < 0$  ou  $\lambda > 0$ , e isso ocorre independente do tamanho de amostra  $n = 30$  ou  $n = 90$  e ocorre também independente do valor de  $\epsilon$  avaliado.



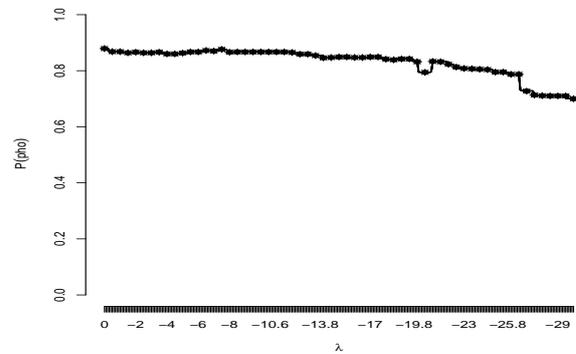
(a)  $d_\rho - n = 30$



(b)  $P(\rho = \rho') - n = 30$



(c)  $d_\rho - n = 90$

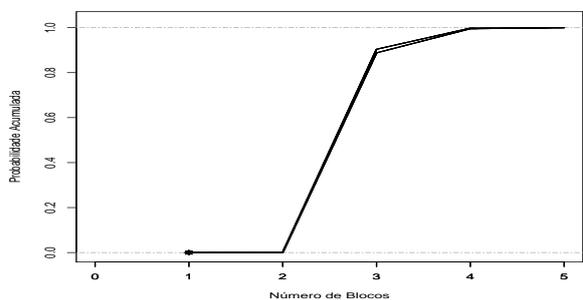


(d)  $P(\rho = \rho') - n = 90$

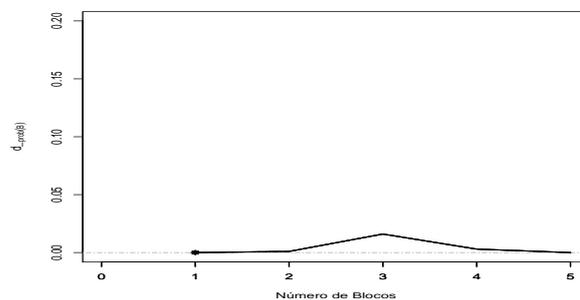
Figura 5.21:  $\lambda \leq 0$  - Cenário escada crescente - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .

Nas Figuras 5.22 e 5.23 são apresentados os resultados para o número de blocos no cenário escada crescente. Nota-se que apesar da partição de interesse ser identificada com menor probabilidade a medida que  $\lambda$  se afasta de zero, o número de blocos é identificado corretamente e com probabilidade alta e próxima para todos os valores de  $\lambda$ , indicando que o modelo fica impreciso em relação ao instante exato da mudança.

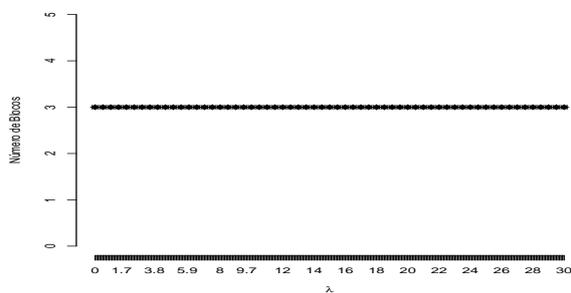
Além disso, a diferença entre as funções de distribuições é pequena e a mesma para todos os valores de  $\epsilon$  considerados neste estudo de robustez, Figuras 5.22 e 5.23 (b) e (f).



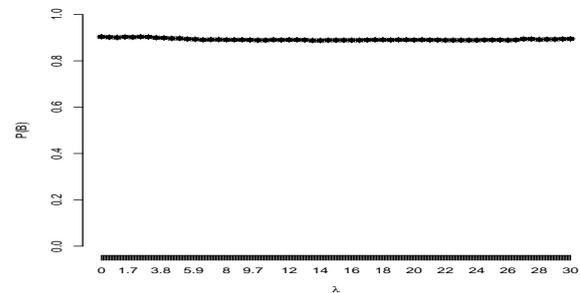
(a)  $P(B \leq b) - n = 30$



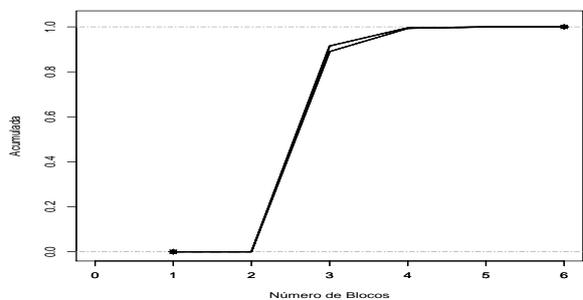
(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 30$



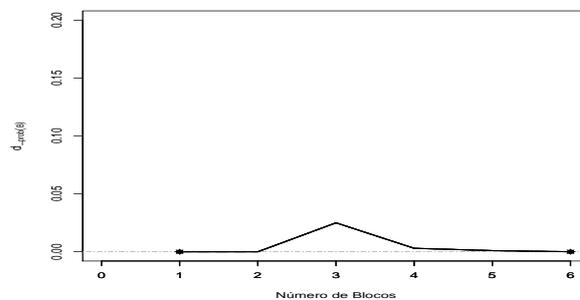
(c) Moda -  $n = 30$



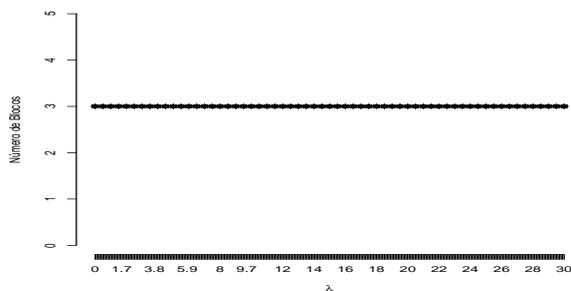
(d)  $P(B=b) - n = 30$



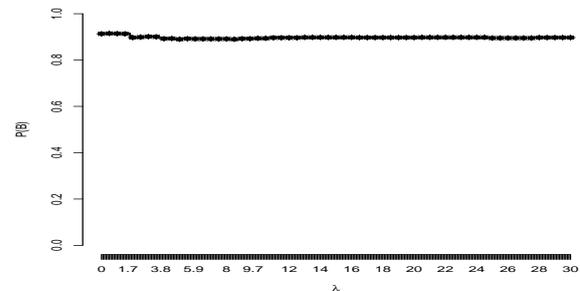
(e)  $P(B \leq b) - n = 90$



(f)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 90$

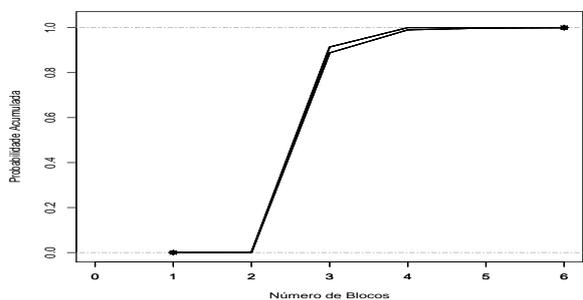


(g) Moda -  $n = 90$

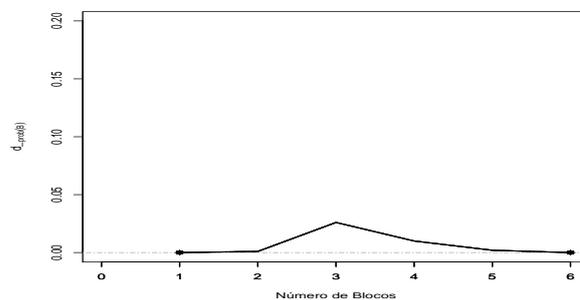


(h)  $P(B=b) - n = 90$

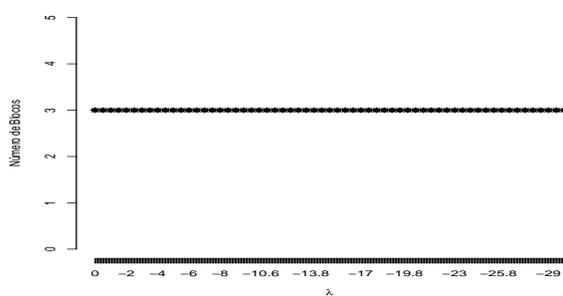
Figura 5.22:  $\lambda \geq 0$  - Cenário escada crescente - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .



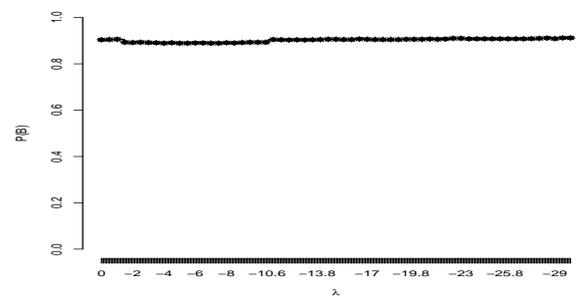
(a)  $P(B \leq b) - n = 30$



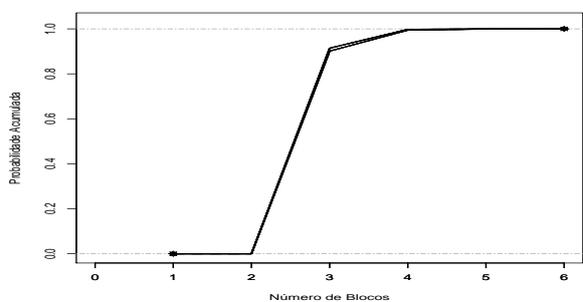
(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 30$



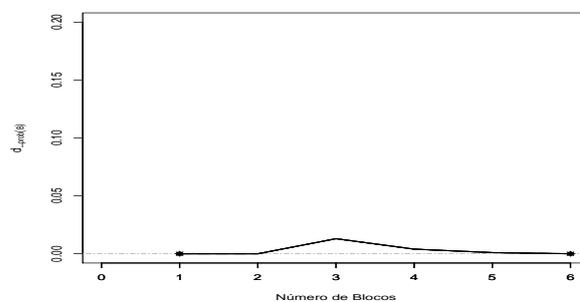
(c) Moda -  $n = 30$



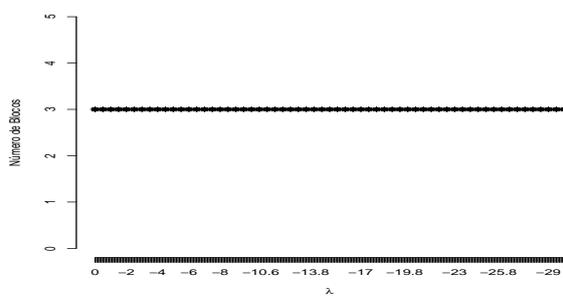
(d)  $P(B=b) - n = 30$



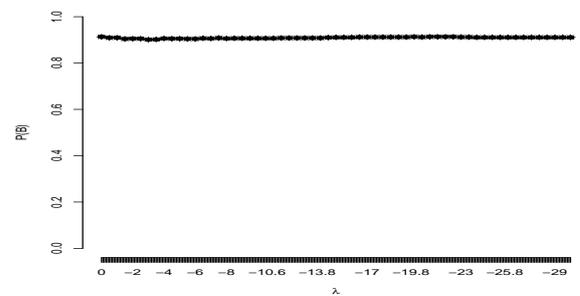
(e)  $P(B \leq b) - n = 90$



(f)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 90$



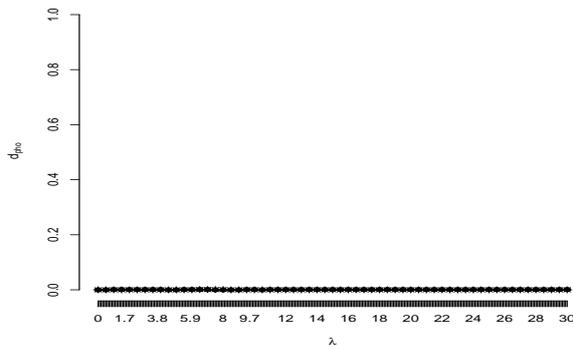
(g) Moda -  $n = 90$



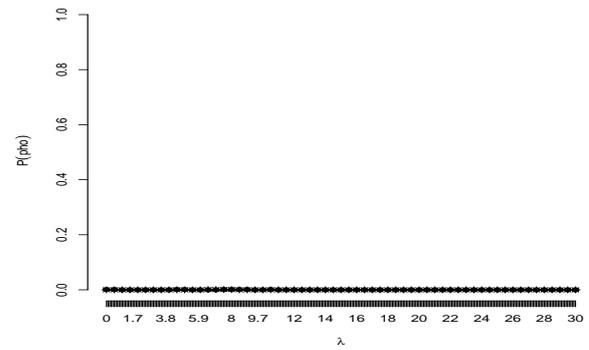
(h)  $P(B=b) - n = 90$

Figura 5.23:  $\lambda \leq 0$  - Cenário escada crescente - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .

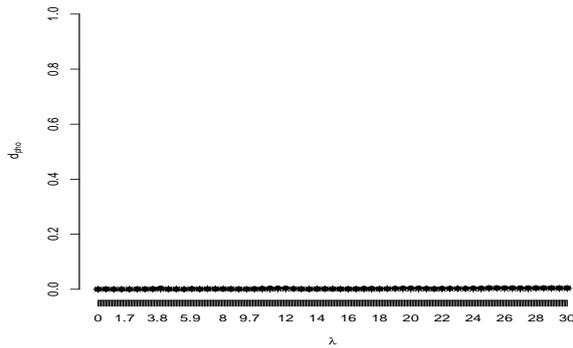
As Figuras 5.24 e 5.25 mostram os resultados da robustez para o cenário epsilon em relação a partição de interesse.



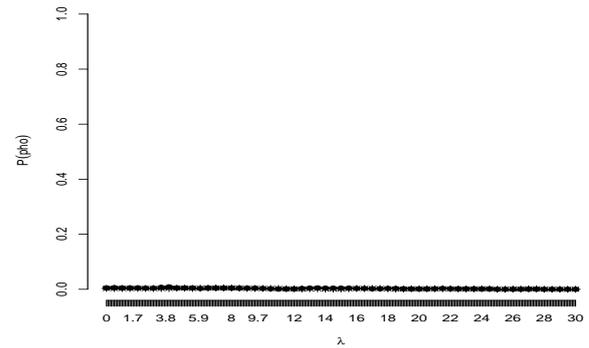
(a)  $d_\rho - n = 30$



(b)  $P(\rho = \rho') - n = 30$



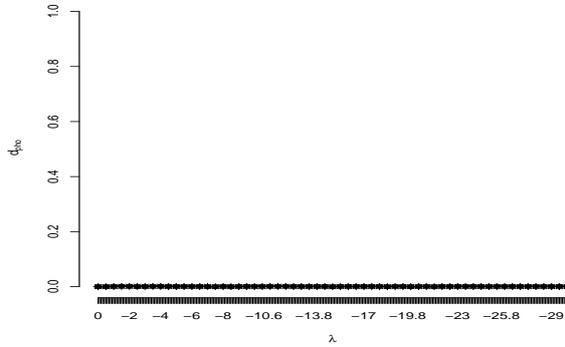
(c)  $d_\rho - n = 90$



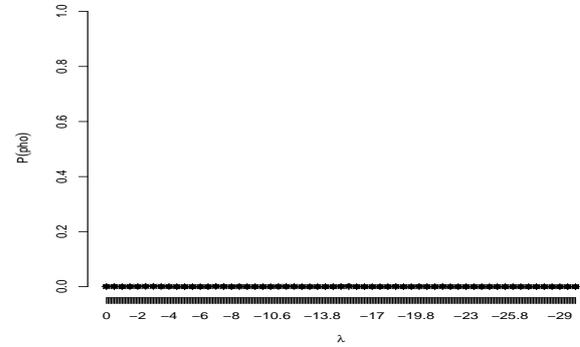
(d)  $P(\rho = \rho') - n = 90$

Figura 5.24:  $\lambda \geq 0$  - Cenário epsilon - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

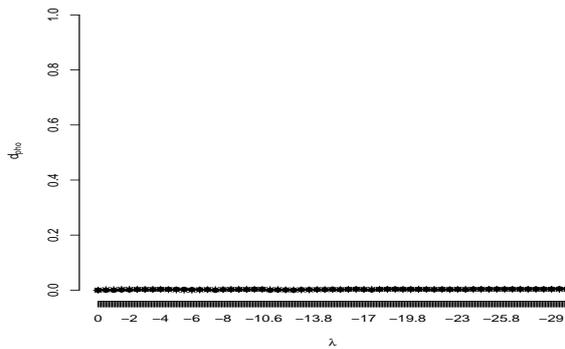
Nota-se que para o cenário epsilon em nenhum momento o modelo foi capaz de identificar corretamente a partição de interesse independente do valor de  $\lambda$ . Mas essa é uma característica do modelo partição produto para esse tipo de cenário, a probabilidade da partição correta ser indicada é muito baixa. Geralmente, as partições estimadas são próximas das reais, mas não são exatamente idênticas, pois a diferença entre os blocos é pequena.



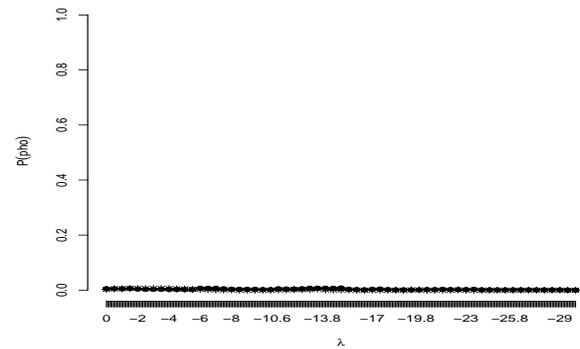
(a)  $d_\rho - n = 30$



(b)  $P(\rho = \rho') - n = 30$



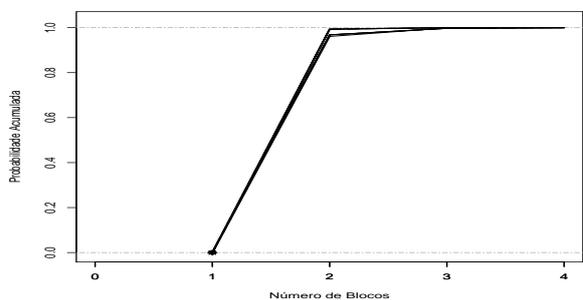
(c)  $d_\rho - n = 90$



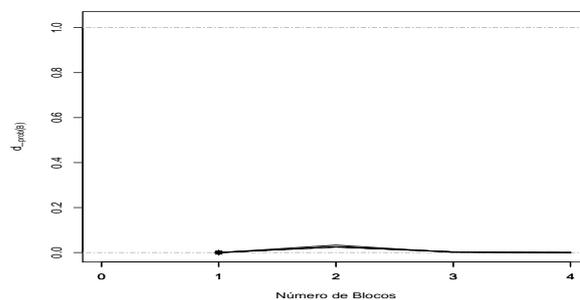
(d)  $P(\rho = \rho') - n = 90$

Figura 5.25:  $\lambda \leq 0$  - Cenário epsilon - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

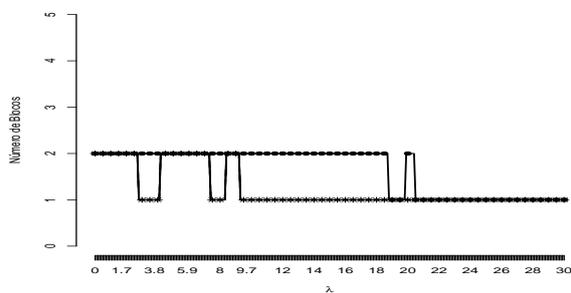
O número de blocos estimado para cada valor do parâmetro de assimetria para o cenário epsilon é ilustrado nas Figuras 5.26 e 5.27. Nota-se que para  $\epsilon = 50$  e  $n = 30$  o modelo teve mais dificuldade em identificar dois blocos a medida que  $\lambda$  se afasta de zero, nos demais casos a partir de  $\lambda = \pm 19$  o modelo tende a indicar apenas um bloco, mas se o tamanho de amostra é  $n = 90$  o número de blocos estimados é 2 para todos os valores de  $\lambda$  avaliados, no entanto, o modelo erra, pois o número correto de blocos é  $B = 3$ .



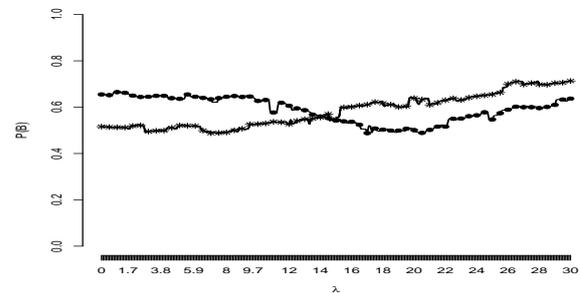
(a)  $P(B \leq b) - n = 30$



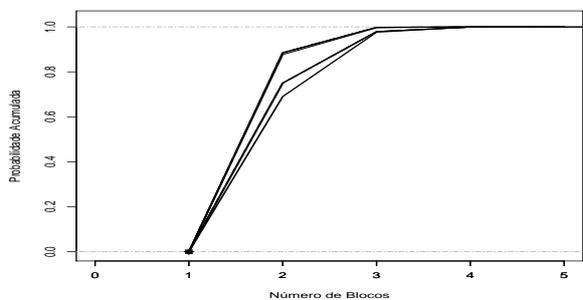
(b)  $|P_{\lambda}(B = b) - P_{\lambda=0}(B = b)| - n = 30$



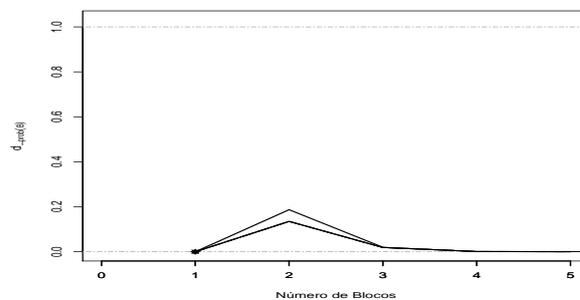
(c) Moda -  $n = 30$



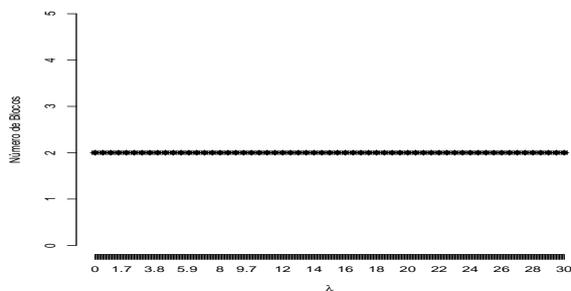
(d)  $P(B=b) - n = 30$



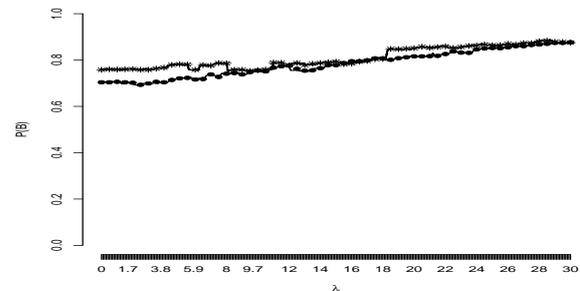
(e)  $P(B \leq b) - n = 90$



(f)  $|P_{\lambda}(B = b) - P_{\lambda=0}(B = b)| - n = 90$

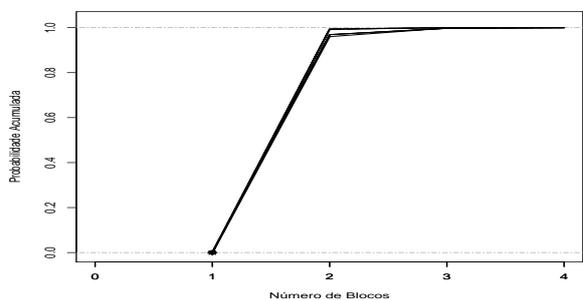


(g) Moda -  $n = 90$

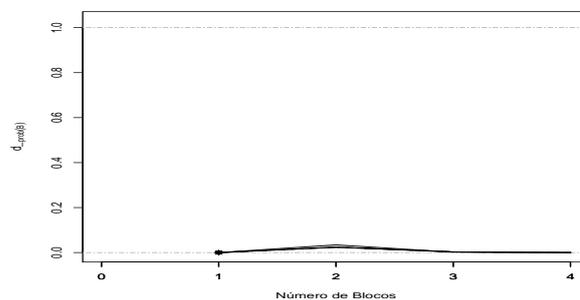


(h)  $P(B=b) - n = 90$

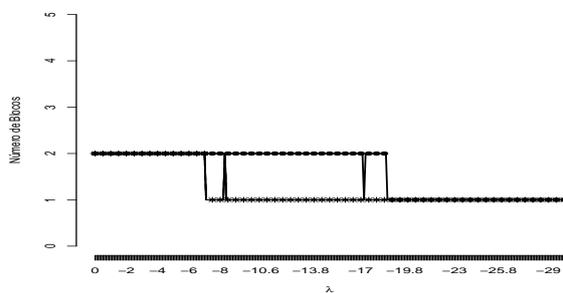
Figura 5.26:  $\lambda \geq 0$  - Cenário epsilon - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com \*  $\epsilon = 50$ .



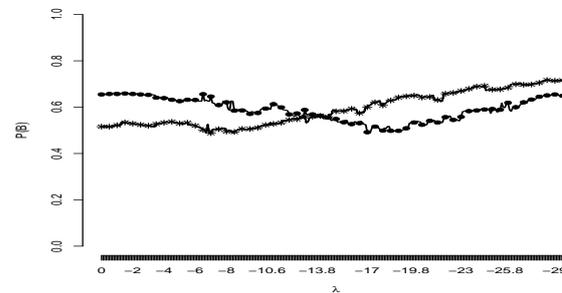
(a)  $P(B \leq b) - n = 30$



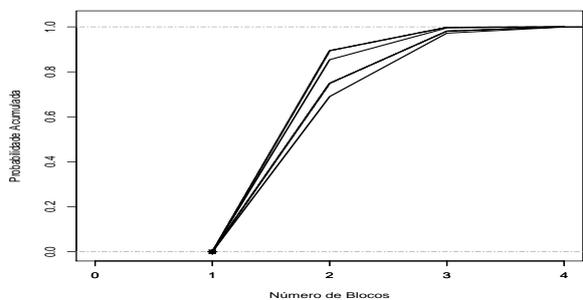
(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 30$



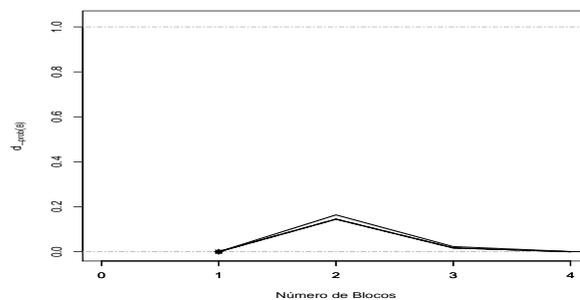
(c) Moda -  $n = 30$



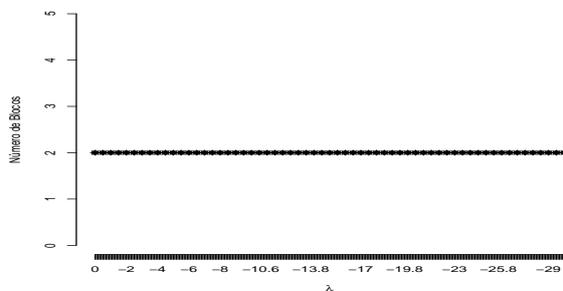
(d)  $P(B=b) - n = 30$



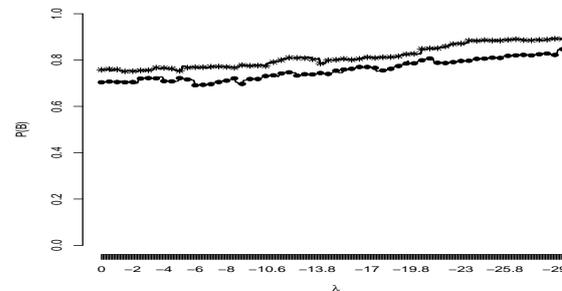
(e)  $P(B \leq b) - n = 90$



(f)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 90$



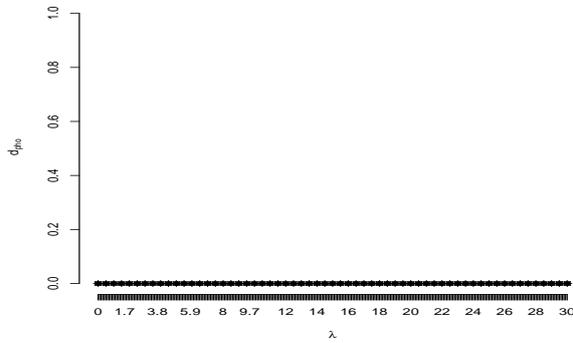
(g) Moda -  $n = 90$



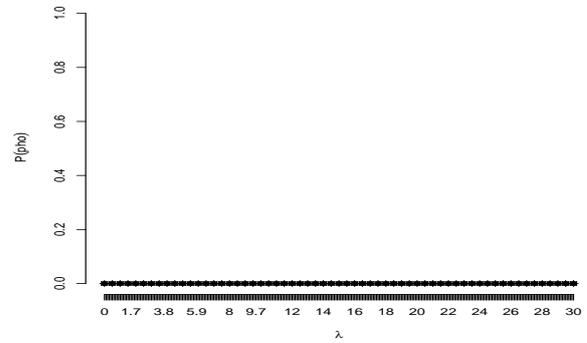
(h)  $P(B=b) - n = 90$

Figura 5.27:  $\lambda \leq 0$  - Cenário epsilon - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com \*  $\epsilon = 50$ .

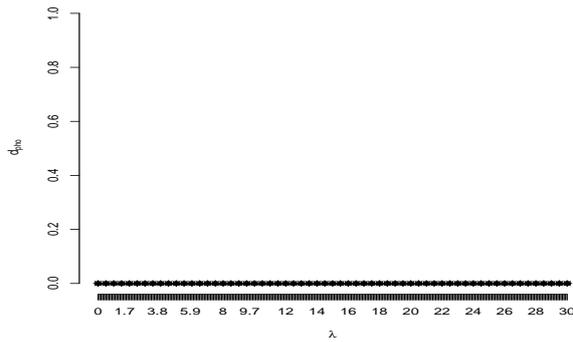
As Figuras 5.28 e 5.29 ilustram os resultados para a partição de interesse no cenário tendência.



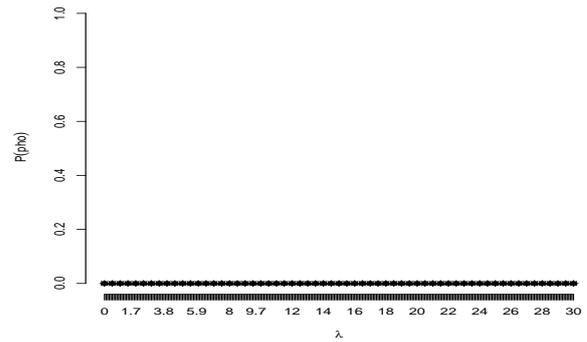
(a)  $d_\rho - n = 40$



(b)  $P(\rho = \rho') - n = 40$



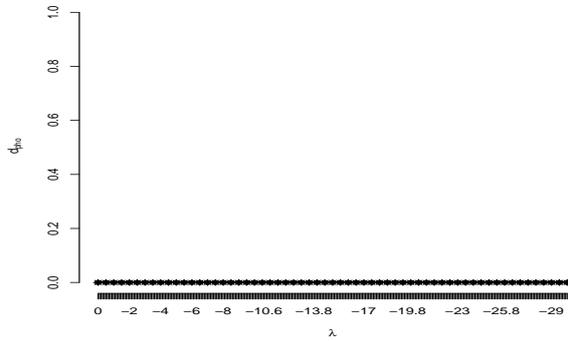
(c)  $d_\rho - n = 120$



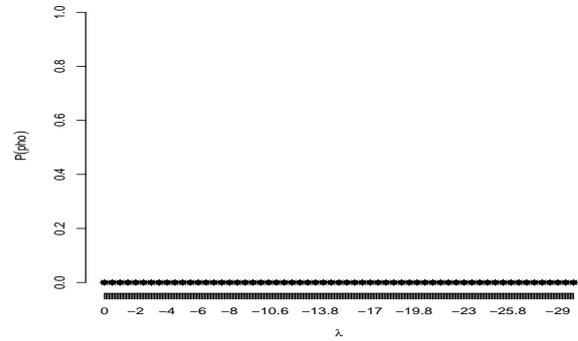
(d)  $P(\rho = \rho') - n = 120$

Figura 5.28:  $\lambda \geq 0$  - Cenário tendência - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

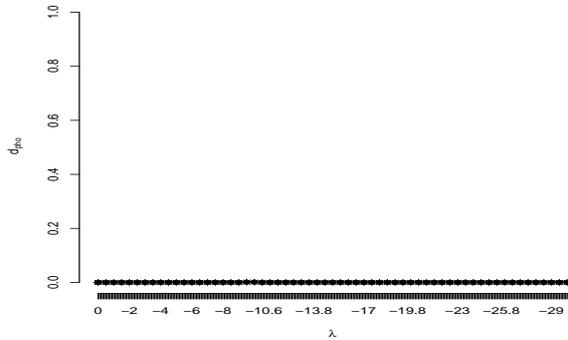
Nota-se que no cenário tendência o modelo também não foi capaz de identificar a partição correta para nenhum valor de  $\lambda$  avaliado, mas nesse tipo de cenário o modelo partição produto também deixa a desejar ao identificar a partição correta, como pode ser observado, mesmo  $\lambda = 0$  não foi capaz de identificar a partição correta.



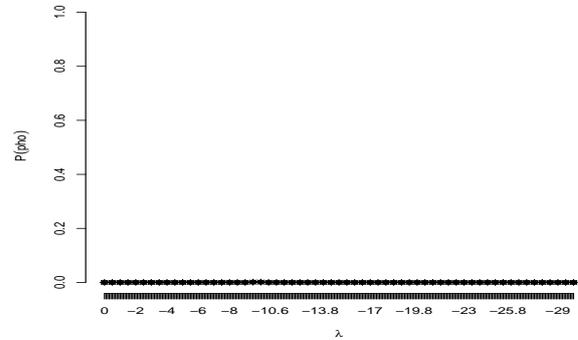
(a)  $d_\rho - n = 40$



(b)  $P(\rho = \rho') - n = 40$



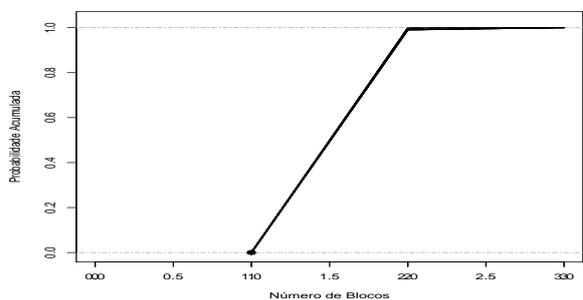
(c)  $d_\rho - n = 120$



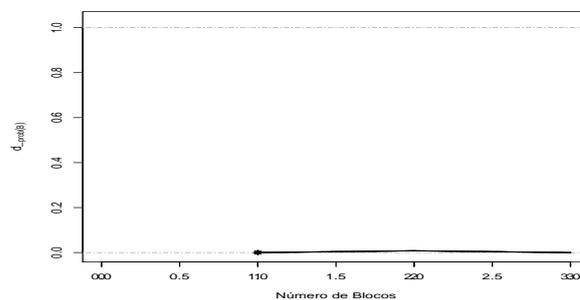
(d)  $P(\rho = \rho') - n = 120$

Figura 5.29:  $\lambda \leq 0$  - Cenário tendência - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

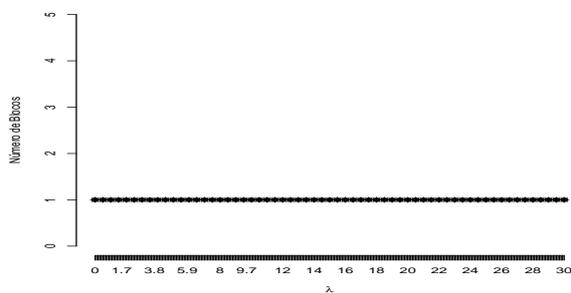
Os resultados de robustez para o cenário tendência em relação ao número de blocos são apresentados nas Figuras 5.30 e 5.31, e representam os resultados para  $\lambda$  positivo e negativo, respectivamente. Nota-se, assim como no cenário epsilon, que o modelo não foi capaz de identificar o número correto de blocos e para  $n = 40$  o modelo tende a indicar um único bloco independente do valor do parâmetro de assimetria. Além disso, a maior variação das funções de distribuição para o número de blocos ocorre para  $\lambda$  positivo e tamanho de amostra 120, nesse caso o valor de  $\epsilon$  fez com que a diferença na probabilidade de indicar  $B = 2$  variasse cerca de 65% para  $\epsilon = 0$ .



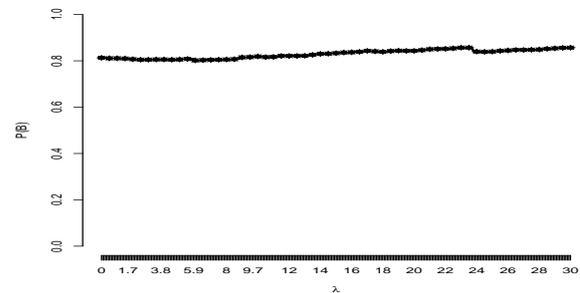
(a)  $P(B \leq b) - n = 40$



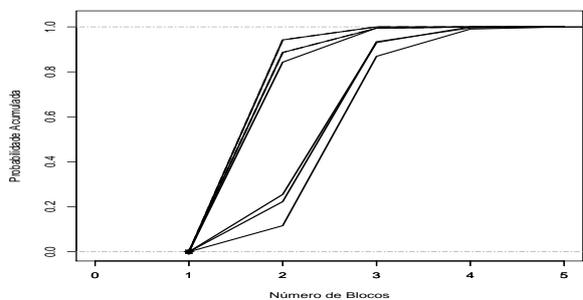
(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 40$



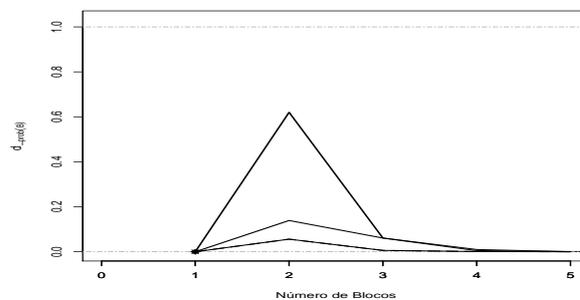
(c) Moda -  $n = 40$



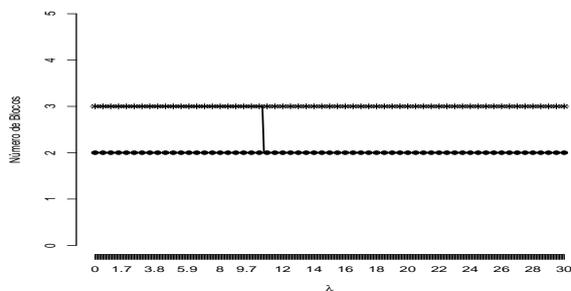
(d)  $P(B=b) - n = 40$



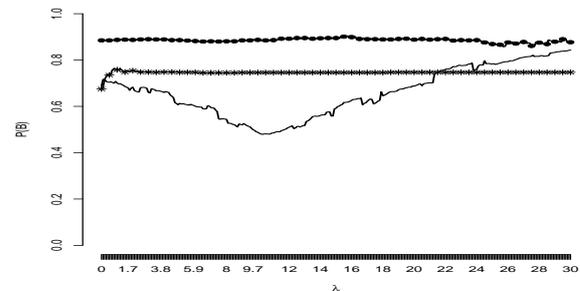
(e)  $P(B \leq b) - n = 120$



(f)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 120$

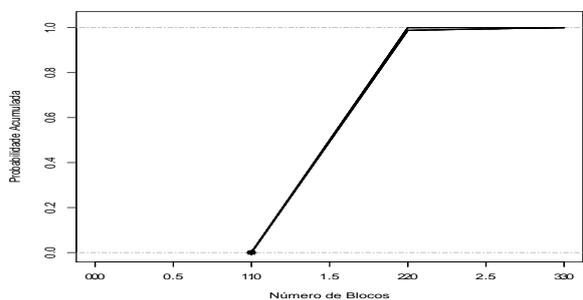


(g) Moda -  $n = 120$

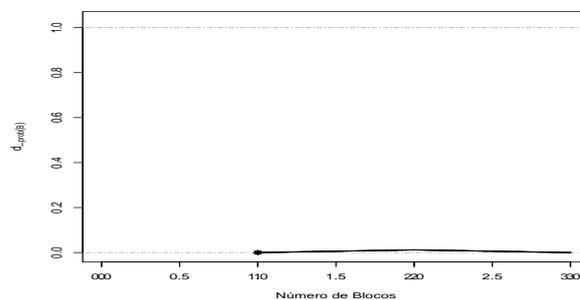


(h)  $P(B=b) - n = 120$

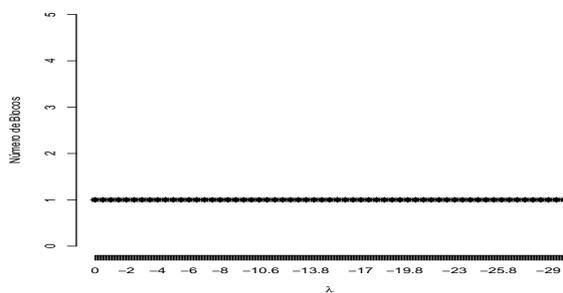
Figura 5.30:  $\lambda \geq 0$  - Cenário tendência - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com \*  $\epsilon = 50$ .



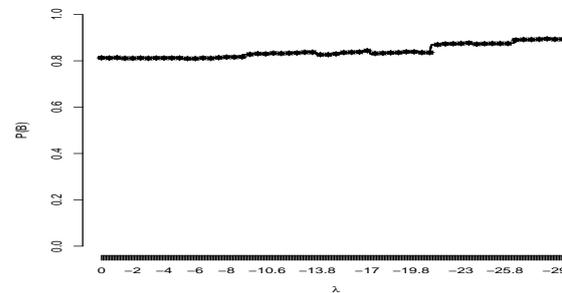
(a)  $P(B \leq b) - n = 40$



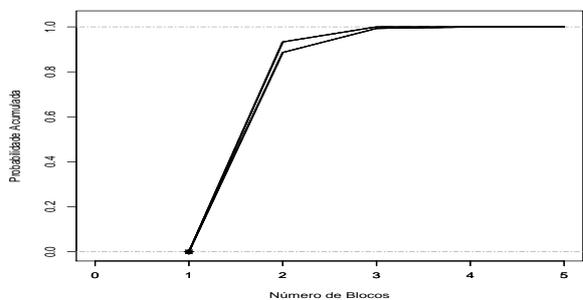
(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 40$



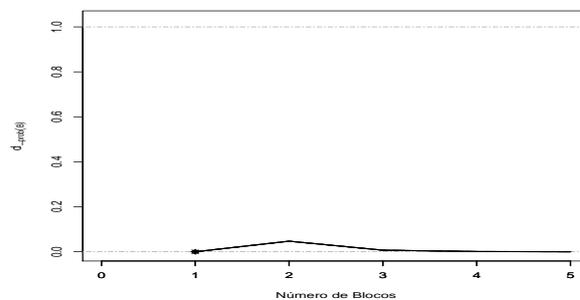
(c) Moda -  $n = 40$



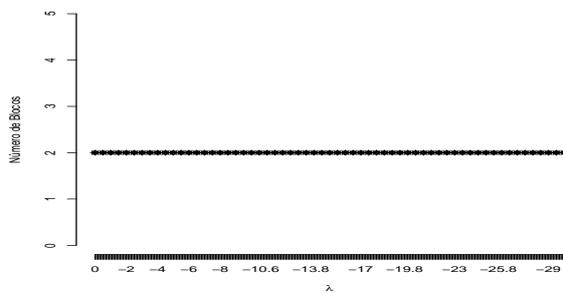
(d)  $P(B=b) - n = 40$



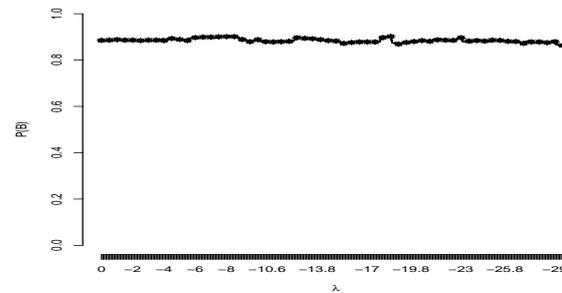
(e)  $P(B \leq b) - n = 120$



(f)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)| - n = 120$



(g) Moda -  $n = 120$



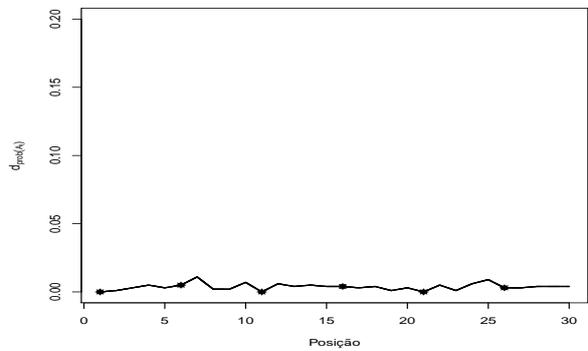
(h)  $P(B=b) - n = 120$

Figura 5.31:  $\lambda \leq 0$  - Cenário tendência - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com \*  $\epsilon = 50$ .

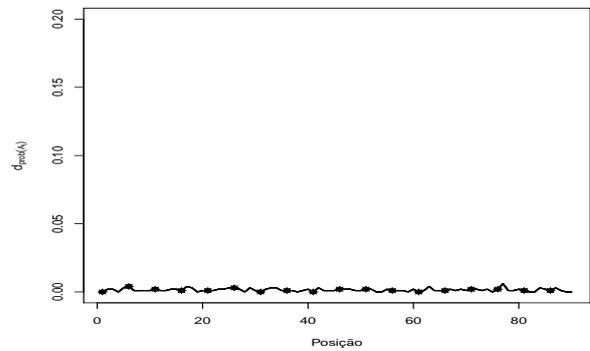
A próxima seção mostra os resultados da robustez no modelo partição produto para os instante em que ocorrem as mudanças.

### 5.1.3 Robustez na probabilidade de mudança em cada instante

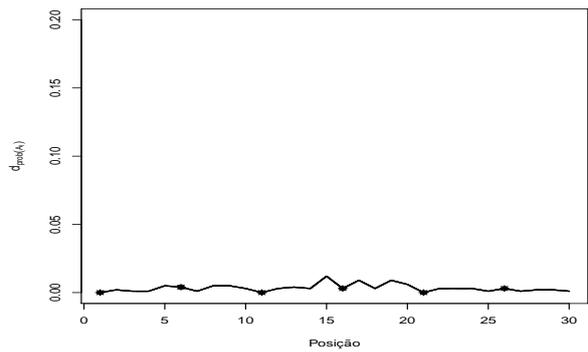
Nesta seção encontram-se os resultados da robustez global para a probabilidade de mudança em cada instante, os gráficos apresentados aqui reference a medida  $d_{P(A_t)}$  definida na Seção 4.2. A Figura 5.32 fornece os resultados da robustez de  $P(A_t)$  no cenário degrau não simétrico.



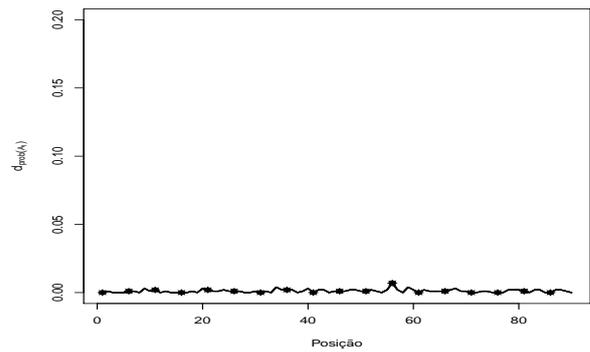
(a)  $d_{P(A_t)}$  com  $\lambda \leq 0$  e  $n = 30$



(b)  $d_{P(A_t)}$  com  $\lambda \leq 0$  e  $n = 90$



(c)  $d_{P(A_t)}$  com  $\lambda \geq 0$  e  $n = 30$

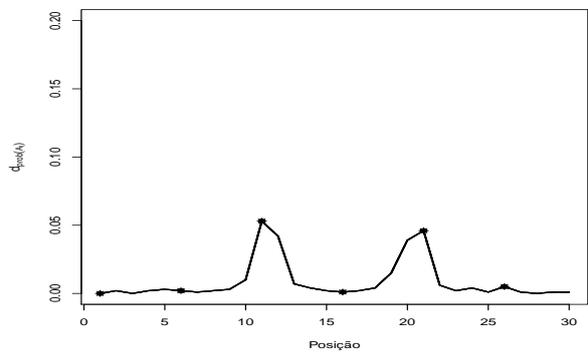


(d)  $d_{P(A_t)}$  com  $\lambda \geq 0$  e  $n = 90$

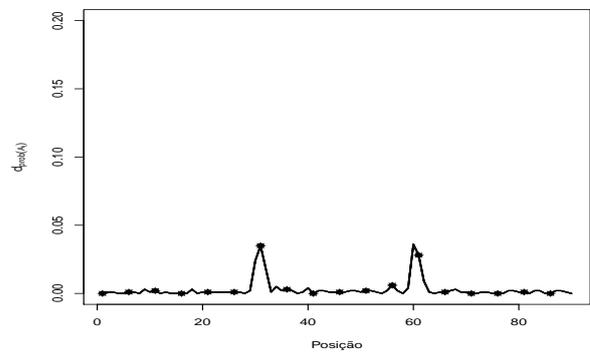
Figura 5.32: Cenário degrau não simétrico - Magnitude da medida  $d_{P(A_t)}$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

A diferença  $d_{P(A_I)}$  apresentada na Figura 5.32, referente ao cenário degrau não simétrico, foi praticamente nula, mesmo próxima aos pontos de mudança, repetindo o comportamento observado para as distâncias  $d_M$  e  $d_{var}$  na Seção 5.1.1.

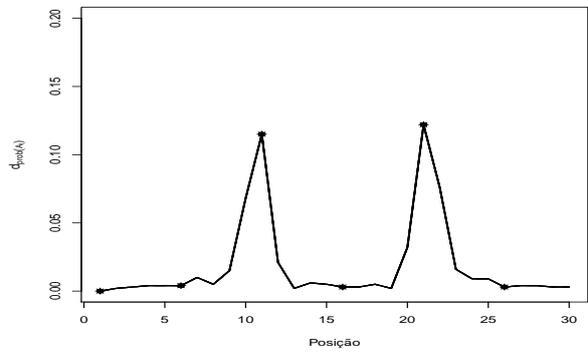
A Figura 5.33 ilustra os resultados da medida de robustez da probabilidade de mudança em cada instante para o cenário degrau positivo.



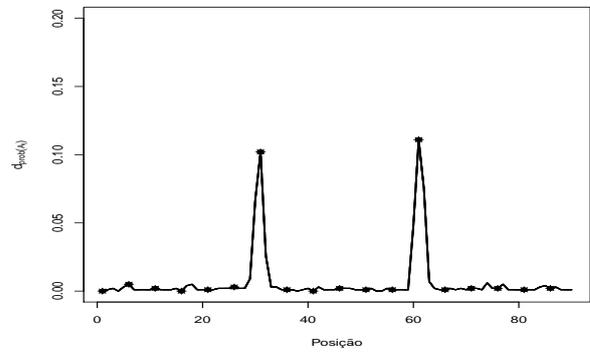
(a)  $d_{P(A_I)}$  com  $\lambda \geq 0$  e  $n = 30$



(b)  $d_{P(A_I)}$  com  $\lambda \geq 0$  e  $n = 90$



(c)  $d_{P(A_I)}$  com  $\lambda \leq 0$  e  $n = 30$



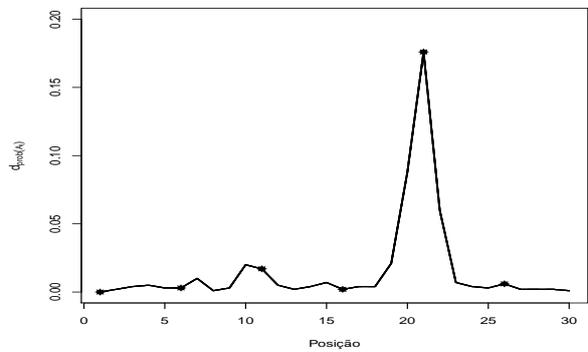
(d)  $d_{P(A_I)}$  com  $\lambda \leq 0$  e  $n = 90$

Figura 5.33: Cenário degrau positivo - Magnitude da medida  $d_{P(A_I)}$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

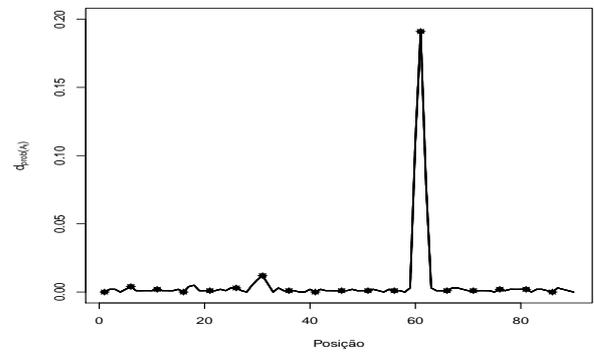
Nota-se na Figura 5.33(c) e (d) que a medida  $d_{P(A_I)}$  apresenta maiores variações próximo aos pontos de mudança para  $\lambda < 0$  se comparado a  $\lambda > 0$ . Além disso, a magnitude da medida  $d_{P(A_I)}$

é ligeiramente maior para  $n = 30$  se comparada a  $n = 90$ . Observa-se também que as diferenças são as mesmas independente da informação amostral, pois para todo  $\epsilon$  as medidas foram as mesmas.

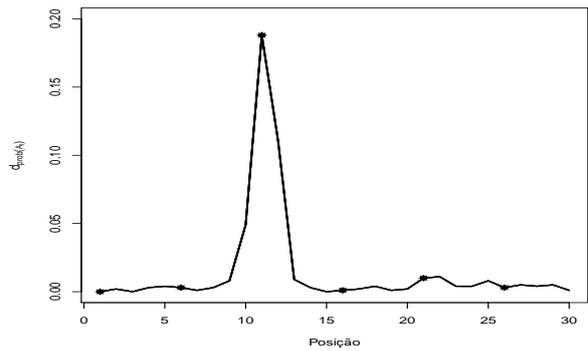
Os resultados para o cenário escada crescente estão na Figura 5.34. Observa-se, nesse cenário, que a medida  $d_{P(A_l)}$  chega a apresentar uma variação próxima a 20% em torno do primeiro ponto de mudança se  $\lambda > 0$  e se  $\lambda < 0$  essa variação ocorre próxima ao segundo ponto de mudança. Mais uma vez o valor da média dos blocos não influenciou na medida  $d_{P(A_l)}$ , pois para todo valor de  $\epsilon$  as distâncias foram as mesmas.



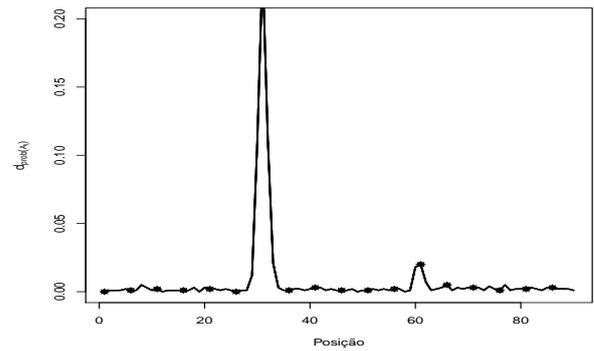
(a)  $d_{P(A_l)}$  com  $\lambda \leq 0$  e  $n = 30$



(b)  $d_{P(A_l)}$  com  $\lambda \leq 0$  e  $n = 90$



(c)  $d_{P(A_l)}$  com  $\lambda \geq 0$  e  $n = 30$



(d)  $d_{P(A_l)}$  com  $\lambda \geq 0$  e  $n = 90$

Figura 5.34: Cenário escada crescente - Magnitude da medida  $d_{P(A_l)}$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .

Os resultados com respeito a magnitude da medida  $d_{P(A_I)}$  para o cenário epsilon são ilustrados na Figura 5.35.

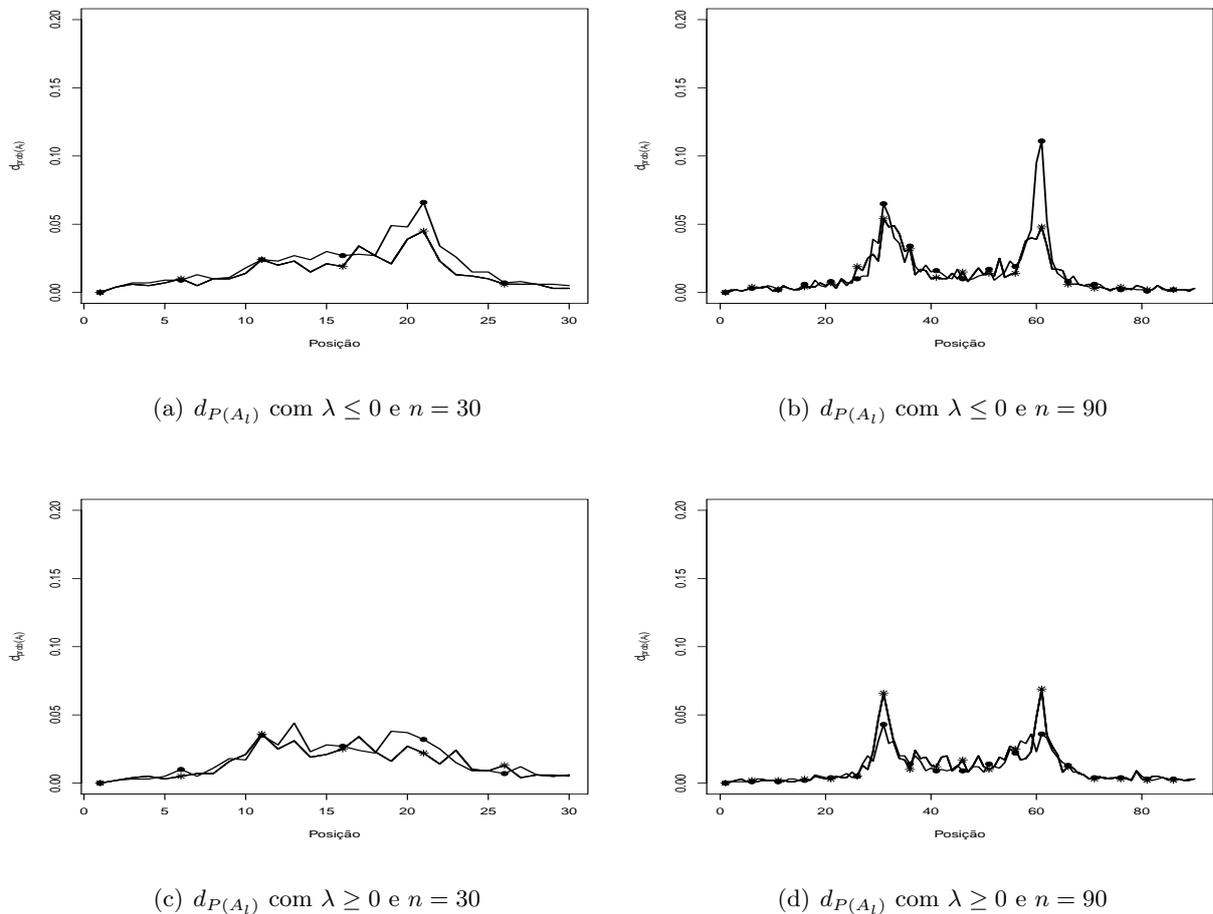


Figura 5.35: Cenário epsilon - Magnitude da medida  $d_{P(A_I)}$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

Observa-se na Figura 5.35 que a magnitude da diferença de  $d_{P(A_I)}$  é inferior a 0.06 para  $n = 30$  e para  $n = 90$  essa diferença é um pouco maior que 0.10, no entanto, como foi visto na seção anterior, nem o número de blocos nem a partição de interesse são estimados corretamente, logo a conclusão que se chega é que os valores de  $\lambda$  tendem a estimar os mesmos pontos de mudança equivocada.

A Figura 5.36 apresenta os resultados da probabilidade de mudança em cada instante para o cenário

tendência.

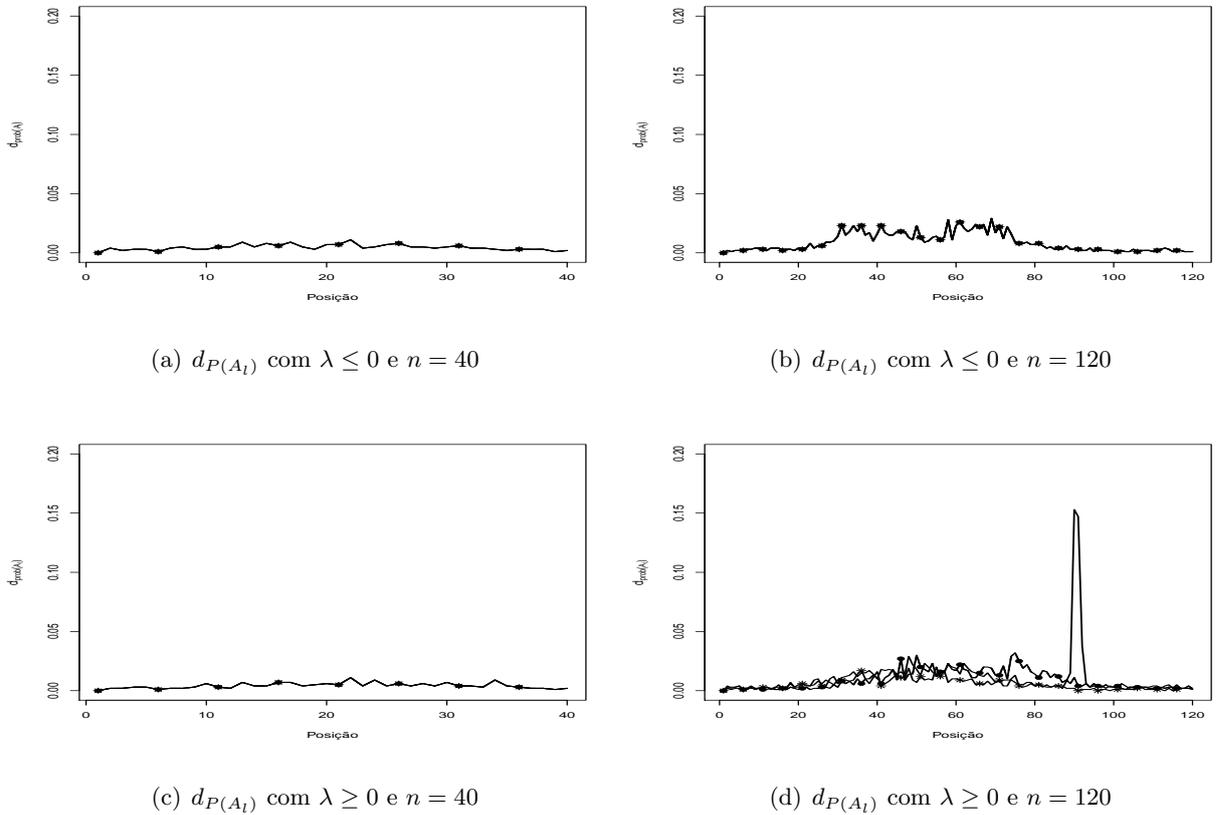


Figura 5.36: Cenário tendência - Magnitude da medida  $d_{P(A_t)}$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

Nota-se no cenário tendência que a diferença na estimativa da probabilidade de mudança para os diversos valores de  $\lambda$  são próximas a zero para  $n = 40$ , e para  $n = 120$  essa diferença é próxima de 0.04 com exceção do ponto 91 quando  $\lambda$  é positivo.

A próxima seção faz um resumo dos principais resultados observados para o estudo de robustez levando em consideração a contaminação de forma multiplicativa na distribuição *a priori*.

#### 5.1.4 Considerações do estudo de robustez global no Modelo I

De forma geral, observa-se que as distâncias  $d_M$  e  $d_{var}$  para 30 observações são ligeiramente maiores que nos casos onde o tamanho de amostra é 90, indicando que o número de observações influencia nas estimativas, no sentido que aumentando o tamanho da amostra as diferenças tendem a diminuir.

Os gráficos apresentados para as medidas  $d_M$  e  $d_{var}$  não possibilitam indicar se a distância é maior a medida que  $\lambda$  se afasta de zero, logo não permite indicar em que sentido a diferença é maior, mas permite visualizar que essa diferença em nenhum momento foi superior a 0.6.

Nota-se que a distância  $d_{var}$ , para todos os cenários avaliados, apresenta diferenças maiores próximo aos pontos de mudança. Nessa mesma direção as diferenças  $d_M$  tendem a ser levemente alteradas próximos aos pontos de mudança, com exceção do cenário degrau não simétrico  $(-3, 7, 0)$ , cuja a magnitude da diferença entre os blocos é maior.

Os gráficos com a medida de robustez para a probabilidade de mudança em cada instante apresentam resultados similares aos encontrados para as medida  $d_M$  e  $d_{var}$ , ou seja, as maiores diferenças foram observadas próximo aos pontos de mudança e os cenários que apresentam as maiores diferenças são os cenários epsilon e tendência.

Quando foi acrescido a cada bloco os valores 2, 5, 50 o comportamento das distâncias não foi alterado nos cenários escada e degrau, indicando que o modelo é robusto a translações constantes nos dados e, portanto, às as estimativas não se alteram levando em consideração os valores de  $\lambda$  avaliados. Mas no caso dos cenários tendência e epsilon a quantidade  $\epsilon$  adicionada influencia.

Os cenários escada e degrau indicam que o modelo identifica corretamente o número de blocos ( $B = 3$ ) e a diferença entre as funções de distribuição é pequena, mesmo quando a direção da assimetria é oposta ao valor das observações. Mas, o número de blocos é identificado corretamente com maior probabilidade para  $\lambda = 0$  em todos os cenários independente se o tamanho da amostra é 30 ou 90. No caso dos cenários tendência e epsilon o número de blocos é subestimado.

Em relação a probabilidade do modelo identificar a partição corretamente verifica-se que para ambos os tamanhos de amostra  $n = 30$  e  $n = 90$ , os melhores resultados ocorrem quando  $\lambda = 0$  em todos os cenários, com exceção do cenário tendência, independente se o sinal dos dados concorda ou não com o parâmetro de assimetria. E no caso do cenário epsilon a partição de interesse não foi identificada corretamente.

Os gráficos com o número de blocos apresentam um comportamento similar aos gráficos que retratam a partição, como era esperado, pois se a partição é identificada corretamente, consequentemente, o número de blocos é identificado corretamente. Além disso, nos cenários escada e degrau a probabilidade da moda *a posteriori* ser 3 foi indicada com uma probabilidade superior a 0.85 para todos os valores de  $\lambda$  avaliados.

A próxima seção é dedicada ao estudo das mesmas medidas de robustez discutidas aqui, no entanto assumindo agora que a contaminação ocorre na distribuição amostral.

## 5.2 Robustez no Modelo II

Os resultados apresentados nessa seção referem-se as medidas de robustez global definidas no Capítulo 4 para o contexto do Modelo II, onde a função de verossimilhança pertence a classe de distribuições normais assimétricas.

Serão avaliados nesse estudo os mesmos cenários do Quadro 5.1, degrau positivo, degrau não simétrico, escada crescente, epsilon e tendência. No entanto, para avaliar a robustez global, por questões de custo computacional, foi fixada uma grade de  $\lambda$  nos intervalos  $[-10, 10]$ , ou seja,  $(\lambda = 0, \pm 0.1, \pm 0.2, \pm 0.3, \pm 0.4, \pm 0.5, \dots, \pm 9.8, \pm 9.9, \pm 10)$  e foram considerados os mesmos tamanhos amostrais ( $n = 30, 90$ ). Nesta seção, as médias dos blocos também foram acrescidas da quantidade  $\epsilon$  que assume os valores 2, 5 e 50.

O estudo MCMC foi realizado assumindo uma cadeia de tamanho 60000 iterações e o período de *burn-in* foi de 50000 iterações com um lag de 10 observações, logo o tamanho da amostra *a poste-*

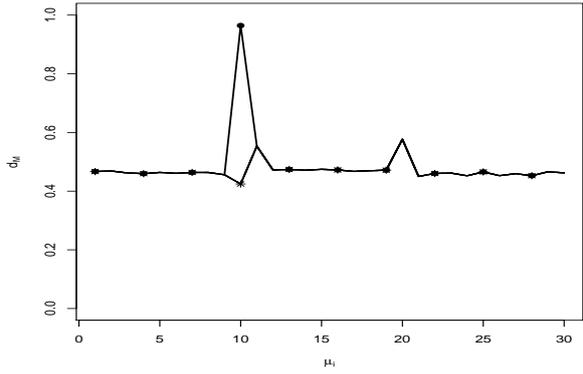
*riori* final foi de 1000 observações.

No estudo de robustez nos dados, ao contrário do estudo de sensibilidade, os dados não foram gerados de uma distribuição normal assimétrica, pois nesse caso originaria varias combinações de resultados, o que tornaria inviável do ponto de vista computacional.

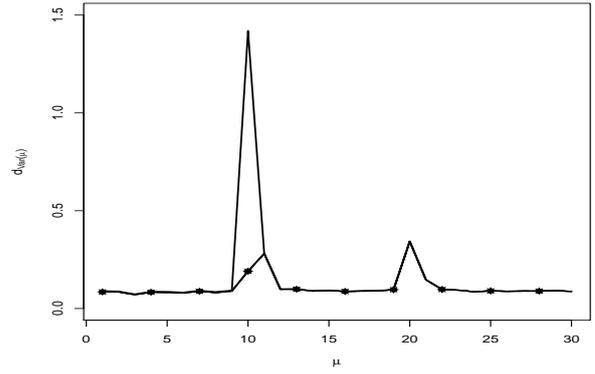
### **5.2.1 Robustez nas estimativas produto e variância *a posteriori* de $\mu_i$**

Os resultados desta seção mostram o comportamento das medidas  $d_M$  e  $d_{var}$  assumindo valores de  $\lambda$  positivos e negativos para a contaminação na distribuição amostral.

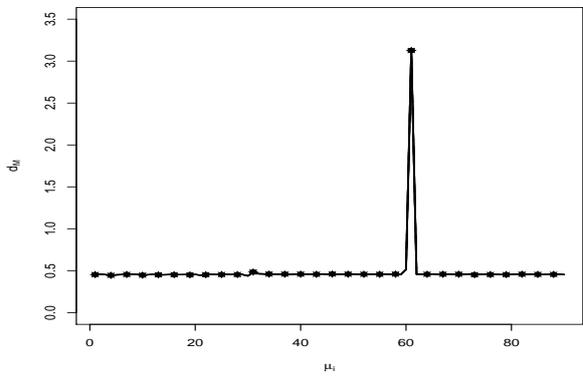
As Figuras 5.37 e 5.38 apresentam os resultados para as medidas  $d_M$  e  $d_{var}$  considerando o cenário degrau positivo.



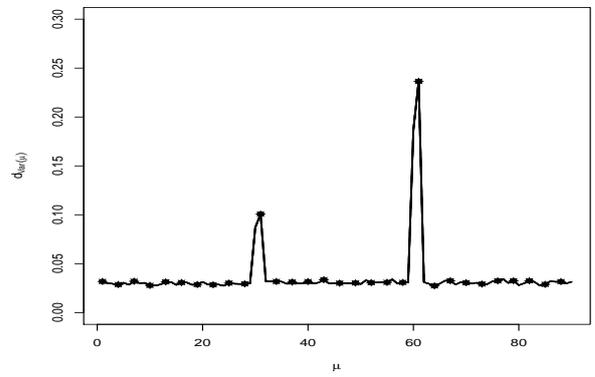
(a)  $\lambda \leq 0 - n = 30$



(b)  $\lambda \leq 0 - n = 30$

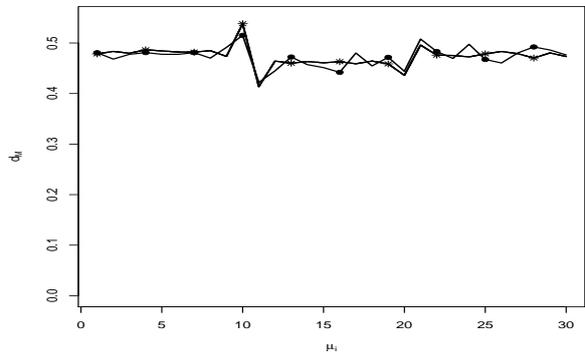


(c)  $\lambda \leq 0 - n = 90$

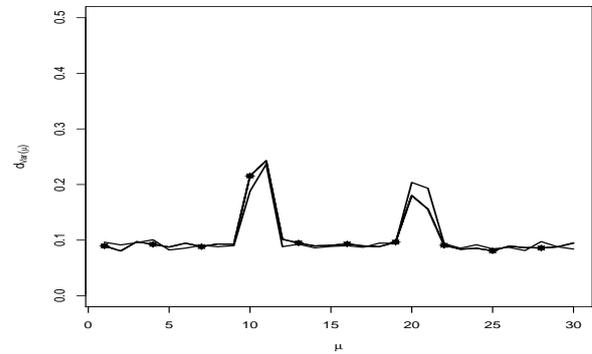


(d)  $\lambda \leq 0 - n = 90$

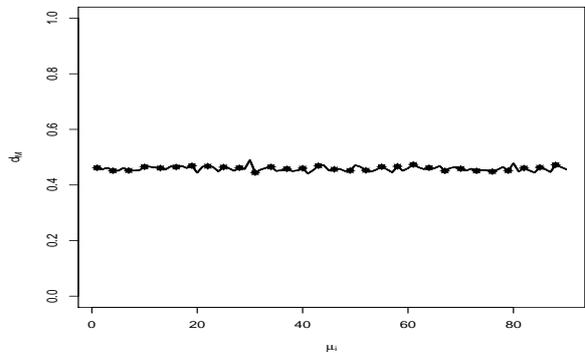
Figura 5.37: Cenário degrau positivo - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .



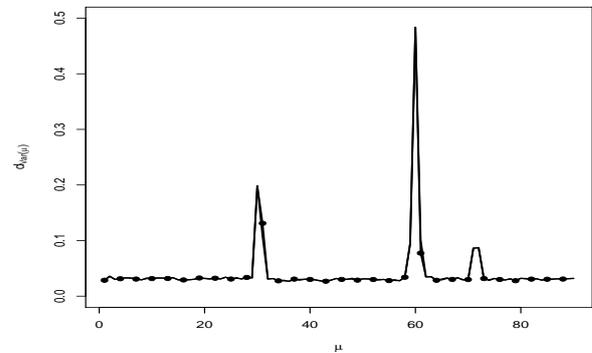
(a)  $\lambda \geq 0 - n = 30$



(b)  $\lambda \geq 0 - n = 30$



(c)  $\lambda \geq 0 - n = 90$

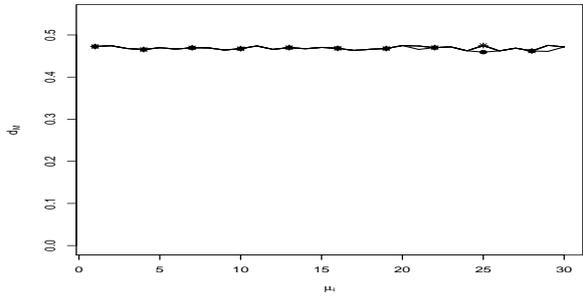


(d)  $\lambda \geq 0 - n = 90$

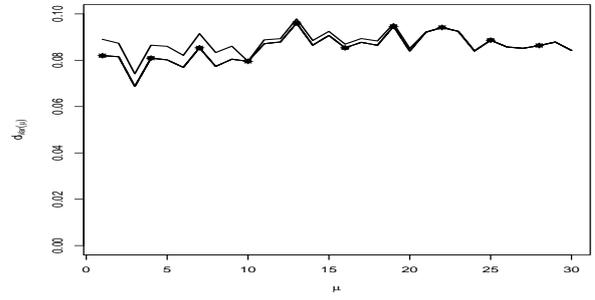
Figura 5.38: Cenário degrau positivo - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

Observa-se na Figura 5.37 (c) que para  $\lambda < 0$  a medida  $d_M$  foi próxima de 3 no instante 61, que representa um ponto de mudança. Ao comparar as Figuras 5.37 e 5.38 nota-se que para  $n = 90$  as medidas  $d_M$  e  $d_{var}$  estão próximas em relação aos valores de  $\epsilon$  avaliados, mas para  $n = 30$  e  $\lambda < 0$  há um pico na medida  $d_{var}$  para  $\epsilon = 0$  e um pico para  $d_M$  quando  $\epsilon = 2$ . Indicando que o modelo pode apresentar estimativas errôneas próximo aos pontos de mudança dependendo dos valores assumidos pela série.

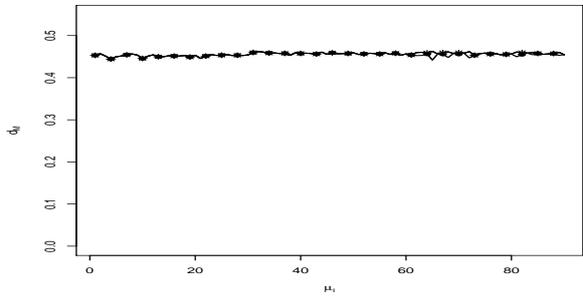
As medidas  $d_M$  e  $d_{var}$  para o cenário degrau não simétrico são apresentadas na Figura 5.39.



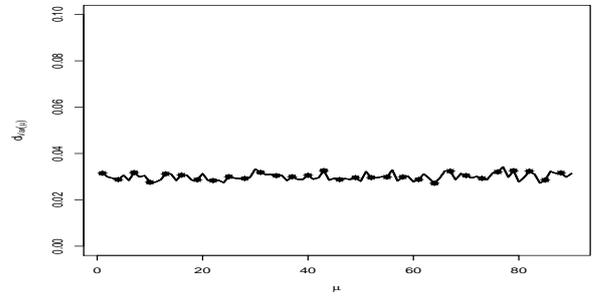
(a)  $\lambda \leq 0 - n = 30$



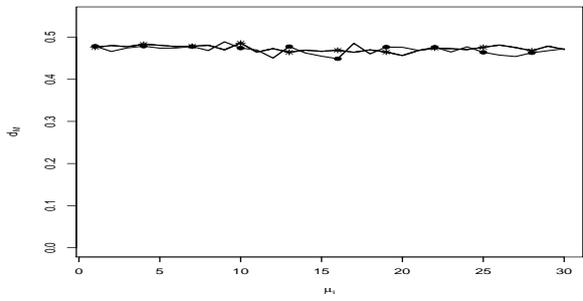
(b)  $\lambda \leq 0 - n = 30$



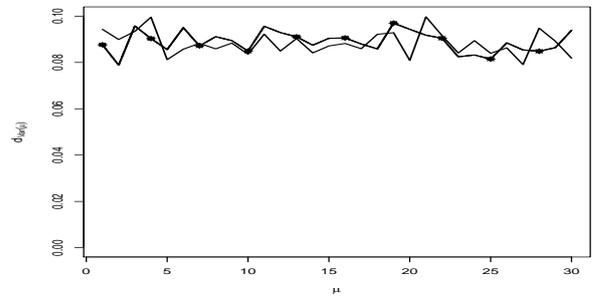
(c)  $\lambda \leq 0 - n = 90$



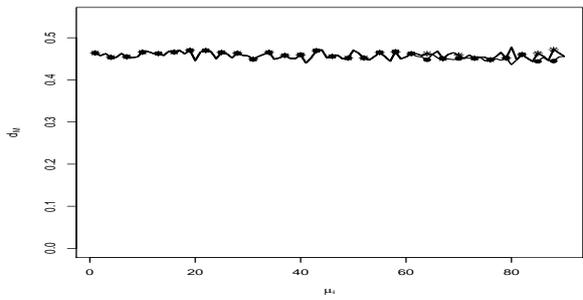
(d)  $\lambda \leq 0 - n = 90$



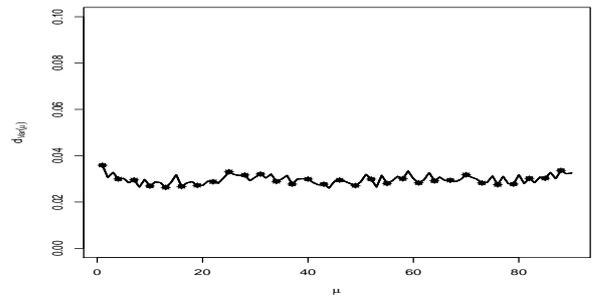
(e)  $\lambda \geq 0 - n = 30$



(f)  $\lambda \geq 0 - n = 30$



(g)  $\lambda \geq 0 - n = 90$



(h)  $\lambda \geq 0 - n = 90$

Figura 5.39: Cenário degrau não simétrico - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .

Nota-se no cenário degrau não simétrico que as diferenças  $d_M$  ficam em torno de 0.45 independente do valor de  $\lambda$ , independente do tamanho da amostra ou independente do valor de  $\epsilon$ , indicando a robustez do MPP para esse cenário onde a diferença entre os blocos é maior que nos demais estudados aqui. A Figura 5.39(b), (d), (f) e (h) mostram que as diferenças  $d_{var}$  diminuem quando  $n = 90$ .

As Figuras 5.40 e 5.41 ilustram os resultados para o cenário escada crescente. Observa-se que a maior magnitude da distância  $d_M$  ocorre para  $n = 30$  e  $\lambda < 0$  próximo ao primeiro ponto de mudança na posição 10. O mesmo ocorre para a medida  $d_{var}$ . No entanto, a medida  $d_{var}$  apresenta maiores oscilações próximo ao ponto de mudança para ambos os tamanhos de amostra e todos os valores de  $\lambda$  avaliados.

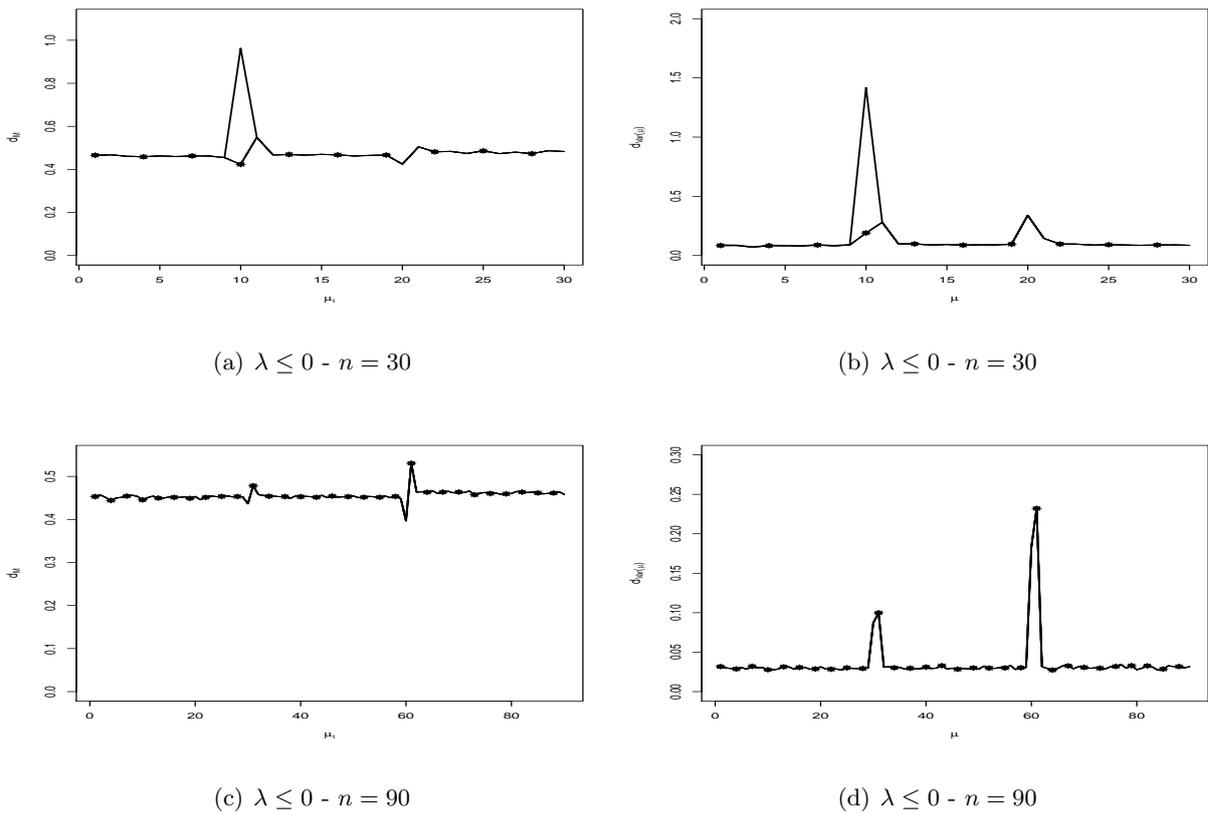
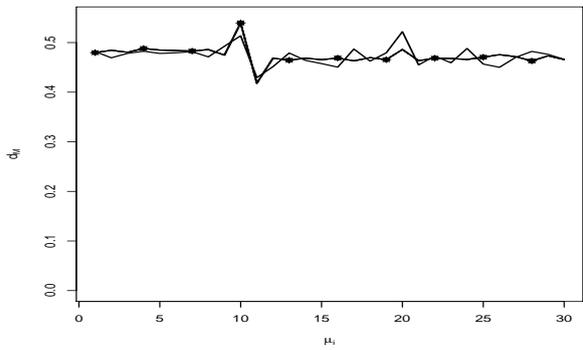
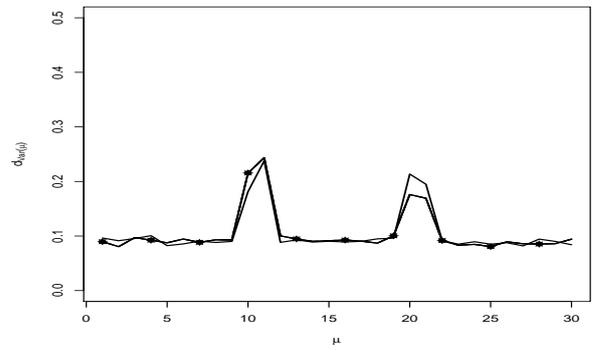


Figura 5.40: Cenário escada crescente - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com \*  $\epsilon = 50$ .

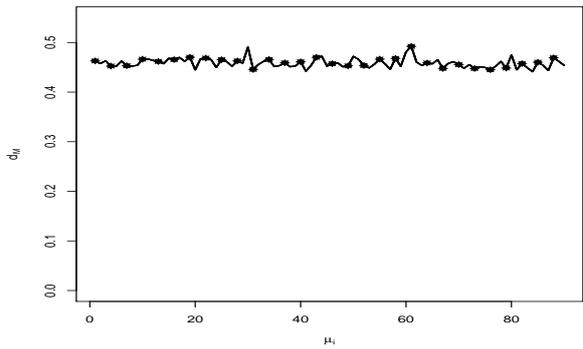
Nota-se também, na Figura 5.40(a) e (b), que para  $\epsilon = 0$  as medidas  $d_M$  e  $d_{var}$  apresentaram maiores magnitudes no instante  $i = 10$ , próximo ao primeiro ponto de mudança.



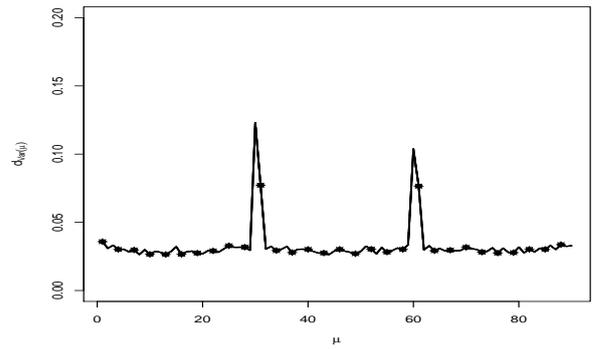
(a)  $\lambda \geq 0$ -  $n = 30$



(b)  $\lambda \geq 0$ -  $n = 30$



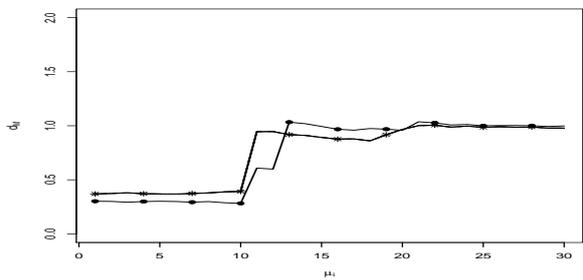
(c)  $\lambda \geq 0$ -  $n = 90$



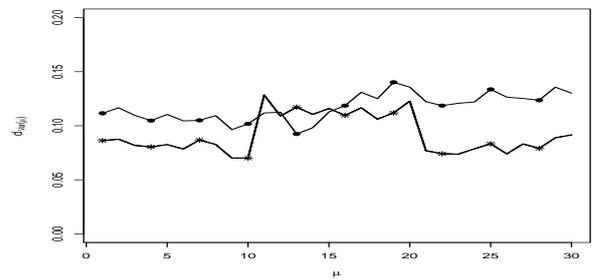
(d)  $\lambda \geq 0$ -  $n = 90$

Figura 5.41: Cenário escada crescente - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .

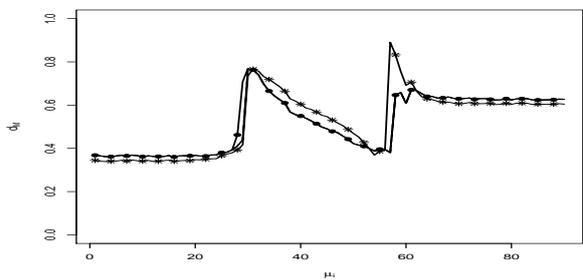
As Figuras 5.42 e 5.43 ilustram os resultados de robustez para o cenário epsilon. Observa-se que para  $n = 30$  há uma maior variação da medida  $d_M$  a partir do primeiro ponto de mudança se  $\lambda \leq 0$ , mas se  $\lambda \geq 0$  a maior variação ocorre até o segundo ponto de mudança. Se  $n = 90$  há um pico próximo aos pontos de mudança para  $\lambda$  negativo, mas se  $\lambda$  é positivo no caso em que  $\epsilon = 0$  o comportamento das estimativas é ao contrário do comportamento para os demais valores de  $\epsilon$ .



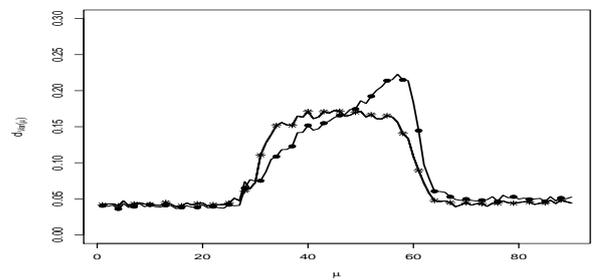
(a)  $\lambda \leq 0 - n = 30$



(b)  $\lambda \leq 0 - n = 30$

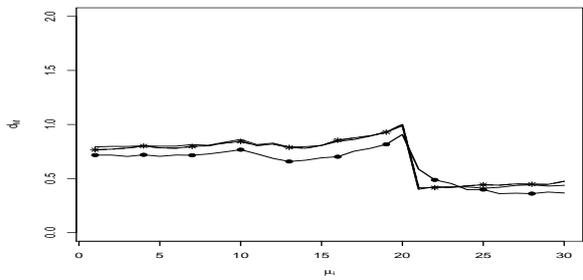


(c)  $\lambda \leq 0 - n = 90$

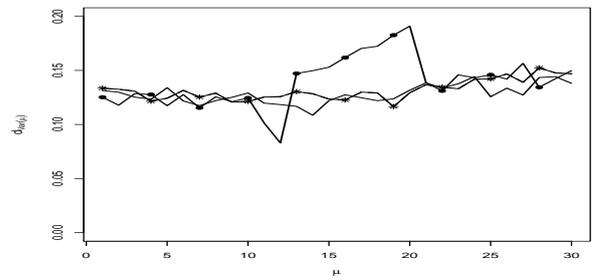


(d)  $\lambda \leq 0 - n = 90$

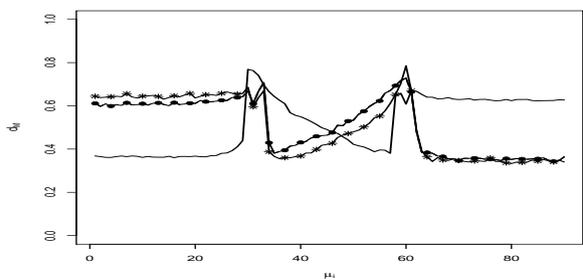
Figura 5.42: Cenário epsilon - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .



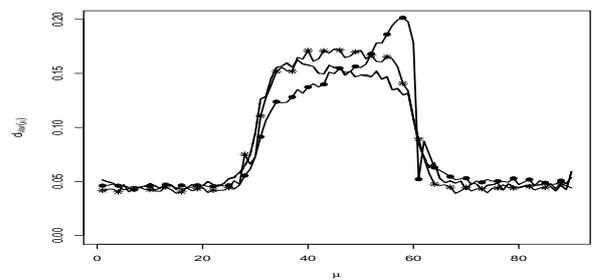
(a)  $\lambda \geq 0 - n = 30$



(b)  $\lambda \geq 0 - n = 30$



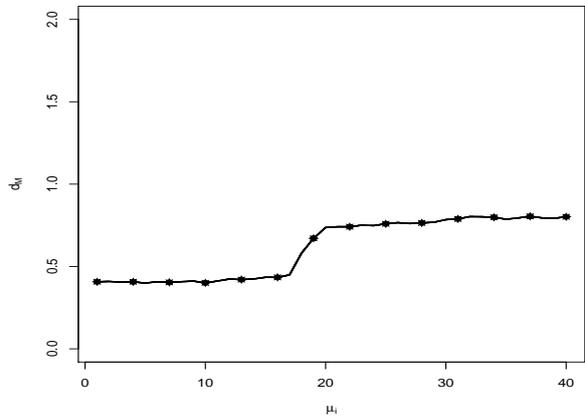
(c)  $\lambda \geq 0 - n = 90$



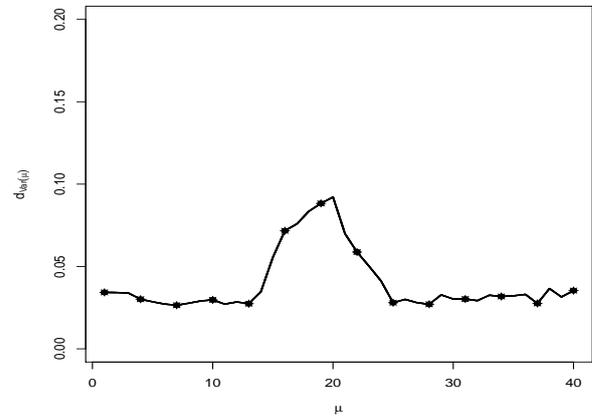
(d)  $\lambda \geq 0 - n = 90$

Figura 5.43: Cenário epsilon - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .

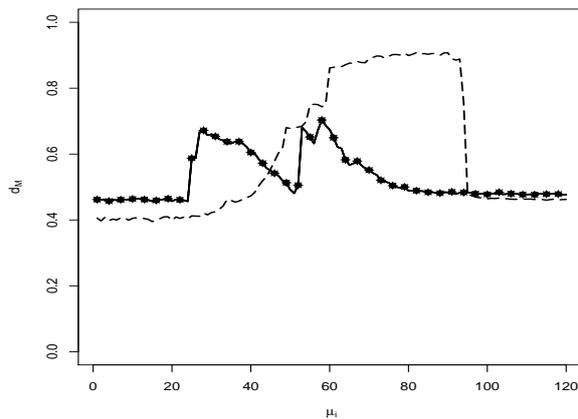
Os resultados para o cenário tendência são apresentados nas Figuras 5.44 e 5.45.



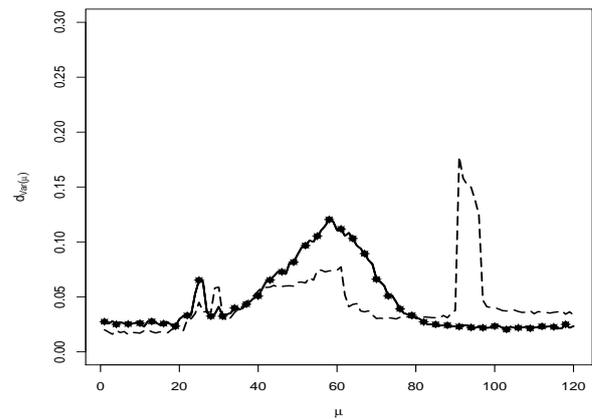
(a)  $\lambda \leq 0 - n = 40$



(b)  $\lambda \leq 0 - n = 40$



(c)  $\lambda \leq 0 - n = 120$

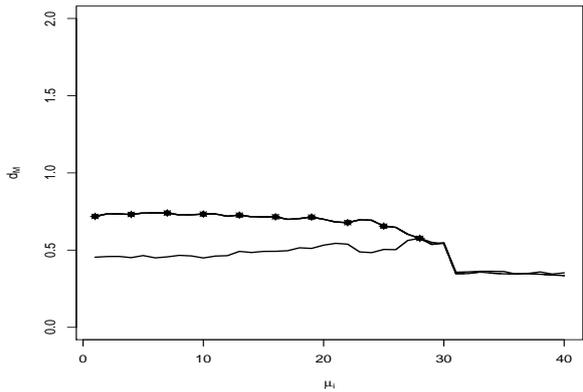


(d)  $\lambda \leq 0 - n = 120$

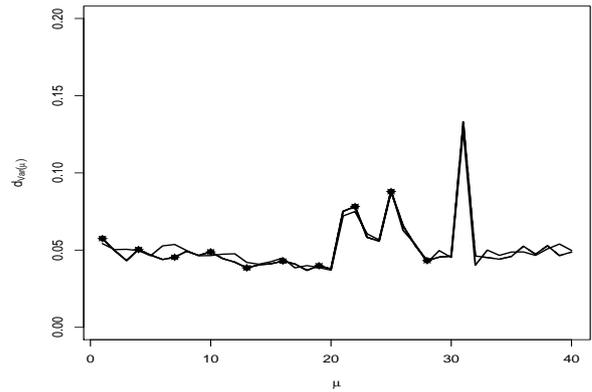
Figura 5.44: Cenário tendência - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com \*  $\epsilon = 50$ .

No cenário tendência nota-se que para  $n = 40$  e  $\lambda \geq 0$  a variação é maior até o terceiro ponto de mudança (posição 31) e para  $\lambda$  negativo a variação é maior a partir do primeiro ponto de mudança (posição 11). Para o tamanho da amostra 120 houve uma maior variação no comportamento das estimativas produto, observa-se que para  $\epsilon = 5$  e  $\lambda$  negativo a variação no terceiro bloco foi maior

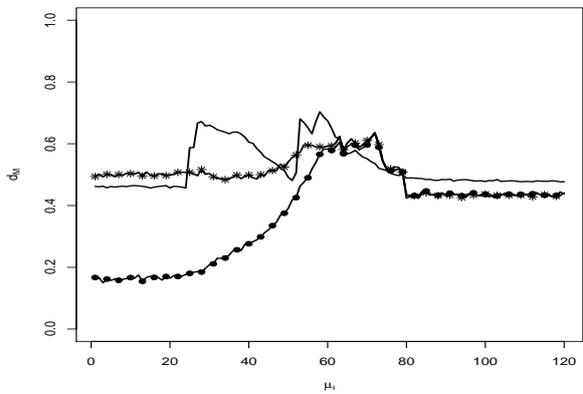
que para os demais valores de  $\epsilon$ . Em relação a medida  $d_{var}$  a variação é maior entre o primeiro e o terceiro ponto de mudança, onde existe a tendência determinística na série, mas o modelo partição produto tende a ser mais sensível para estimar a média nos casos em que a série possui tendência.



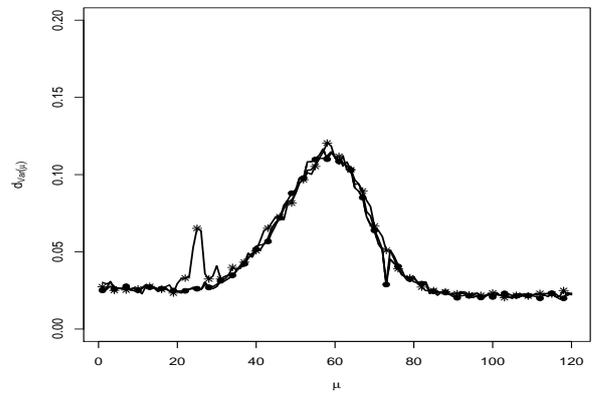
(a)  $\lambda \geq 0$ -  $n = 40$



(b)  $\lambda \geq 0$ -  $n = 40$



(c)  $\lambda \geq 0$ -  $n = 120$



(d)  $\lambda \geq 0$ -  $n = 120$

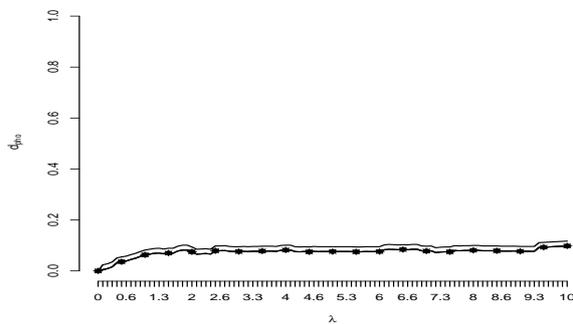
Figura 5.45: Cenário tendência - Magnitude das medidas  $d_M$  e  $d_{var}$  para cada  $\mu_i$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .

A próxima seção destina-se aos resultados da robustez bayesiana em relação ao número de blocos e partição de interesse.

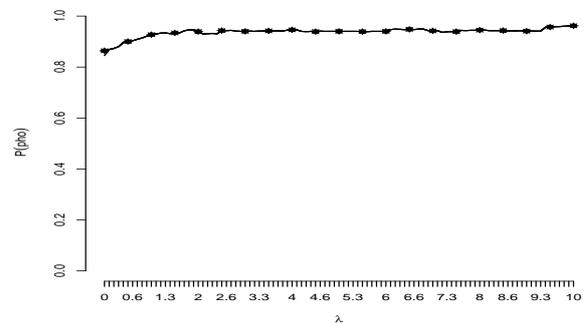
### 5.2.2 Robustez na partição e no número de blocos

Nessa seção são apresentados os resultados da robustez global para a partição de interesse e o número de blocos.

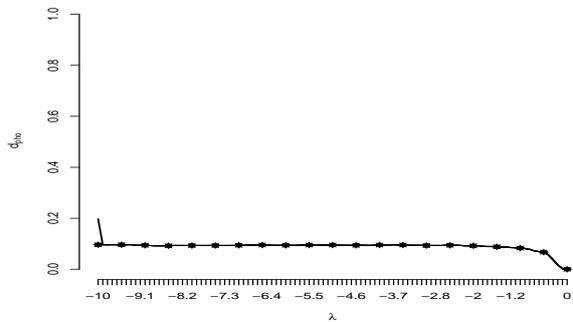
As Figuras 5.46 e 5.47 referem-se ao cenário degrau positivo e apresentam os resultados da partição de interesse.



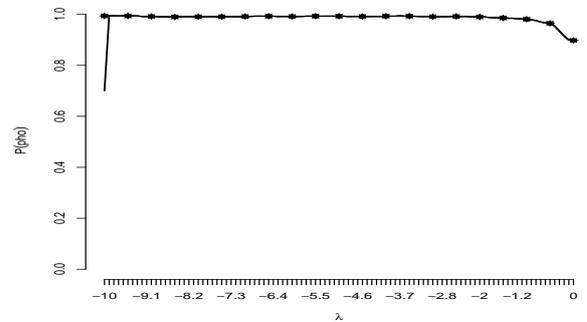
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$

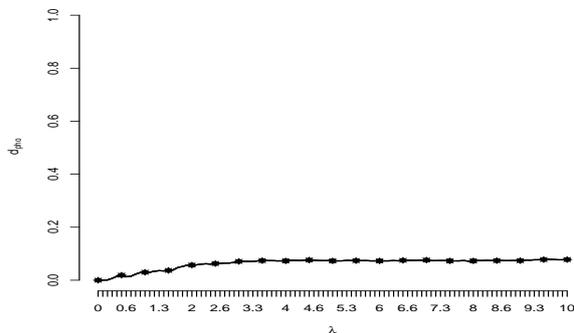


(c)  $d_\rho$  com  $\lambda \leq 0$

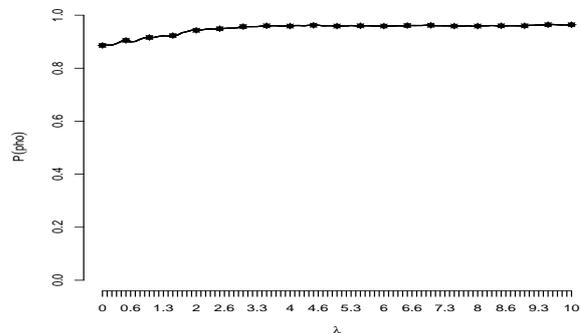


(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

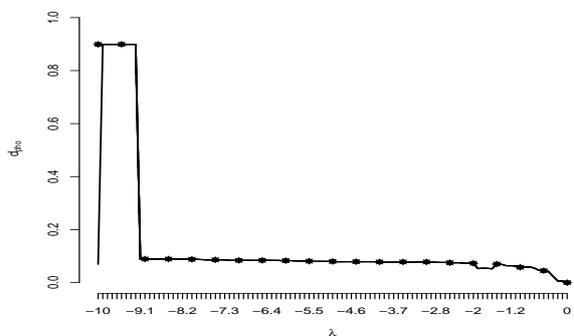
Figura 5.46: Cenário degrau positivo - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$



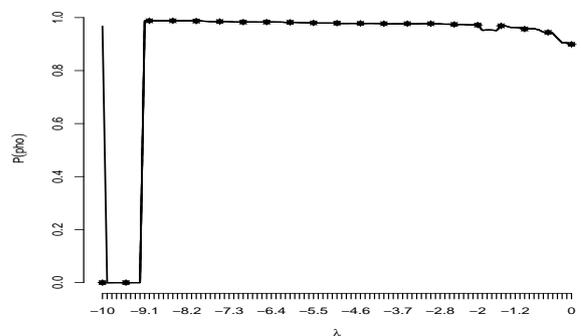
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$



(c)  $d_\rho$  com  $\lambda \leq 0$

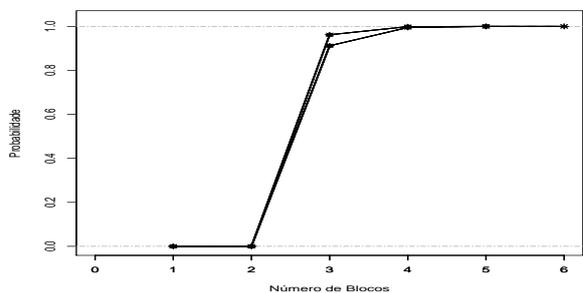


(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

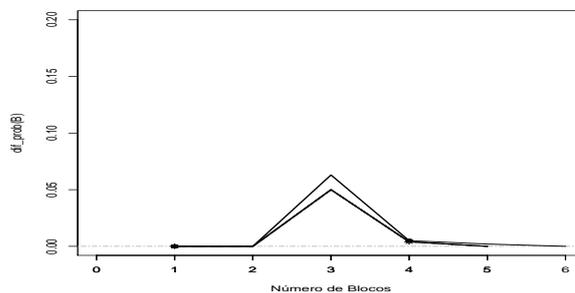
Figura 5.47: Cenário degrau positivo - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$

Se  $\lambda > 0$ , independente do tamanho da amostra a probabilidade da partição de interesse ser identificada corretamente aumenta quando  $\lambda$  se afasta de zero. Mas o modelo com  $\lambda$  negativo se torna impreciso, pois para  $\lambda$  próximo ao intervalo  $(-10, -9)$ , o modelo identificou com probabilidade zero a partição de interesse. Isso ocorreu porque na verdade o modelo indicou a partição  $\rho = \{0, 30, 61\}$  com probabilidade 1 quando na verdade a partição correta era  $\rho = \{0, 30, 60\}$ .

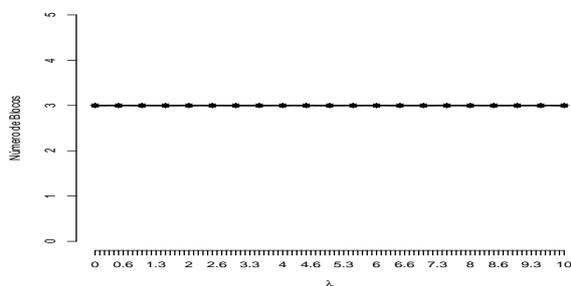
A moda do número de blocos, a função de distribuição do número de blocos e a probabilidade da moda *a posteriori* são apresentadas nas Figuras 5.48, 5.49, 5.50, 5.51.



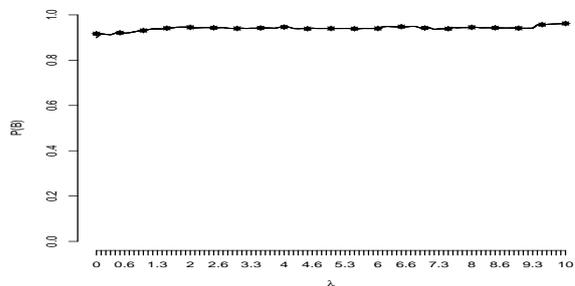
(a)  $P(B \geq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$

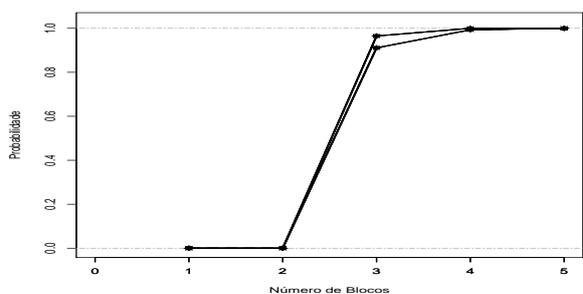


(c) Moda

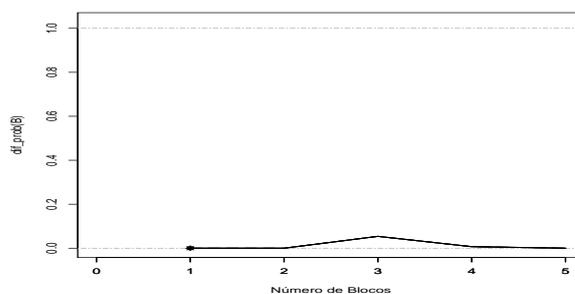


(d)  $P(B=b)$

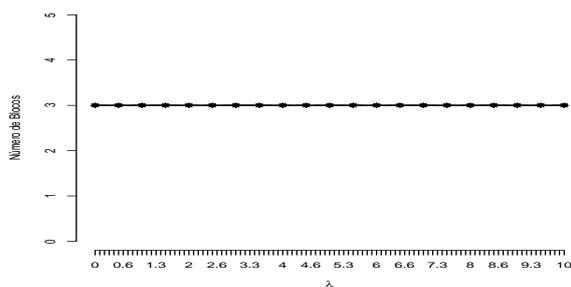
Figura 5.48:  $\lambda \geq 0$  - Cenário degrau positivo - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$



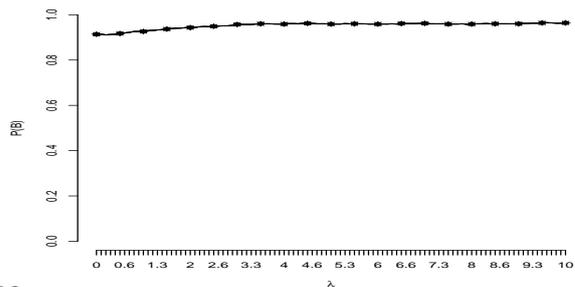
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$

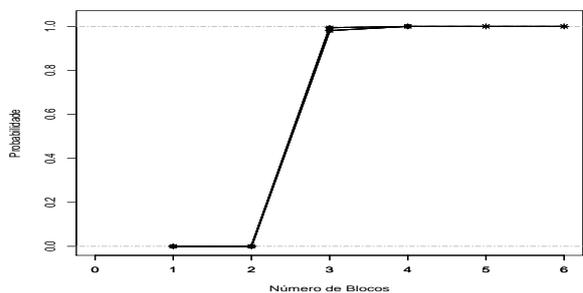


(c) Moda

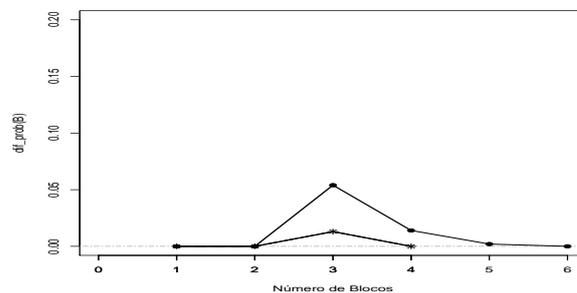


(d)  $P(B=b)$

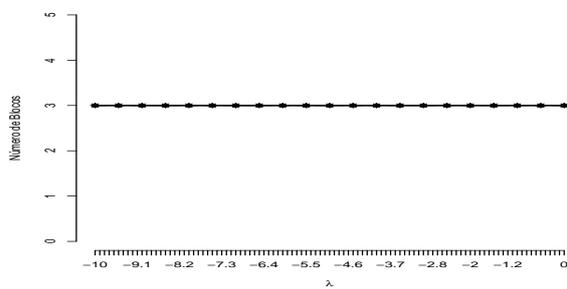
Figura 5.49:  $\lambda \geq 0$  - Cenário degrau positivo - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$



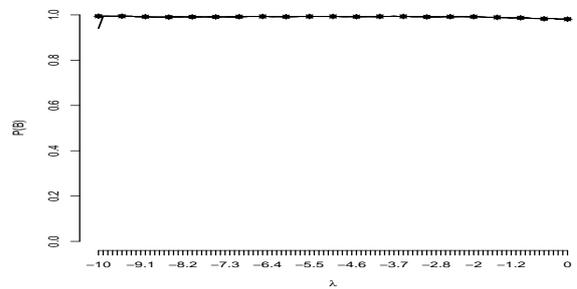
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B=b) - P_{\lambda=0}(B=b)|$

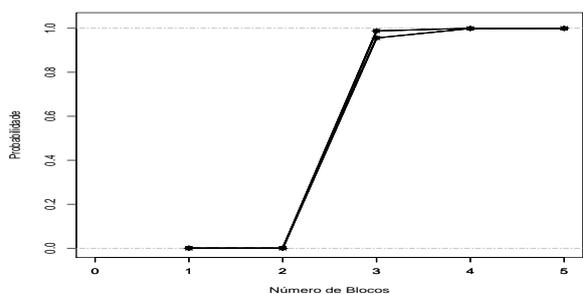


(c) Moda

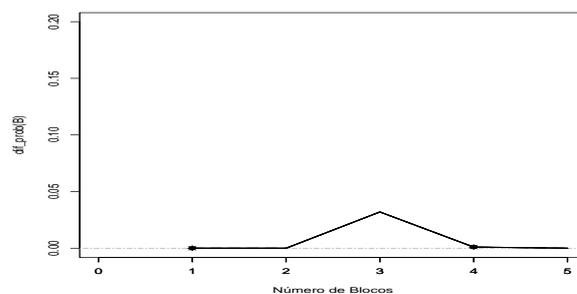


(d)  $P(B=b)$

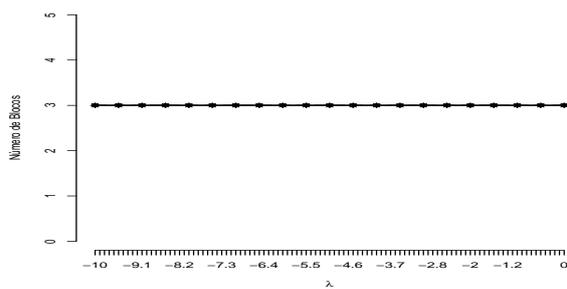
Figura 5.50:  $\lambda \leq 0$  - Cenário degrau positivo - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$



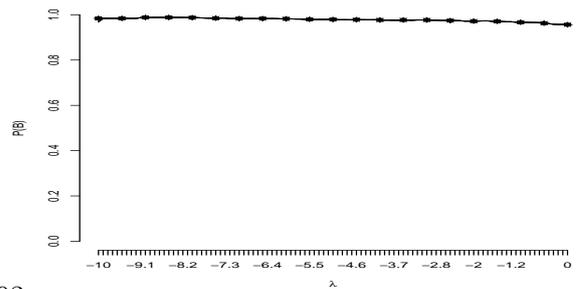
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B=b) - P_{\lambda=0}(B=b)|$



(c) Moda



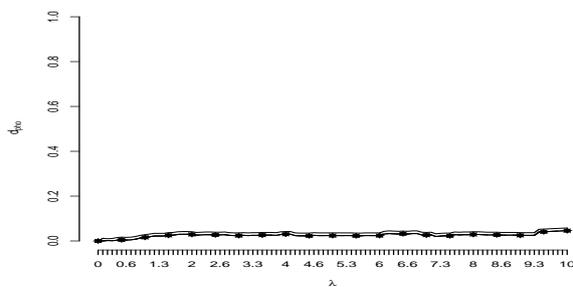
(d)  $P(B=b)$

Figura 5.51:  $\lambda \leq 0$  - Cenário degrau positivo - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

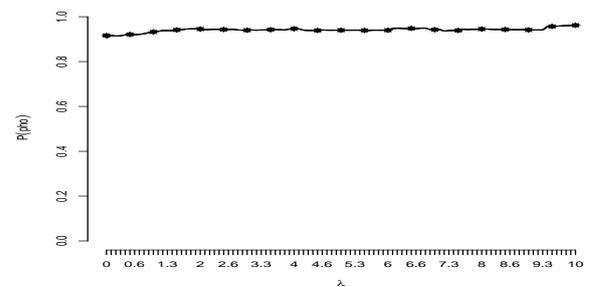
Observa-se que o número de blocos foi identificado corretamente para ambos os tamanho de amostra avaliados e para todos os valores de  $\lambda$  entre  $-10$  e  $10$ . Além disso, a probabilidade da moda *a posteriori* ( $B = 3$ ) aumenta quando  $\lambda$  vai para  $-10$  ou  $10$ .

As funções de distribuição, definidas na expressão (4.5), para  $n = 90$  são praticamente idênticas, conseqüentemente a diferença entre essas fuções são mínimas, Figuras 5.49 e 5.51 (b).

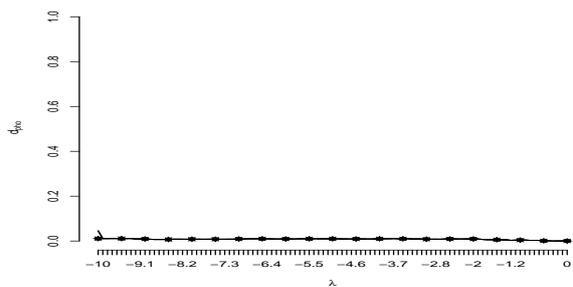
As Figuras 5.52 e 5.53 mostram os resultados para a robustez na partição de interesse para o cenário degrau não simétrico.



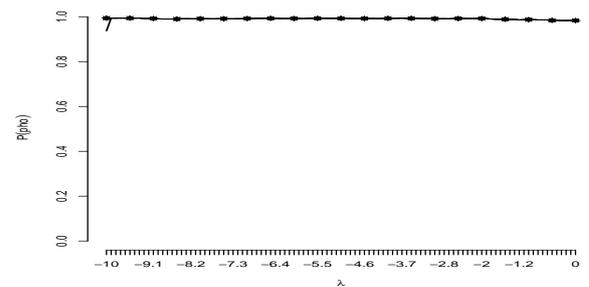
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$

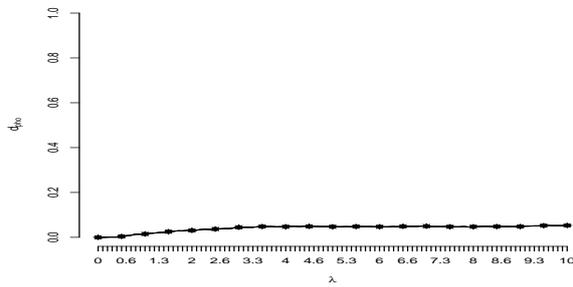


(c)  $d_\rho$  com  $\lambda \leq 0$

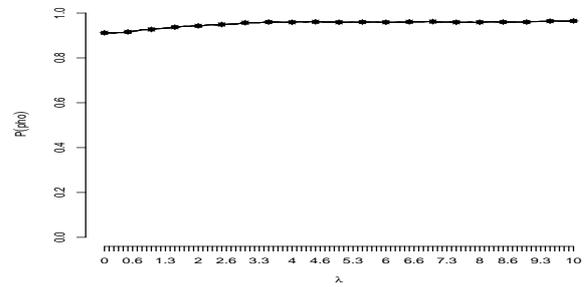


(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

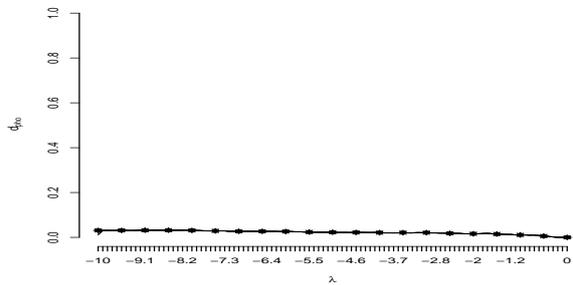
Figura 5.52: Cenário degrau não simétrico - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .



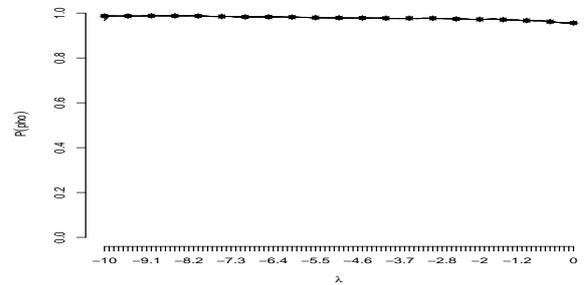
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$



(c)  $d_\rho$  com  $\lambda \leq 0$

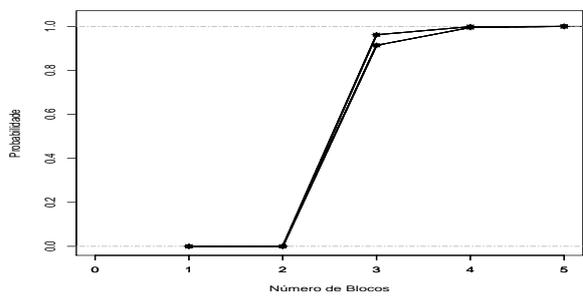


(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

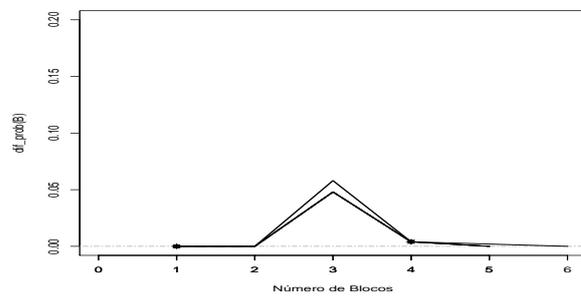
Figura 5.53: Cenário degrau não simétrico - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

A medida de robustez  $d_\rho$  é próxima de zero para  $\lambda < 0$  e tamanho de amostra  $n = 90$ . A probabilidade com que a partição de interesse é indicada aumenta levemente quando o parâmetro de assimetria se afasta de zero.

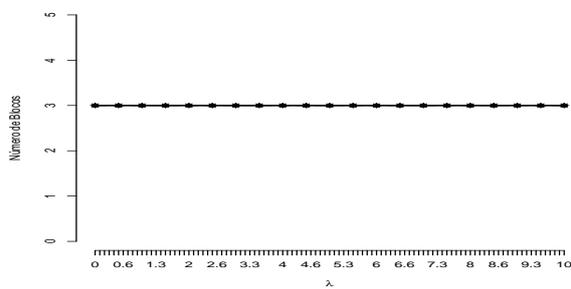
Os resultados para a robustez do número de blocos para o cenário degrau não simétrico estão nas Figuras 5.54, 5.55, 5.56 e 5.57. Observa-se nessas figuras que o número de blocos é identificado corretamente com alta probabilidade para todos os valores de  $\lambda$  avaliados e as diferenças são mínimas para  $\lambda < 0$  nesse cenário. As diferenças entre as funções de distribuição são pequenas e ocorrem próximas de  $B = 3$ .



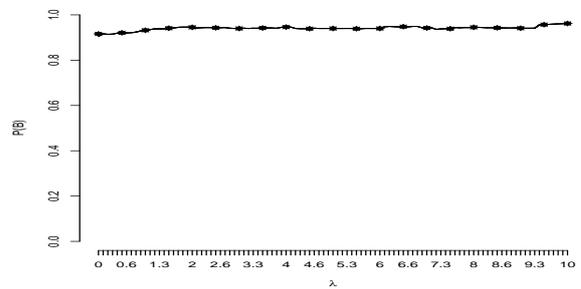
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$

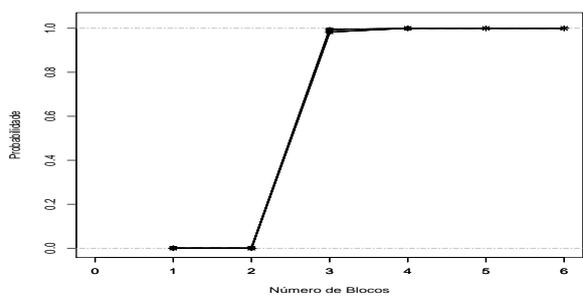


(c) Moda

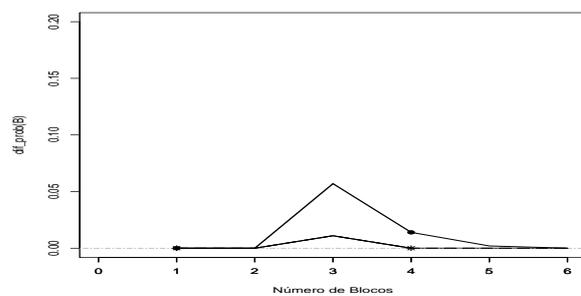


(d)  $P(B=b)$

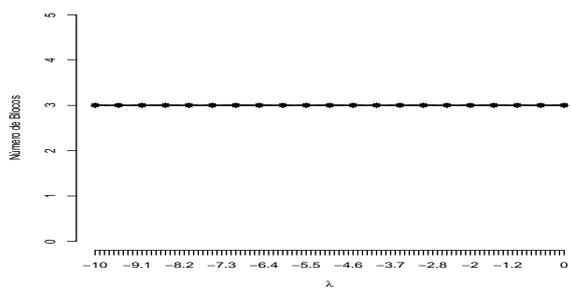
Figura 5.54:  $\lambda \geq 0$  - Cenário degrau não simétrico - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .



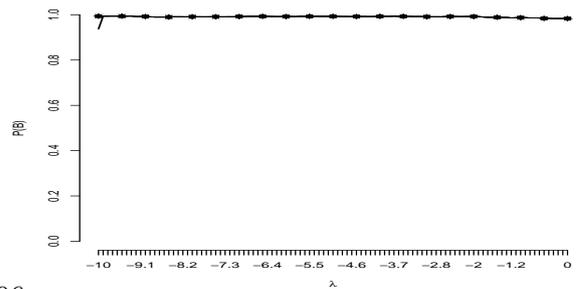
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$

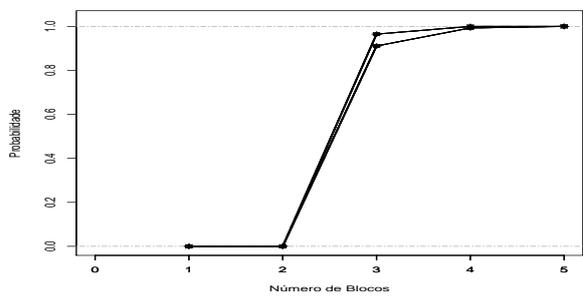


(c) Moda

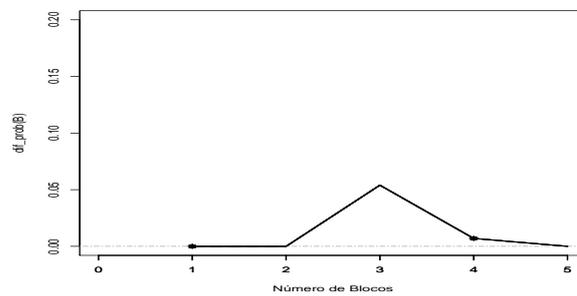


(d)  $P(B=b)$

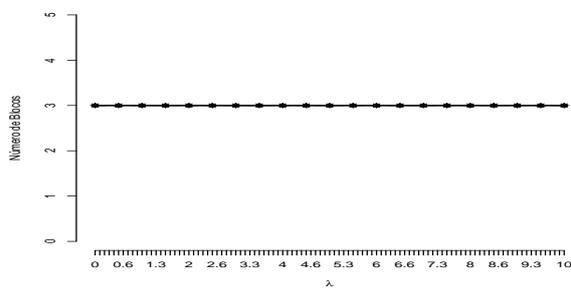
Figura 5.55:  $\lambda \leq 0$  - Cenário degrau não simétrico - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .



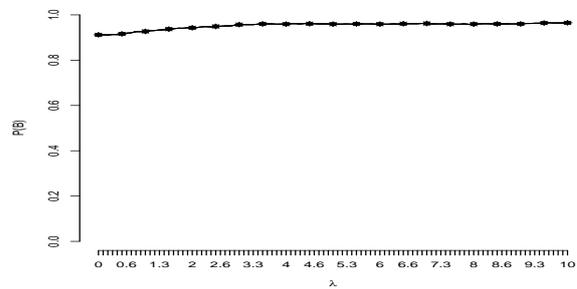
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$

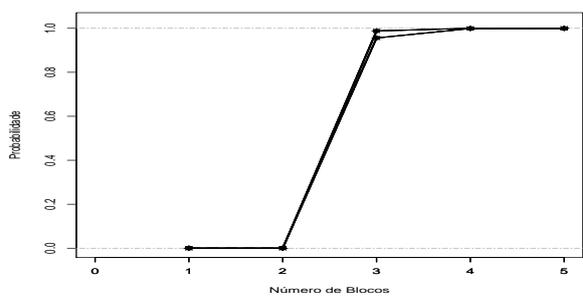


(c) Moda

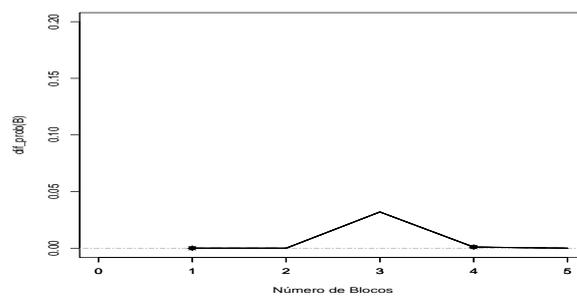


(d)  $P(B=b)$

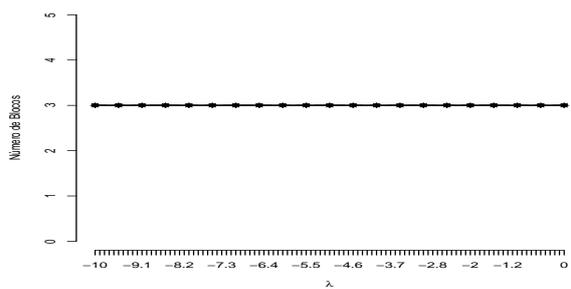
Figura 5.56:  $\lambda \geq 0$  - Cenário degrau não simétrico - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .



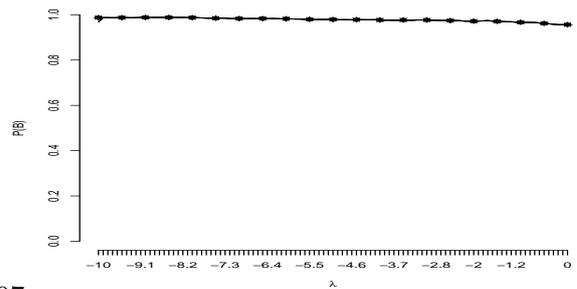
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$



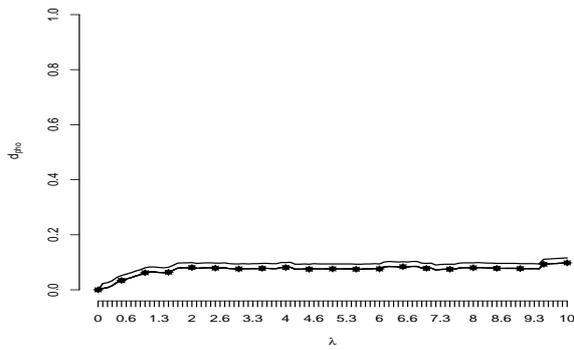
(c) Moda



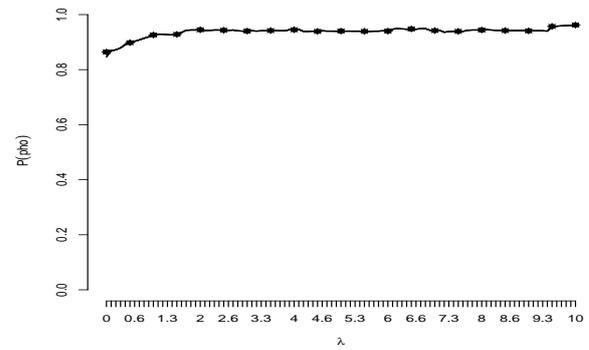
(d)  $P(B=b)$

Figura 5.57:  $\lambda \leq 0$  - Cenário degrau não simétrico - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

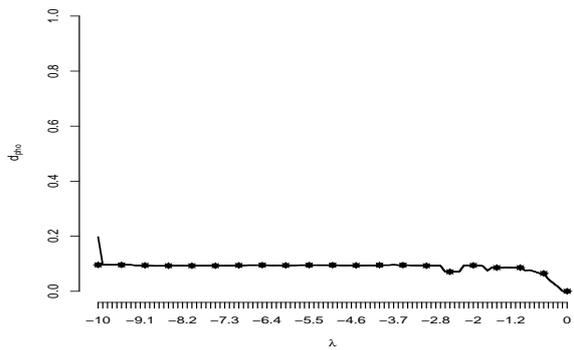
Os resultados para a partição de interesse para o cenário escada crescente encontram-se na Figuras 5.58 e 5.59.



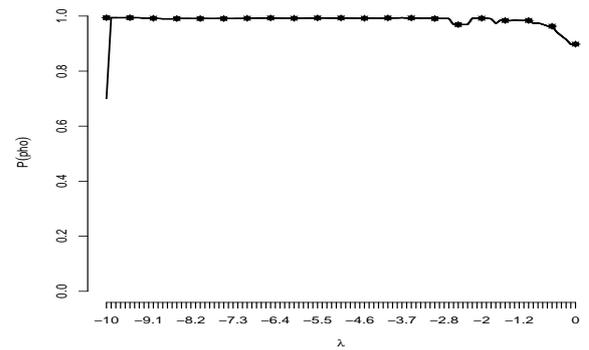
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$



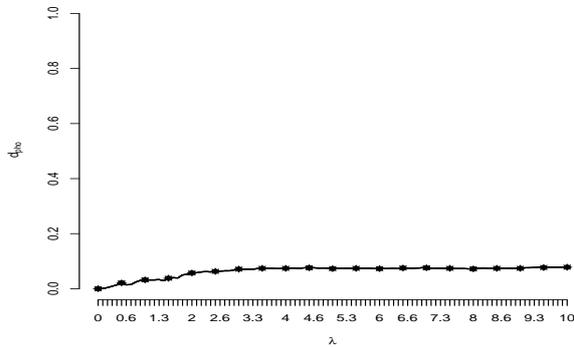
(c)  $d_\rho$  com  $\lambda \leq 0$



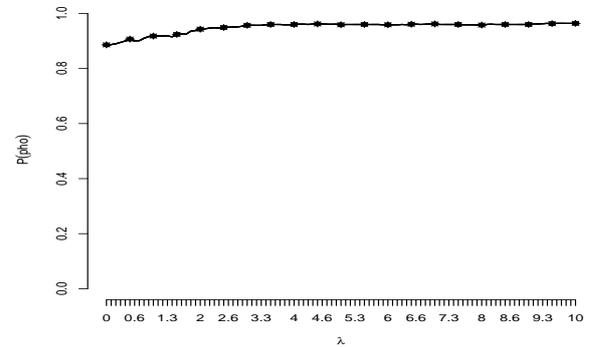
(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

Figura 5.58: Cenário escada crescente - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .

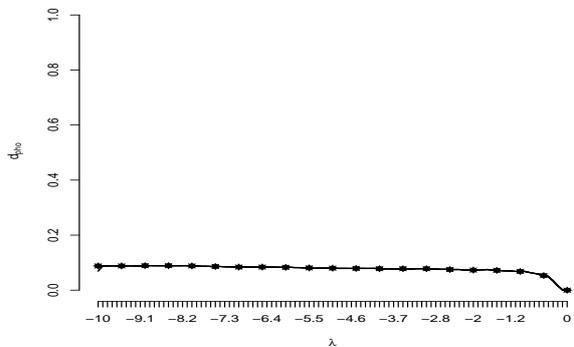
No cenário escada crescente a partição de interesse foi indentificada com menor probabilidade pontualmente para o caso em que  $\epsilon = 0$  e  $\lambda = -10$ . Com exceção desse caso o padrão se mantém e a partição de interesse é identificada com maior probabilidade a medida que  $\lambda$  se afasta de zero.



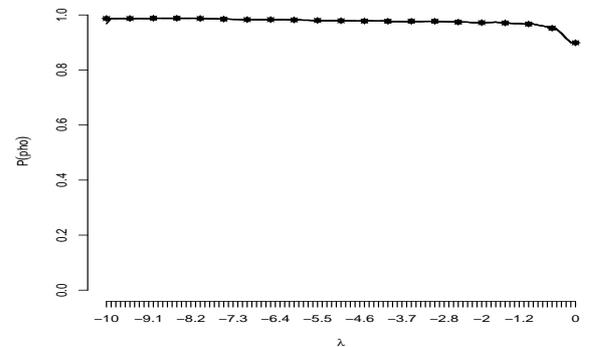
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$



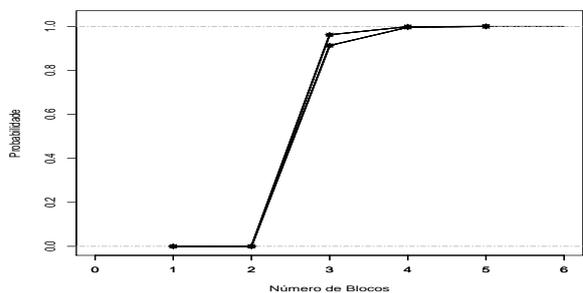
(c)  $d_\rho$  com  $\lambda \leq 0$



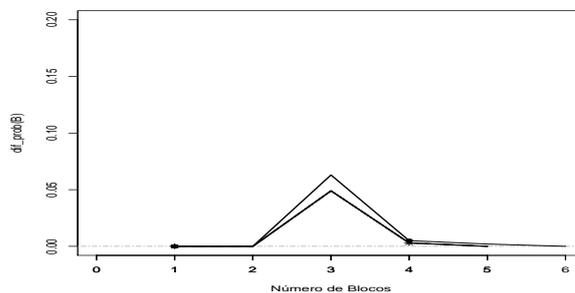
(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

Figura 5.59: Cenário escada crescente - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

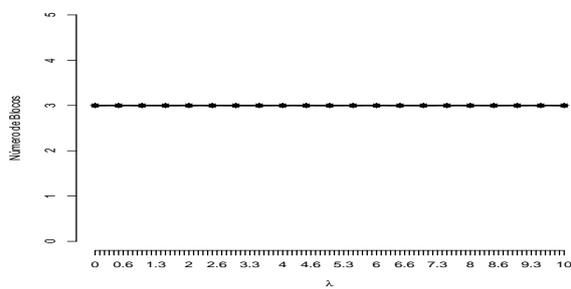
As Figuras 5.60, 5.61, 5.62, 5.63 ilustram os resultados para o número de blocos no cenário escada crescente. Nota-se nessas Figuras que a medida  $d_B$  é menor para  $\lambda < 0$ . Além disso, o modelo identifica corretamente o número de blocos para todos os valores de  $\lambda$ , mas para  $\lambda < 0$  as probabilidades são próximas de 1. Se  $\lambda$  é positivo fica mais evidente que a medida que  $\lambda$  se afasta de zero a probabilidade da moda *a posteriori* aumenta.



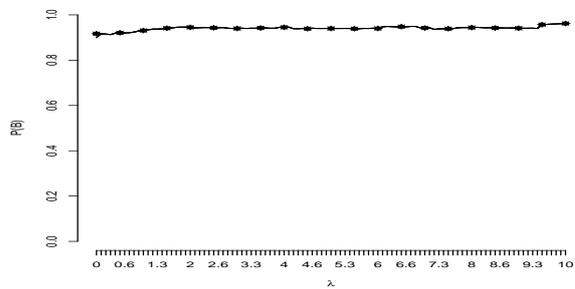
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B=b) - P_{\lambda=0}(B=b)|$

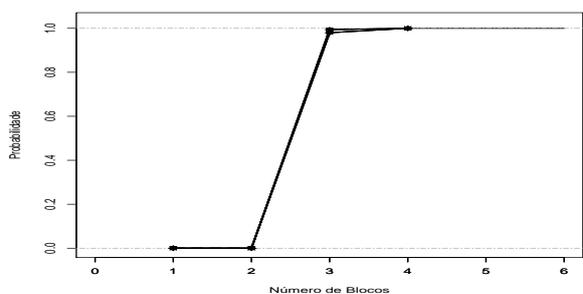


(c) Moda

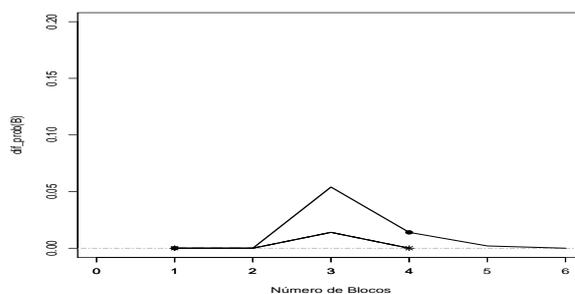


(d)  $P(B=b)$

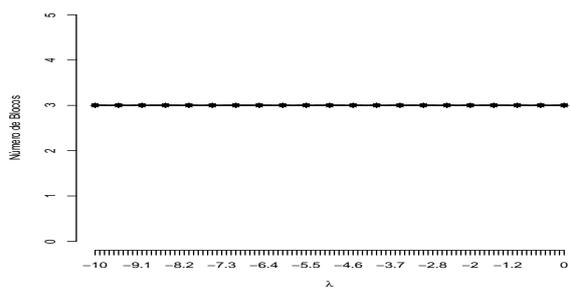
Figura 5.60:  $\lambda \geq 0$  - Cenário escada crescente - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .



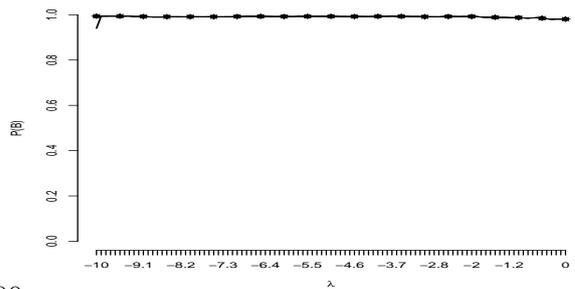
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B=b) - P_{\lambda=0}(B=b)|$

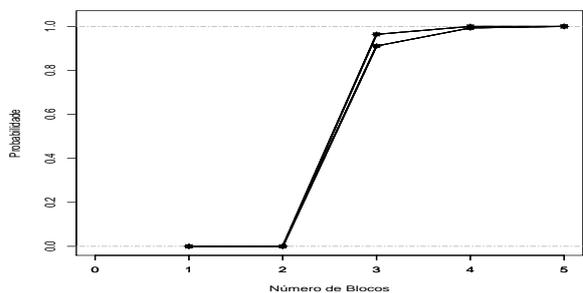


(c) Moda

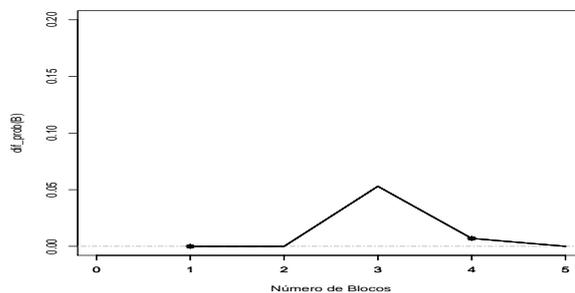


(d)  $P(B=b)$

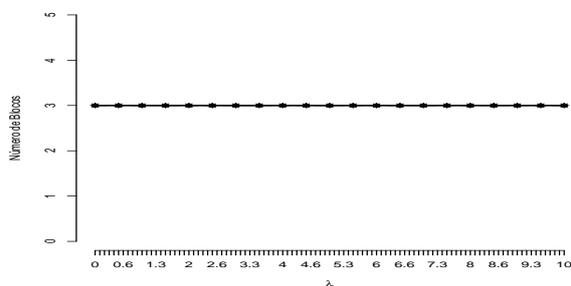
Figura 5.61:  $\lambda \leq 0$  - Cenário escada crescente - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .



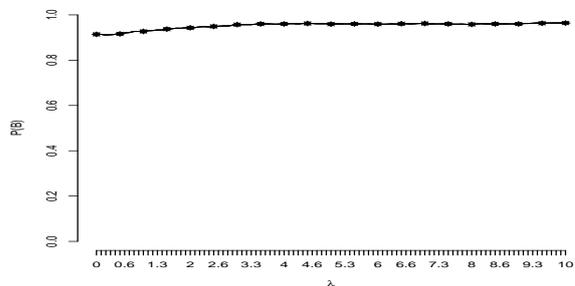
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$

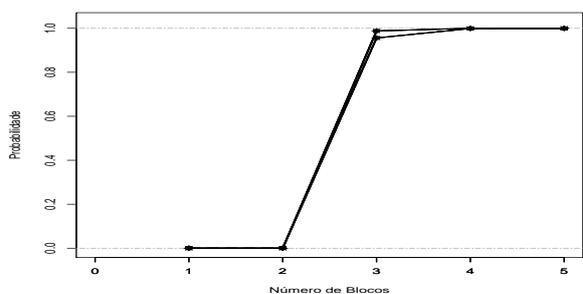


(c) Moda

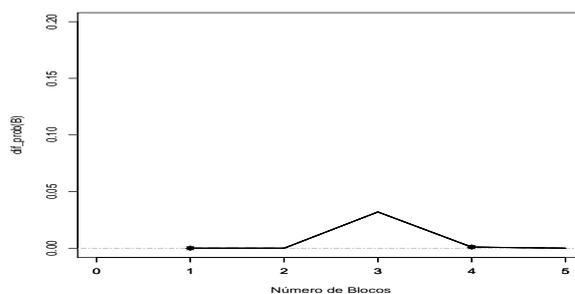


(d)  $P(B=b)$

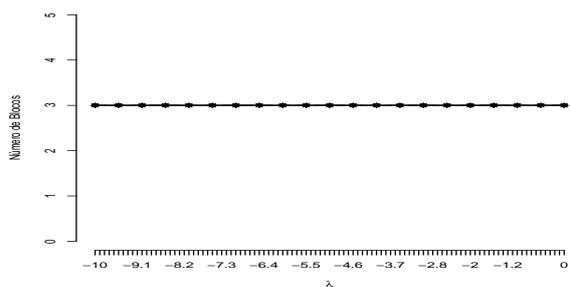
Figura 5.62:  $\lambda \geq 0$  - Cenário escada crescente - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .



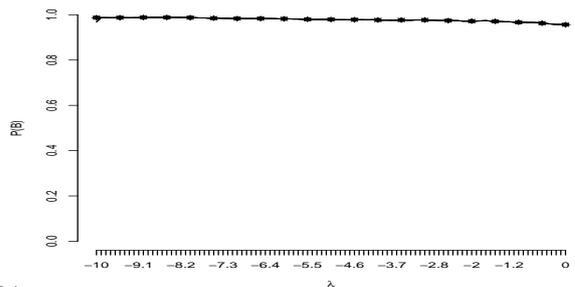
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$



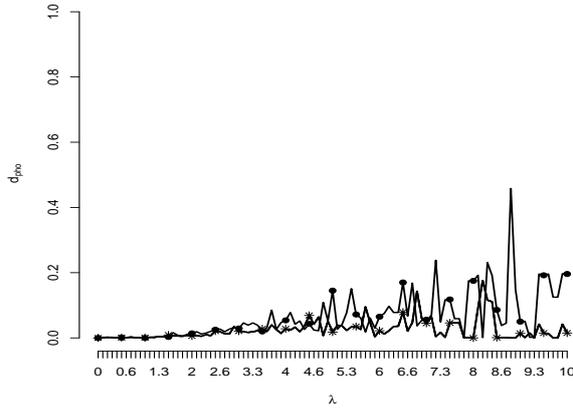
(c) Moda



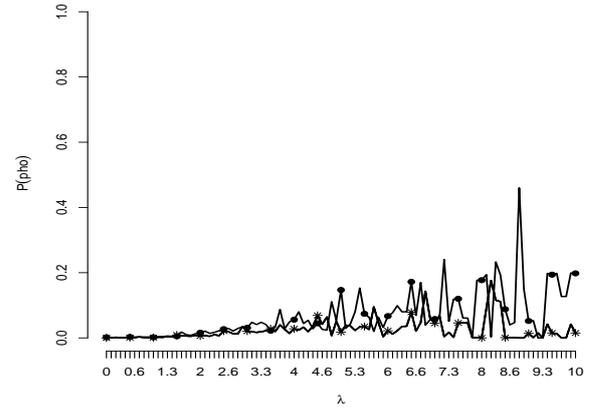
(d)  $P(B=b)$

Figura 5.63:  $\lambda \leq 0$  - Cenário escada crescente - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

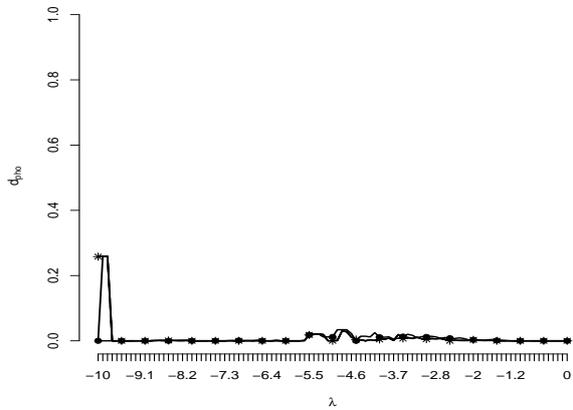
As Figuras 5.64 e 5.65 mostram o comportamento da partição em relação aos valores de  $\lambda$  avaliados no cenário epsilon.



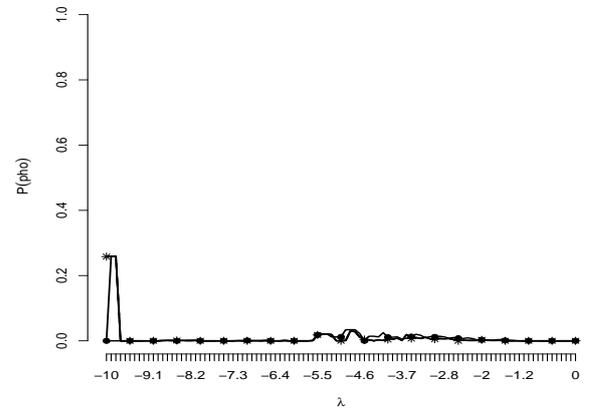
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$



(c)  $d_\rho$  com  $\lambda \leq 0$

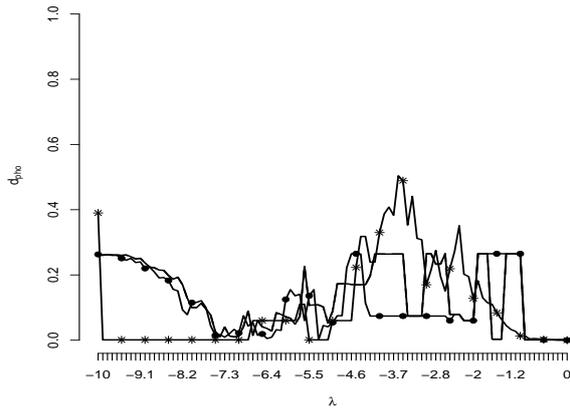


(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

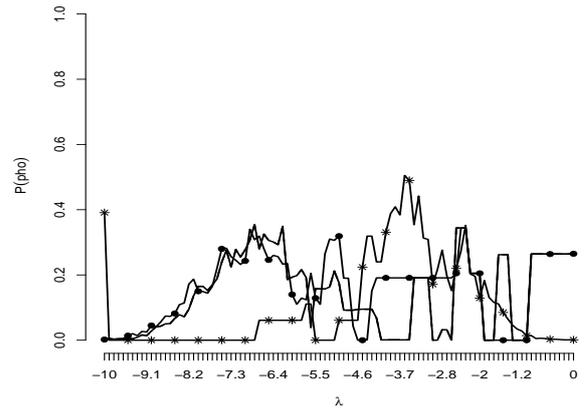
Figura 5.64: Cenário epsilon - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .

Observa-se no cenário epsilon que para  $n = 30$  e  $\lambda$  negativo o modelo não foi capaz de identificar a partição corretamente, no entanto se o tamanho da amostra aumenta a probabilidade de estimar corretamente a partição aumenta, mas não existe um padrão bem definido de quais valores

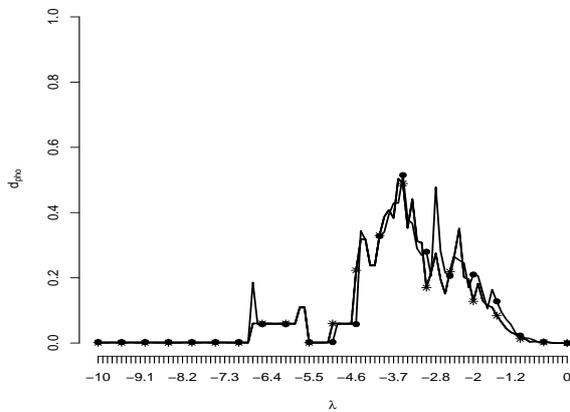
de  $\lambda$  fornecem as melhores estimativas. Se  $\lambda$  é negativo e  $n = 90$  a probabilidade é maior entre  $\lambda = (-4.7, -1)$ , mas se  $\lambda$  é positivo o comportamento da estimativa da partição depende do valor de  $\epsilon$ , mas para  $\epsilon = 50$  o comportamento é similar ao comportamento de  $\lambda \leq 0$ .



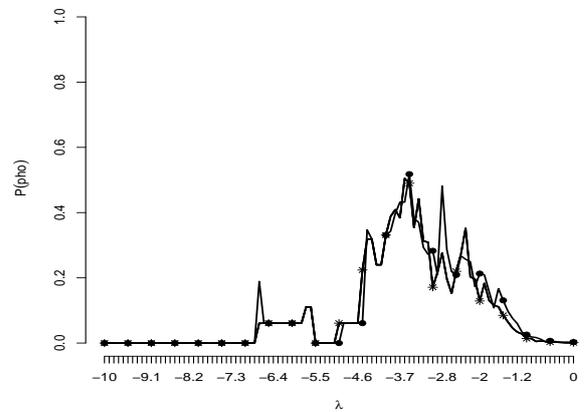
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$



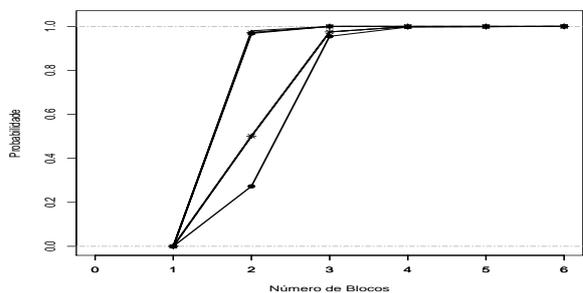
(c)  $d_\rho$  com  $\lambda \leq 0$



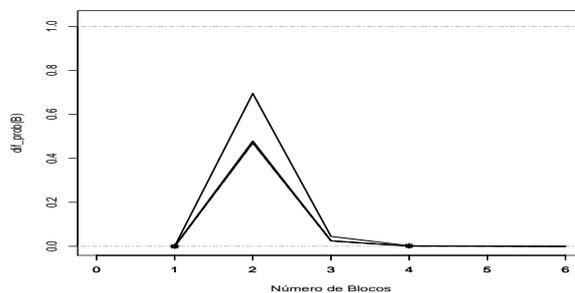
(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

Figura 5.65: Cenário epsilon - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

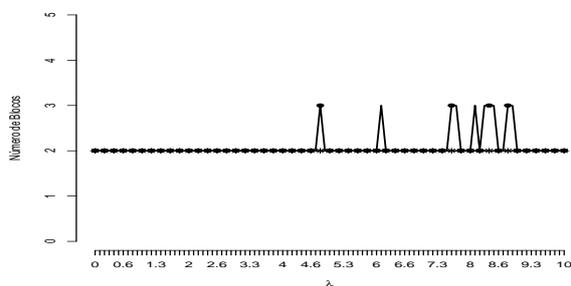
Nas Figuras 5.66, 5.67, 5.68 e 5.69 encontram-se os resultados para o número de blocos no cenário epsilon.



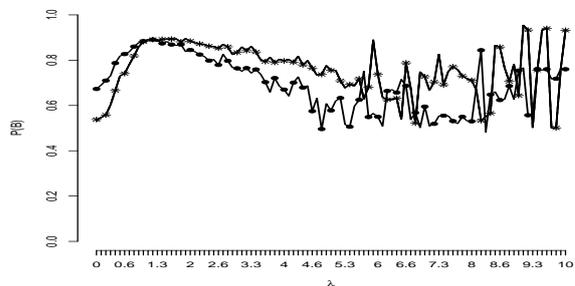
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B=b) - P_{\lambda=0}(B=b)|$

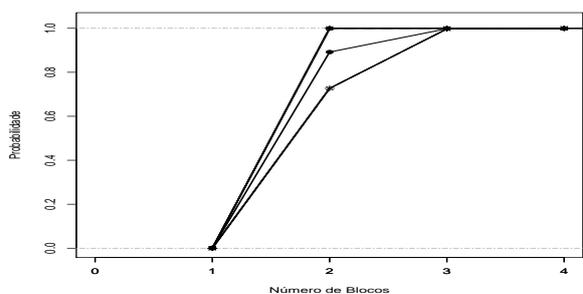


(c) Moda

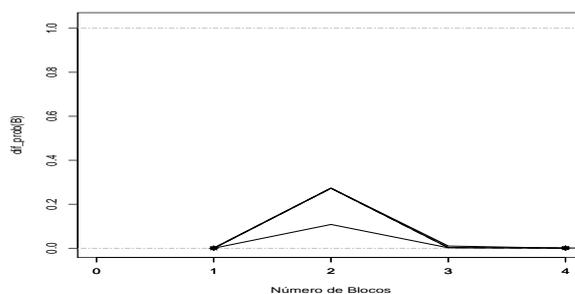


(d)  $P(B=b)$

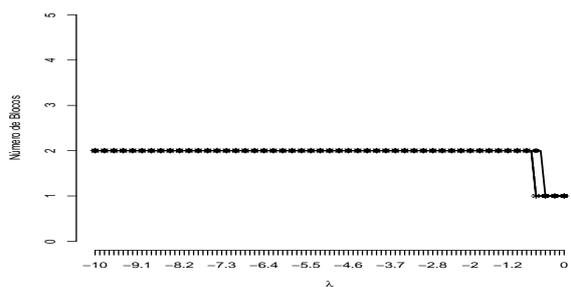
Figura 5.66:  $\lambda \geq 0$  - Cenário epsilon - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .



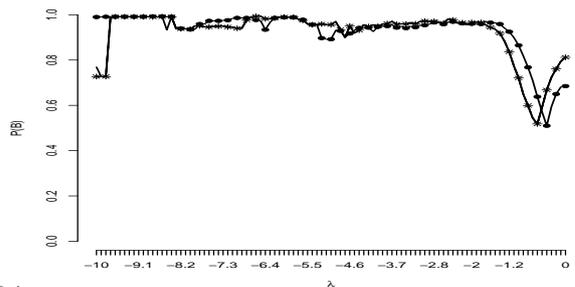
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B=b) - P_{\lambda=0}(B=b)|$

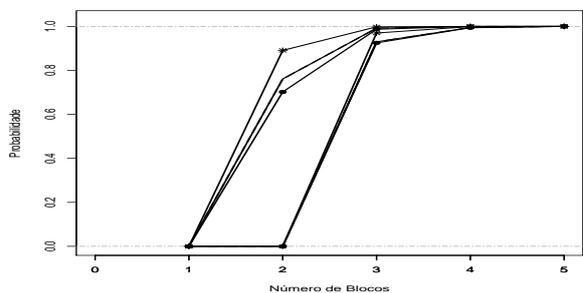


(c) Moda

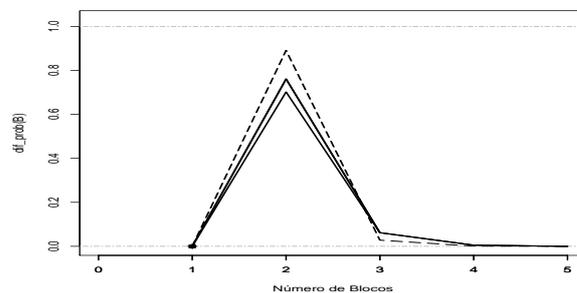


(d)  $P(B=b)$

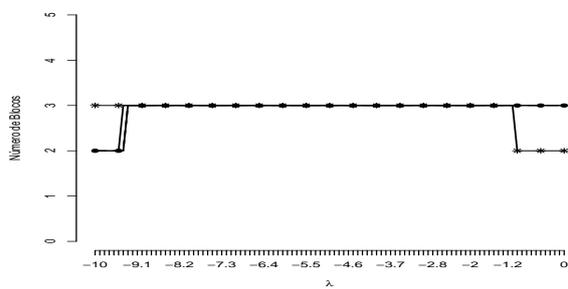
Figura 5.67:  $\lambda \leq 0$  - Cenário epsilon - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 30$ .



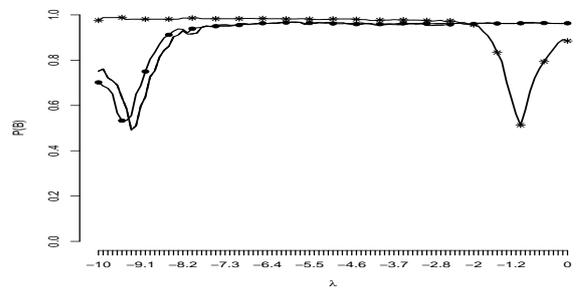
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$

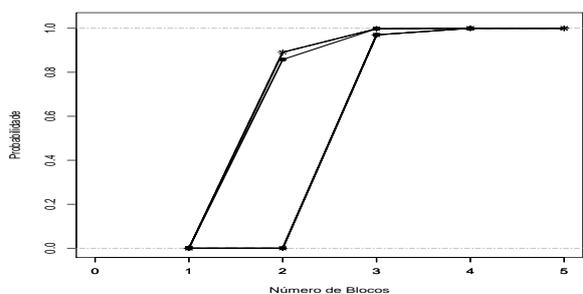


(c) Moda

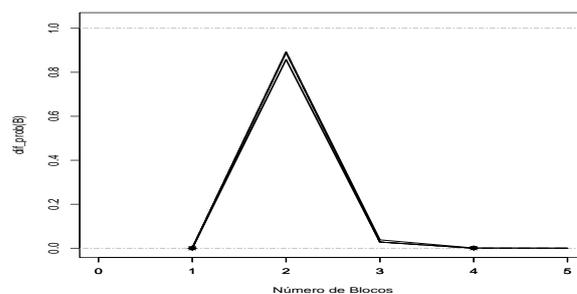


(d)  $P(B=b)$

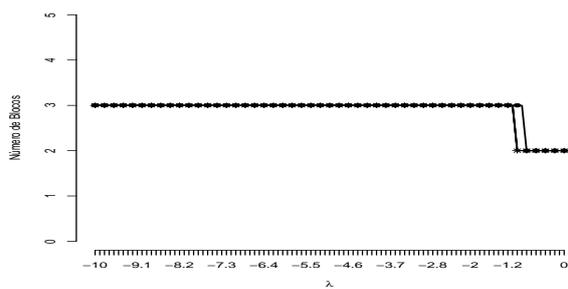
Figura 5.68:  $\lambda \geq 0$  - Cenário epsilon - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .



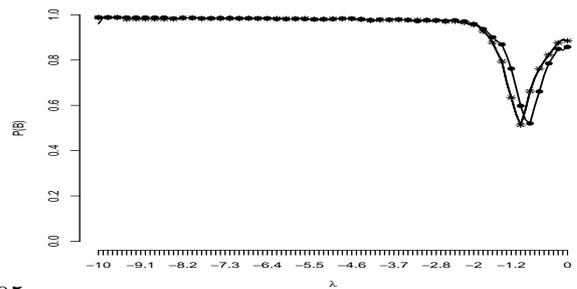
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$



(c) Moda

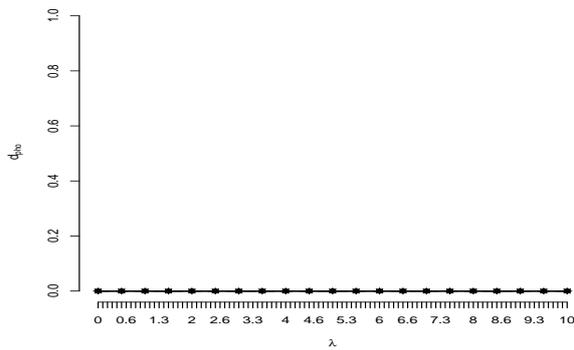


(d)  $P(B=b)$

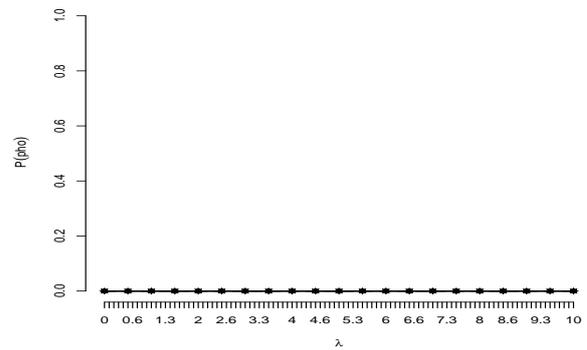
Figura 5.69:  $\lambda \leq 0$  - Cenário epsilon - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 90$ .

O número de blocos no cenário epsilon foi estimado corretamente para valores de  $\lambda \leq -1$  com alta probabilidade para  $n = 30$ , mas se o tamanho da amostra aumenta o número de blocos para  $\lambda \leq -1$  é subestimado.

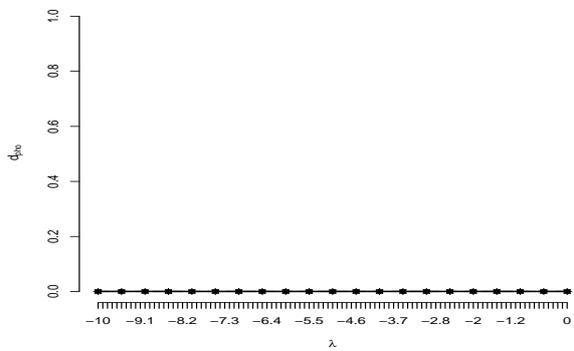
As Figuras 5.70 e 5.71 mostram o estudo de robustez da partição para o cenário tendência.



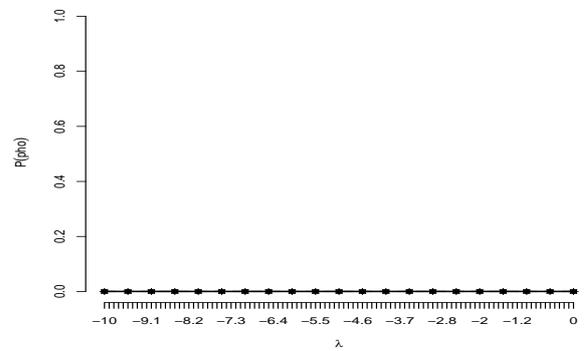
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$



(c)  $d_\rho$  com  $\lambda \leq 0$

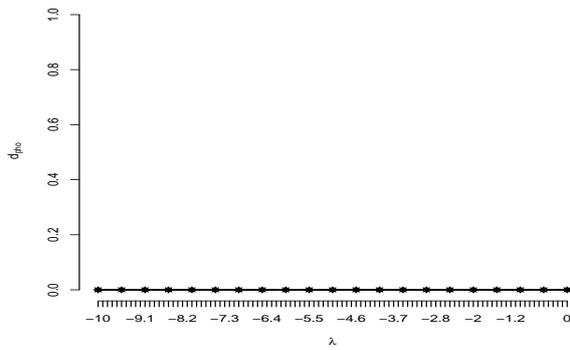


(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

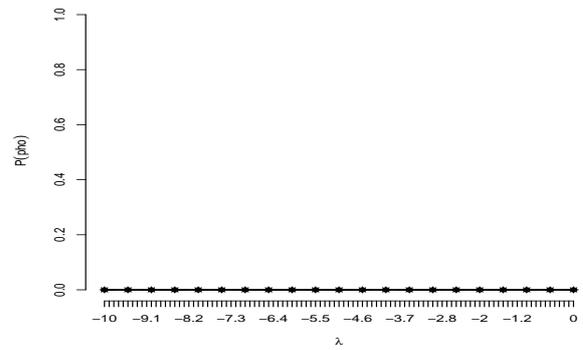
Figura 5.70: Cenário tendência - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 40$ .

No cenário tendência observa-se que mesmo com o tamanho da amostra maior o modelo não foi capaz de identificar corretamente a partição de interesse, ou seja, o mesmo comportamento obser-

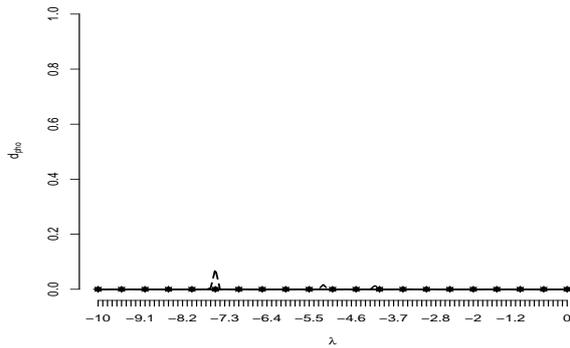
vado no estudo de robustez na distribuição *a priori* se repete no Modelo II, mas como mencionado anteriormente, para esse tipo de cenário o modelo partição produto não apresenta bom desempenho.



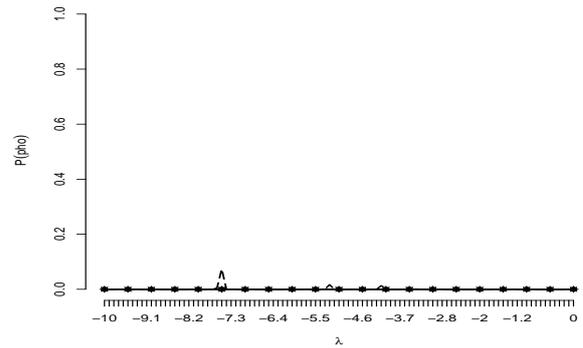
(a)  $d_\rho$  com  $\lambda \geq 0$



(b)  $P(\rho = \rho')$  com  $\lambda \geq 0$



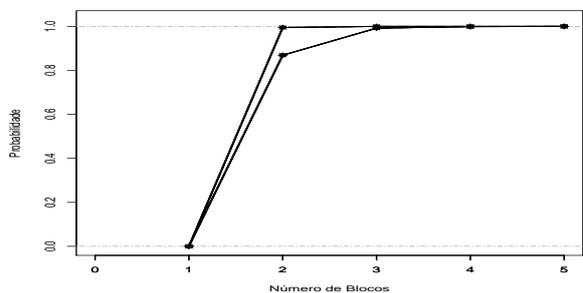
(c)  $d_\rho$  com  $\lambda \leq 0$



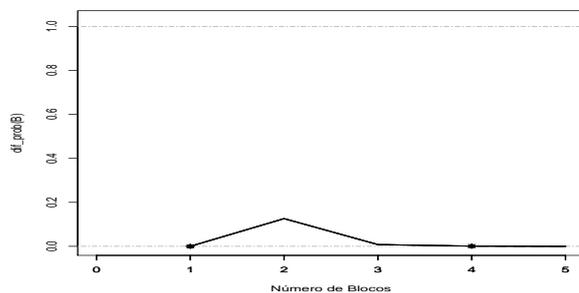
(d)  $P(\rho = \rho')$  com  $\lambda \leq 0$

Figura 5.71: Cenário tendência - Medidas de robustez para  $\rho$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ ,  $n = 120$ .

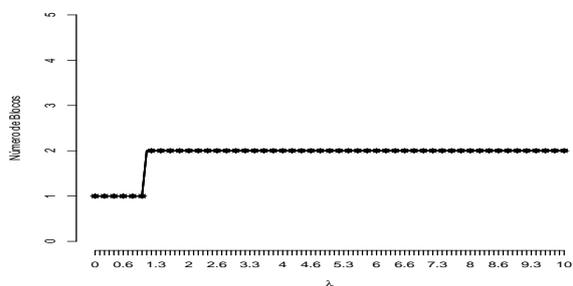
As Figuras 5.72, 5.73, 5.74 e 5.75 referem-se aos resultados para o número de blocos no cenário tendência.



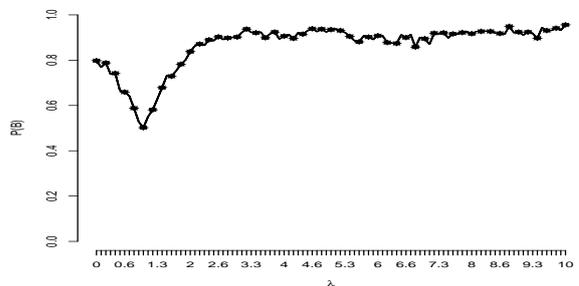
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B=b) - P_{\lambda=0}(B=b)|$

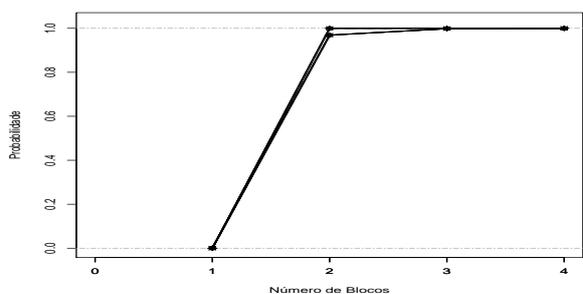


(c) Moda

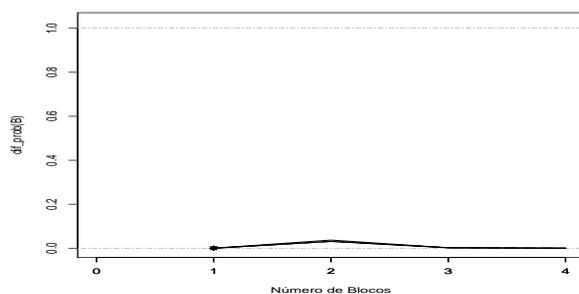


(d)  $P(B=b)$

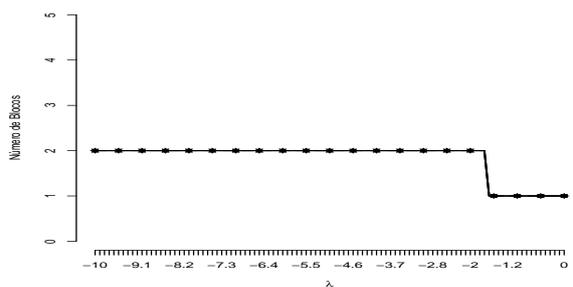
Figura 5.72:  $\lambda \geq 0$  - Cenário tendência - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 40$ .



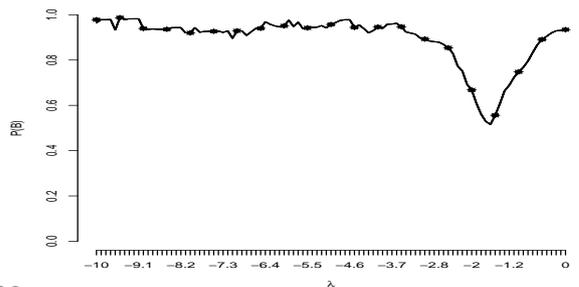
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B=b) - P_{\lambda=0}(B=b)|$

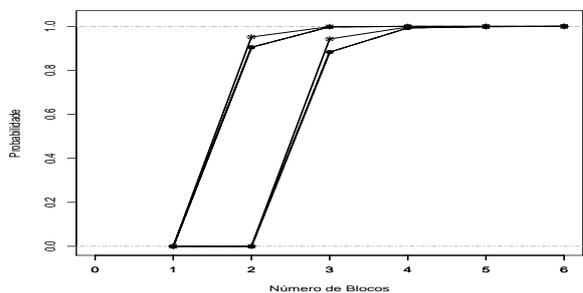


(c) Moda

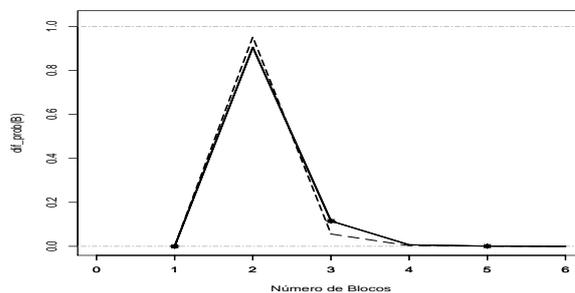


(d)  $P(B=b)$

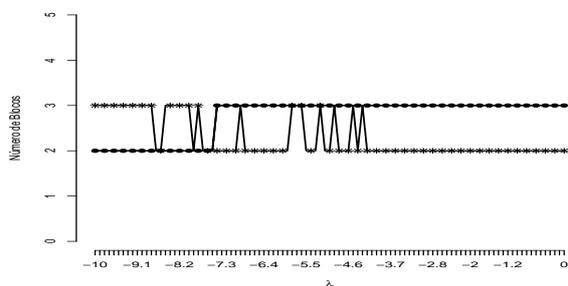
Figura 5.73:  $\lambda \leq 0$  - Cenário tendência - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 40$ .



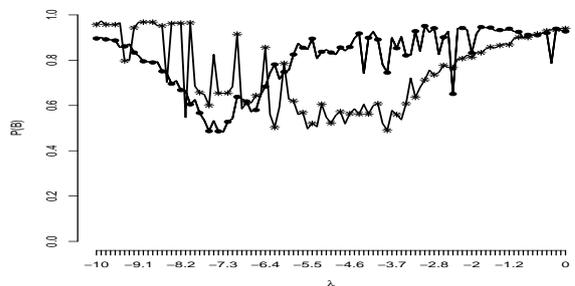
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$

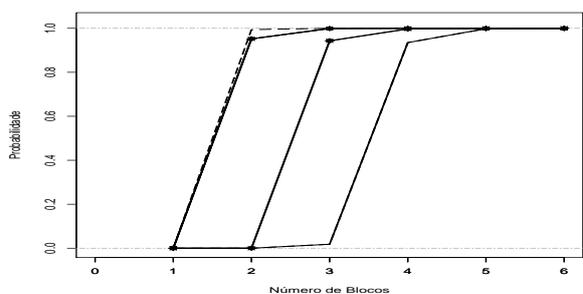


(c) Moda

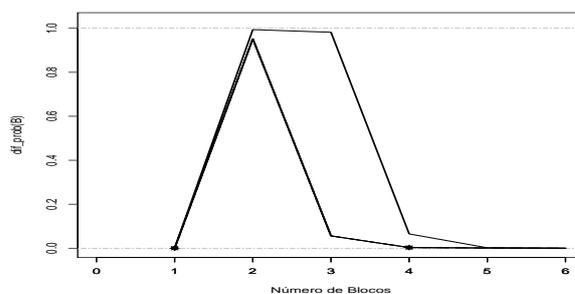


(d)  $P(B=b)$

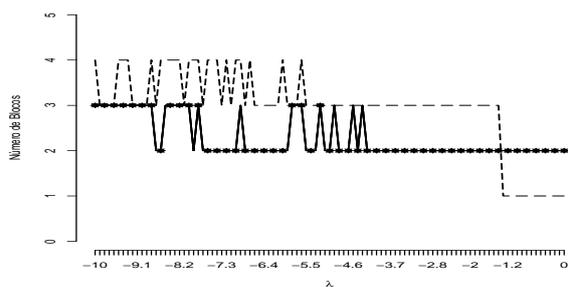
Figura 5.74:  $\lambda \geq 0$  - Cenário tendência - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 120$ .



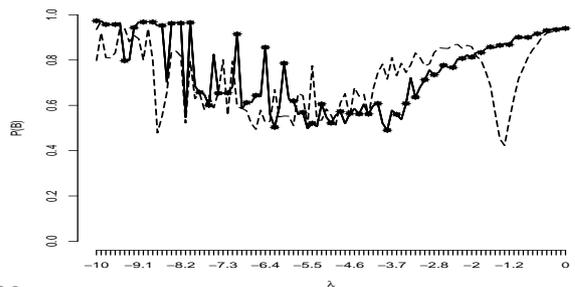
(a)  $P(B \leq b)$



(b)  $|P_\lambda(B = b) - P_{\lambda=0}(B = b)|$



(c) Moda



(d)  $P(B=b)$

Figura 5.75:  $\lambda \leq 0$  - Cenário tendência - Medidas de robustez para o número de blocos - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ ,  $n = 120$ .

Nota-se que para o tamanho de amostra 40 o número de blocos foi subestimado, no entanto observa-se que quando  $\lambda$  se afasta de zero as estimativas tendem a melhorar, mas não chegam a identificar o número correto de blocos. Para o tamanho de amostra 120 e  $\lambda$  positivo o modelo tende a acertar mais o número de blocos, mas não tem um padrão definido em relação a  $\lambda$ , nem em relação ao valor de  $\epsilon$  adicionado a série.

### 5.2.3 Robustez na probabilidade de mudança em cada instante

Finalizando o estudo de robustez, nessa seção serão mostradas como as probabilidade de mudança em cada instante são alteradas pelo parâmetro de assimetria do Modelo II.

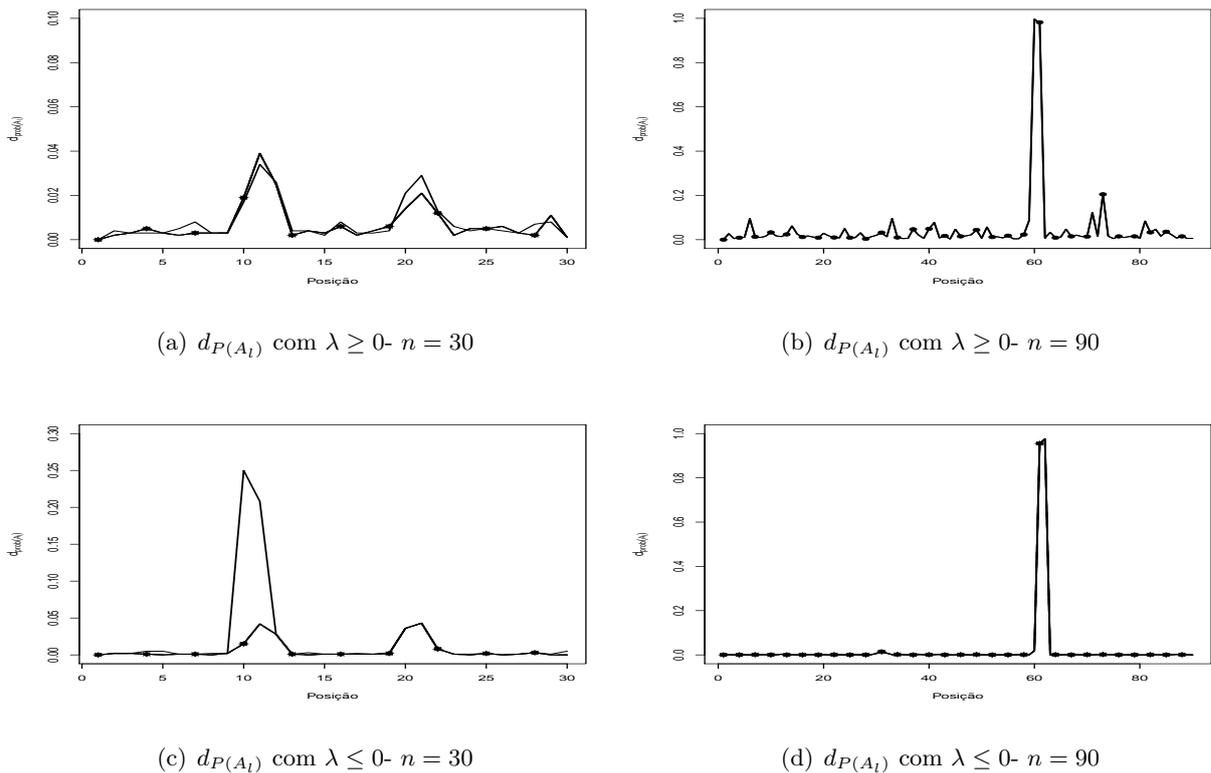


Figura 5.76: Cenário degrau positivo - Magnitude da medida  $d_{P(A_t)}$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

Observa-se na Figura 5.76 que a distância  $d_{P(A_t)}$  atingiu seu valor máximo nos instantes 61 e 62 para  $n = 90$  e  $\lambda$  positivo e negativo. Para o tamanho de amostra  $n = 30$  a maior diferença foi para  $\lambda < 0$  e  $\epsilon = 0$ , com exceção desse caso para  $n = 30$  as demais diferenças foram em torno de 0.04 próximos aos pontos de mudança.

A Figura 5.77 apresenta os resultados para o cenário degrau não simétrico.

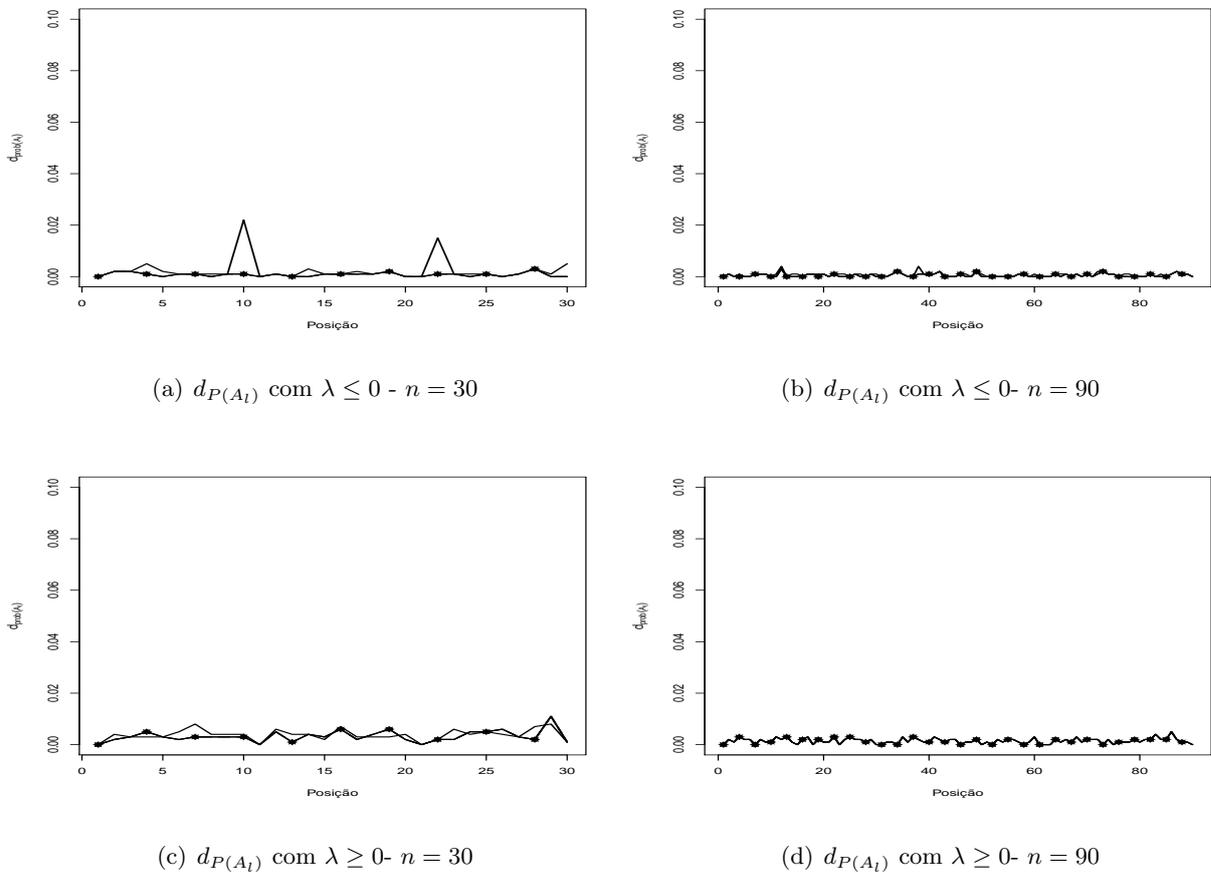


Figura 5.77: Cenário degrau não simétrico - Magnitude da medida  $d_{P(A_t)}$  - Linha cheia  $\epsilon = 0$ , linha com  $\bullet$   $\epsilon = 2$ , linha tracejada  $\epsilon = 5$ , Linha com  $*$   $\epsilon = 50$ .

As medidas  $d_{P(A_t)}$  no cenário degrau não simétrico são próximas de zero para  $\lambda > 0$  ou  $\lambda < 0$ . A maior distância observada ocorre quando  $\epsilon = 0$  e  $\lambda > 0$ , Figura 5.77 (a), no entanto essa oscilação é próxima aos pontos de mudança e em torno de 0.02.

O comportamento dos instantes das mudanças para o cenário escada crescente são apresentadas na Figura 5.78.

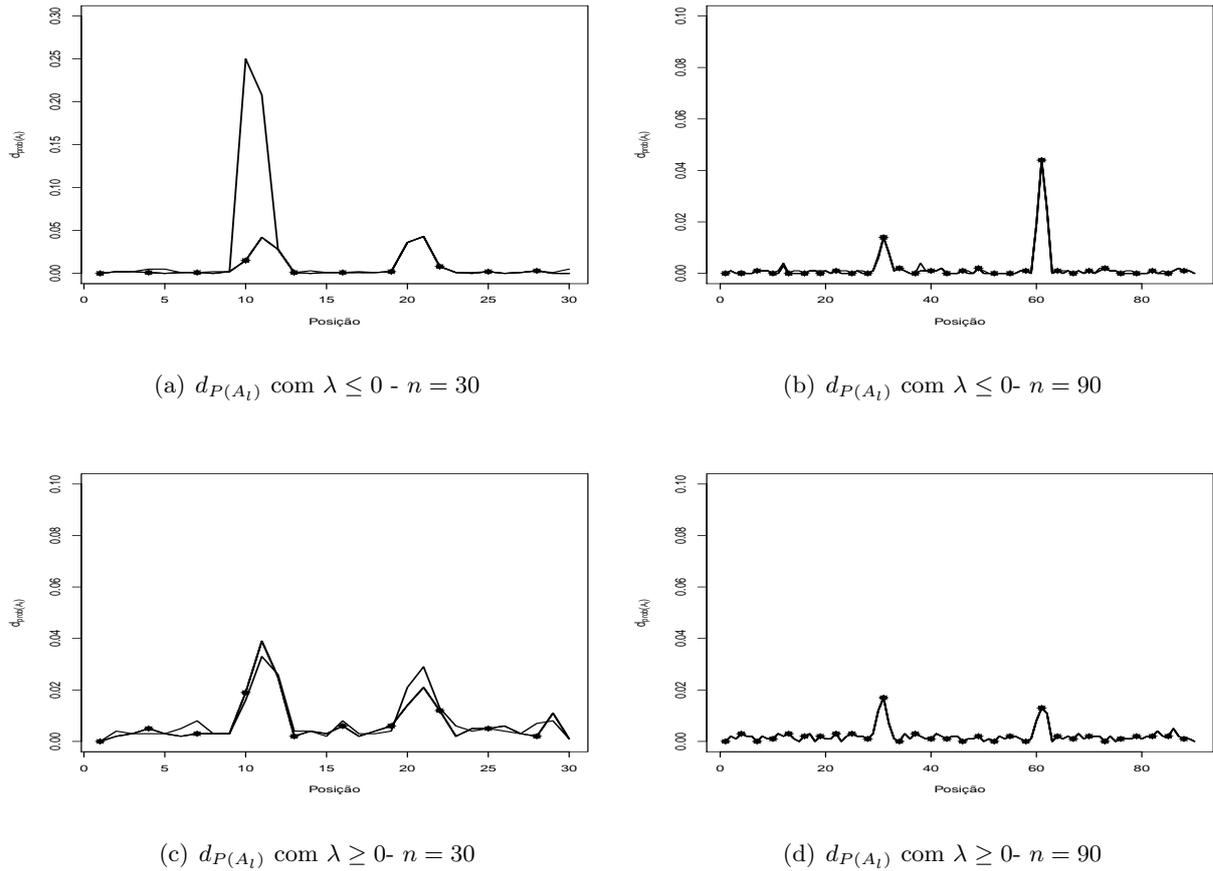
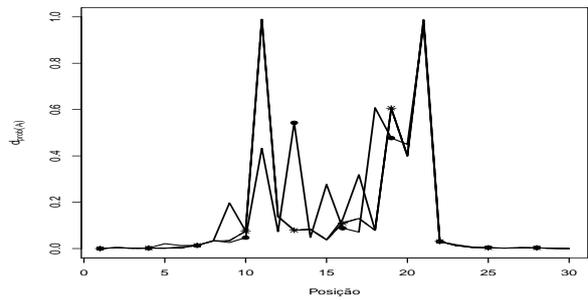


Figura 5.78: Cenário escada crescente - Magnitude da medida  $d_{P(A_t)}$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

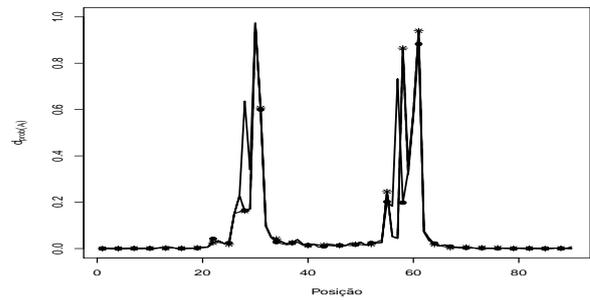
No cenário escada crescente as variações são maiores próximas aos pontos de mudança, no entanto, com exceção de  $\epsilon = 0$  para  $\lambda < 0$  as oscilações foram em torno de 0.05 para  $n = 30$  e 0.02 para  $n = 90$  exceto no instante 61 na Figura 5.78 (b).

As Figuras 5.79 e 5.80 retratam o comportamento da probabilidade de mudança em cada instante para o cenário epsilon e o cenário tendência, respectivamente.

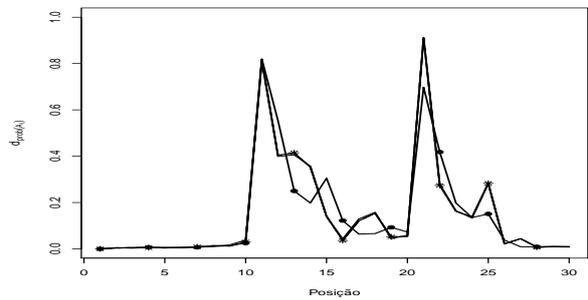
Observa-se na Figura 5.79 que o cenário epsilon tende a confundir o momento exato da mudança em virtude do pequeno salto na média de um bloco para o outro, desta forma, a magnitude da diferença atinge o seu valor máximo próximo aos pontos de mudança. Na Figura 5.80 que representa o cenário tendência, nota-se uma variação maior da medida  $d_{P(A_t)}$  entre o primeiro e o terceiro ponto de mudança, onde existe a tendência determinística.



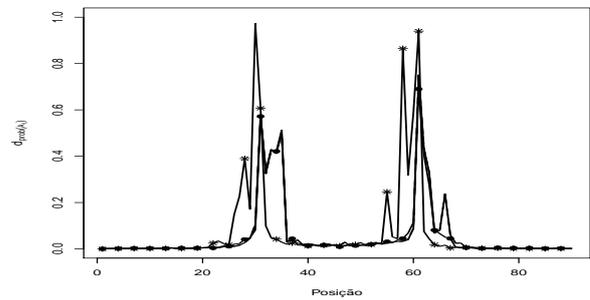
(a)  $d_{P(A_t)}$  com  $\lambda \leq 0$  -  $n = 30$



(b)  $d_{P(A_t)}$  com  $\lambda \leq 0$  -  $n = 90$

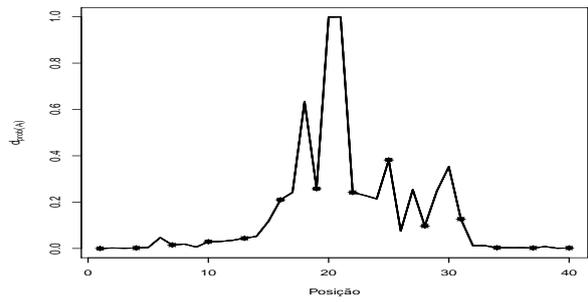


(c)  $d_{P(A_t)}$  com  $\lambda \geq 0$  -  $n = 30$

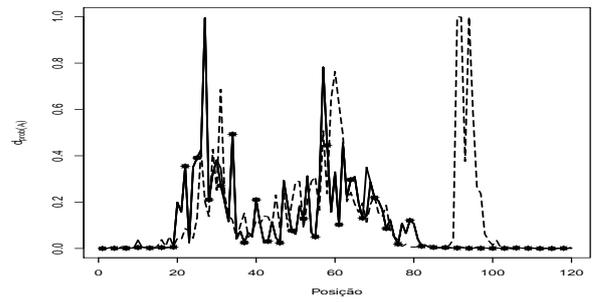


(d)  $d_{P(A_t)}$  com  $\lambda \geq 0$  -  $n = 90$

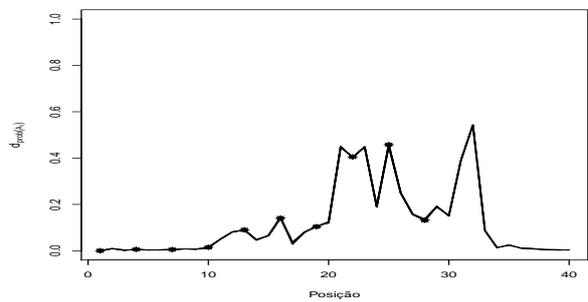
Figura 5.79: Cenário epsilon - Magnitude da medida  $d_{P(A_t)}$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .



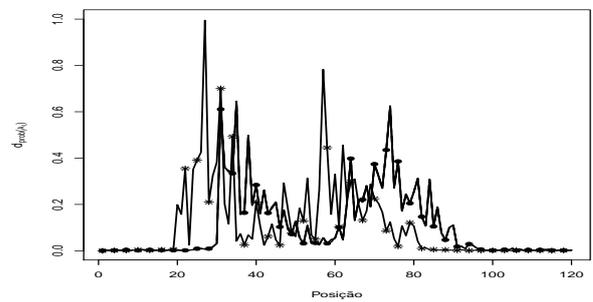
(a)  $d_{P(A_t)}$  com  $\lambda \leq 0$  -  $n = 40$



(b)  $d_{P(A_t)}$  com  $\lambda \leq 0$  -  $n = 120$



(c)  $d_{P(A_t)}$  com  $\lambda \geq 0$  -  $n = 40$



(d)  $d_{P(A_t)}$  com  $\lambda \geq 0$  -  $n = 120$

Figura 5.80: Cenário tendência - Magnitude da medida  $d_{P(A_t)}$  - Linha cheia  $\epsilon = 0$  , linha com  $\bullet$   $\epsilon = 2$  , linha tracejada  $\epsilon = 5$  , Linha com  $*$   $\epsilon = 50$ .

### 5.2.4 Considerações do estudo de robustez global no Modelo II

Como discutido anteriormente, no modelo partição produto é difícil mostrar o comportamento da média para todo  $\lambda$ , pois o número de médias é igual ao número de observações, consequentemente mostrar o comportamento de todas as médias estimadas em função do valor de  $\lambda$  avaliado seria impossível, mas para ilustrar o comportamento de algumas médias  $\mu_i$  em relação ao parâmetro de assimetria, observe a Figura 5.81 a seguir. Essa Figura mostra como as estimativas *a posteriori* de  $\mu_{30}$ ,  $\mu_{61}$  e  $\mu_{90}$  se comportam para valores de  $\lambda = (-10, 0)$  no estudo de robustez global do cenário escada crescente.

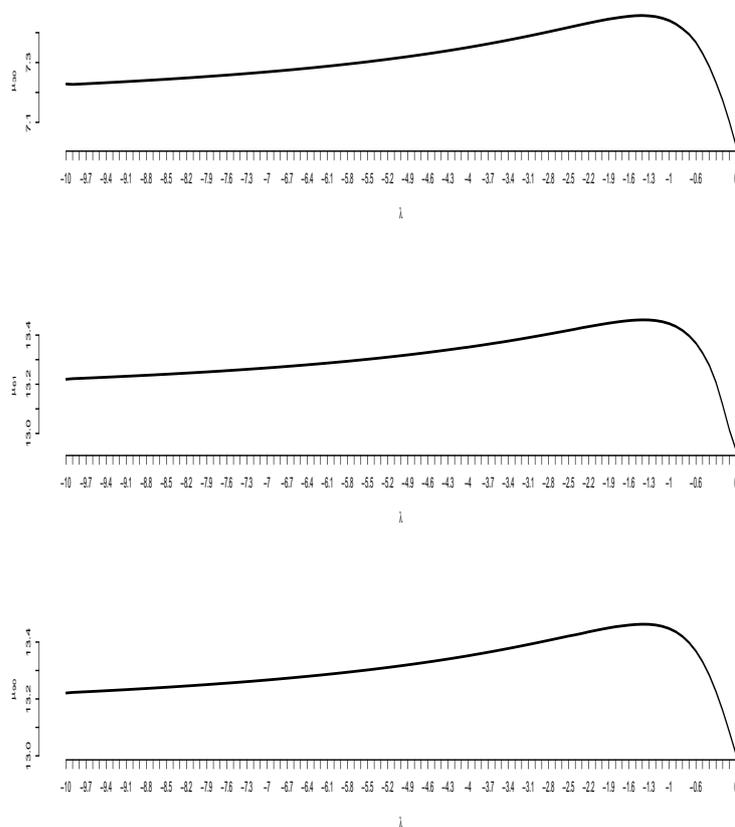


Figura 5.81: Comportamento de  $\mu_i$  no cenário escada crescente e  $\lambda \leq 0$ , para as posições 30 (topo), 61 (meio) e 90 (base).

Nota-se que a medida que  $\lambda$  se afasta de zero as estimativas produto são superestimadas, carac-

terizando uma perda nas estimativas produto a medida que  $\lambda$  se distancia de zero. No entanto, na análise de robustez global considerando a contaminação na função de verossimilhança observou-se um modesto ganho na estimativa da partição de interesse ao considerar assimetria.

Apesar de não ser possível mostrar as estimativas para todo  $\mu_i$ , vale ressaltar que o comportamento apresentado na Figura 5.81 se aplica para as demais médias e demais cenários avaliados no estudo de robustez global. Uma possível explicação para esse fenômeno está na distribuição condicional completa de  $\mu_i$

$$\mu_i | \mathbf{Y}, \mathbf{T}, \rho, p \sim \mathcal{N} \left( \frac{\sigma_\mu^2 \sum_{l=i_j+1}^{i_{j+1}} (y_l - \delta \sigma_y |t_l|) + \sigma_y^2 \mu_0 (1 - \delta^2)}{\sigma_\mu^2 (i_{j+1} + 1 - i_j) + \sigma_y^2 (1 - \delta^2)}; \frac{\sigma_y^2 \sigma_\mu^2 (1 - \delta^2)}{\sigma_\mu^2 (i_{j+1} + 1 - i_j) + \sigma_y^2 (1 - \delta^2)} \right).$$

Note que a variância da distribuição condicional completa de  $\mu_i$  é

$$\frac{\sigma_y^2 \sigma_\mu^2 (1 - \delta^2)}{\sigma_\mu^2 (i_{j+1} + 1 - i_j) + \sigma_y^2 (1 - \delta^2)} \quad (5.1)$$

Conseqüentemente, a perda observada na Figura 5.81 ocorre devido a variância ser menor para valores de  $\lambda$  maiores em valor absoluto, pois a medida que  $\lambda$  aumenta a variância dada em 5.1 tende a zero. Além disso, a primeira e a segunda derivada da expressão 5.1 em relação a  $\delta$  são dadas, respectivamente, por

$$\frac{d'}{d\delta} = \frac{-2 * \delta \sigma_y^2 \sigma_\mu^2 (i_{j+1} - i_j + 1)}{(\sigma_\mu (i_{j+1} - i_j + 1) + \sigma_y^2 (\sigma_y^2 \sigma_\mu^2 (1 - \delta^2)))^2}, \quad (5.2)$$

$$\begin{aligned} \frac{d''}{d\delta} &= \frac{-2\sigma_y^2 \sigma_\mu^2 \sigma_\mu^2 (i_{j+1} - i_j + 1) (\sigma_\mu^2 (i_{j+1} - i_j + 1) + \sigma_y^2 (\sigma_y^2 (1 - \delta^2)))^2}{(\sigma_\mu^2 (i_{j+1} - i_j + 1) + \sigma_y^2 (1 - \delta^2))^2} \\ &\times \frac{-2(\sigma_\mu (i_{j+1} - i_j + 1) + \sigma_y^2 (1 - \delta^2)) (-2\delta \sigma_y^2)}{(\sigma_\mu^2 (i_{j+1} - i_j + 1) + \sigma_y^2 (1 - \delta^2))^2}. \end{aligned} \quad (5.3)$$

Observe que a primeira derivada é zero se, e somente se,  $\delta = 0$  e o sinal da segunda derivada é negativo, indicando que o ponto de máximo da variância da condicional completa de  $\mu_i$  ocorre

quando  $\delta = 0$ , ou seja, quando  $\lambda = 0$ . Como ilustra-se na Figura 5.82.

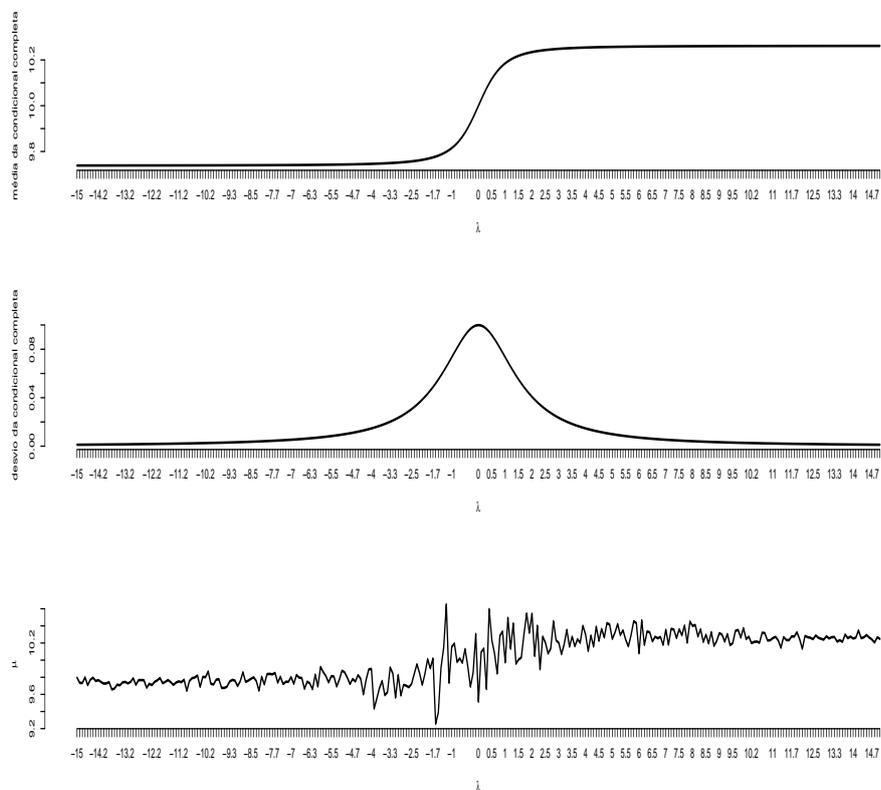


Figura 5.82: Comportamento de  $m^*$  (topo),  $var^*$  (centro) e  $\mu_i$  (base). Para um bloco fixo e observações fixadas em 10.

Contudo, como a variância tende a ser próxima de zero, as estimativas das médias serão extremamente centradas na média do bloco e a variabilidade entre um bloco e outro tende a ser menor o que favorece para que a partição seja estimada com alta probabilidade. No entanto, existe uma perda das estimativas produto.

## 5.3 Análise Crítica do Estudo de Robustez

### 5.3.1 Robustez no Modelo I

No cenário degrau positivo a variação maior da medida  $d_M$  é em torno do ponto de mudança, no entanto, ainda é abaixo de 0.3 para  $n = 30$ , quando o tamanho de amostra aumenta, a magnitude da diferença fica em torno de 0.2 para  $\lambda \geq 0$ , mas se  $\lambda$  é negativo as distâncias chegam a 0.4 e a variabilidade das estimativas aumenta próximo aos ponto de mudança chegando a 0.6 para  $\lambda < 0$  e 0.3 para  $\lambda > 0$ . No caso do cenário degrau não simétrico, como o tamanho do salto é maior, a magnitude das distâncias foi praticamente zero para todos os valores de  $\lambda$ .

Para o cenário escada crescente se  $\lambda \leq 0$  há um pico maior na magnitude da diferença do segundo ponto de mudança, e se  $\lambda \geq 0$  a maior diferença ocorre no primeiro ponto de mudança, o mesmo comportamento é observado para a variabilidade das estimativas de  $\mu_i$ . Nos cenários epsilon e tendência, as medidas  $d_M$  apresentaram variações baixas, mas no estudo de sensibilidade foi possível avaliar que as estimativas de  $\mu_i$  ficam comprometidas na maioria dos casos avaliados para esses cenários.

Em relação a partição de interesse, no cenário degrau não simétrico, para  $n = 30$  à medida que  $\lambda$  se afasta de zero há uma ligeira perda na probabilidade de estimar a partição correta, no entanto, para  $n = 90$  essa perda é quase imperceptível. No cenário escada crescente, independente do tamanho da amostra e independente do sinal do parâmetro de assimetria, a medida que  $\lambda$  se afasta de zero a probabilidade do modelo indicar corretamente a partição de interesse diminui cerca de 11%, conseqüentemente, a distância  $d_p$  aumenta.

No caso do degrau positivo, quando a especificação *a priori* está em desacordo com a verossimilhança, a estimação da partição de interesse fica comprometida. A medida que  $\lambda$  se afasta de zero a probabilidade do modelo com  $\lambda = -30$  indicar a partição correta é cerca de 15% menor que a probabilidade indicada quando  $\lambda = 0$ , considerando ambos os tamanhos de amostras avaliados. No entanto, esse fenômeno também ocorre quando  $\lambda \geq 0$ , ou seja, quando a distribuição *a priori* está em acordo com a verossimilhança, mas nesse caso a perda é menor, e está em torno de 4% quando

comparada com  $\lambda = 0$ .

Nos cenários epsilon e tendência, o modelo não foi capaz de identificar a partição de interesse e o número de blocos tende a ser subestimado para todo  $\lambda$ . No entanto, o aumento do tamanho da amostra melhora as estimativas, mas não o suficiente para estimar corretamente o número de blocos. Logo, nos cenários degrau positivo, escada crescente e degrau não simétrico a maior diferença entre as funções de distribuição ocorrem para  $B = 3$ , sendo que para  $n = 90$  no cenário degrau positivo com  $\lambda \geq 0$  a diferença foi quase nula, mas nos cenários tendência e epsilon a maior diferença ocorre em  $B = 2$ , devido a subestimação do número de blocos nesses cenários.

A diferença entre a probabilidade de mudança em cada instante no cenário degrau não simétrico foi praticamente nula, mesmo próximo aos pontos de mudança. No cenário degrau positivo, nota-se que as diferenças foram menores para  $\lambda \geq 0$ , neste caso as maiores diferenças ocorrem próximas aos pontos de mudança e ficaram próximas a 9%, já para  $\lambda \leq 0$  essas diferenças foram maiores, em torno de 15%. No cenário escada os picos ocorrem no primeiro ponto de mudança para  $\lambda \geq 0$  e no segundo ponto de mudança para  $\lambda \leq 0$ , a magnitude da diferença nesses picos é em torno de 20%.

Nos cenários tendência e epsilon, as diferenças da medida  $d_P(A_l)$  foram inferiores a 0.05 para a maioria dos pontos de mudança, mas lembrando que o número de blocos e partição de interesse não foram estimados de forma satisfatória nesses cenários.

Desta forma, é possível avaliar por meio desse estudo de robustez que o ganho em considerar a assimetria na distribuição *a priori* das médias praticamente não existe, e se levado em consideração o custo computacional por trás do modelo, a conclusão que se chega é que não vale a pena incluir a assimetria. Além disso, nos casos dos cenários degrau positivo, degrau não simétrico e escada crescente, o que ocorreu foi uma perda em estimar a partição de interesse a medida que  $\lambda$  se afasta de zero. Nos cenários epsilon e tendência, já era esperado que as estimativas ficassem comprometidas pela forma como esses cenários são constituídos, mas nota-se que assumir assimetria não trouxe nenhum ganho para esses cenários.

### 5.3.2 Robustez no Modelo II

Em relação a média, no cenário degrau positivo, nota-se que as diferenças ficam em torno de 0.5, com exceção de um pico no ponto 11, para  $n = 30$  e  $\lambda < 0$  e no ponto 61 para  $n = 90$  e  $\lambda \leq 0$ , para  $\lambda \geq 0$  a variação também é em torno de 0.5, mas sem picos. No cenário degrau não simétrico, a magnitude das mudanças também ficaram em torno de 0.5, mas sem picos próximos aos pontos de mudança, independente se  $\lambda$  é negativo ou positivo. No cenário escada crescente, as diferenças ficaram em torno de 0.5, com exceção de um pico nas diferenças do primeiro ponto de mudança para  $n = 30$  e  $\lambda \leq 0$ , mas o pico aconteceu para  $\epsilon = 0$ .

Nos cenários epsilon e tendência, a magnitude da medida  $d_M$  foi maior que 0.5, mas abaixo de 1, no entanto, nesses cenários a diferença entre os blocos as vezes é inferior a 1 o que indica que a medida  $d_M$  nesses casos pode ser considerada grande. Além disso, nota-se que a quantidade  $\epsilon$  adicionada influenciou nas estimativas, principalmente se o tamanho da amostra aumenta.

Em relação a partição de interesse, no cenário degrau positivo, se  $\lambda > 0$  a probabilidade da partição de interesse ser indicada corretamente aumenta 10% a medida que  $\lambda$  se afasta de zero, mas se  $\lambda$  é negativo o contrário acontece e a probabilidade diminui 10% quando  $\lambda$  se afasta de zero. No cenário degrau não simétrico, a partição de interesse é estimada com um ganho de 3% a 5% quando  $\lambda$  se afasta de zero, independente do tamanho da amostra e do sinal do parâmetro de assimetria. No cenário escada crescente o comportamento é similar, no entanto, o ganho é em torno de 10%. Mas no cenário tendência nenhum valor de  $\lambda$  indicou a partição de interesse corretamente e no cenário epsilon a partição não apresenta um comportamento bem definido, ou seja, ora  $\lambda$  distante de zero produz um ganho nas estimativas, ora  $\lambda$  próximo a zero estima melhor a partição.

O mesmo comportamento, observado para o número de blocos no estudo de robustez no Modelo I é observado no Modelo II. A maior variação para o número de blocos ocorre para  $B = 3$ , variando em até 5% a probabilidade do modelo indicar corretamente  $B = 3$ , com exceção dos cenários epsilon e tendência que tendem a subestimar o número de blocos e, conseqüentemente, nesses cenários a maior diferença ocorre para  $B = 2$ .

A probabilidade de mudança em cada instante no cenário degrau não simétrico apresenta diferenças inferiores a 2% para  $n = 30$  e inferior a 0.5% para  $n = 90$ . No cenário degrau positivo as maiores variações na probabilidade de mudança em cada instante ocorrem quando  $n = 30$ ,  $\epsilon = 0$  e o parâmetro de assimetria é negativo, nos demais casos a variação é inferior a 5% para  $n = 30$ . Já para  $n = 90$  a variação chegou a 100% nos pontos de mudança 60 e 61, essa variação ocorreu porque o modelo ficou em dúvida entre o instante correto da mudança, ou seja, alguns valores de  $\lambda$  indicaram 60 e o correto era 61, por isso a diferença de 100%.

No cenário escada crescente a variação da probabilidade de mudança em cada instante é maior perto dos pontos de mudança. Para  $\lambda$  negativo a variação é menor que 5% para ambos os tamanhos de amostra, mas houve um pico para  $\epsilon = 0$ . No caso em que  $\lambda$  é positivo as variações são menores que 5% para  $n = 30$  e menores que 2% para  $n = 90$ . Mas o comportamento para os cenários epsilon e tendência é bem diferente, indicando grande variação da medida  $d_{P(A_t)}$  e indicando que não há robustez nesses cenários.

Portanto, o que observa-se nesse estudo de robustez no Modelo II é que se o objetivo for estimar a partição e o número de blocos há um ganho significativo em considerar a assimetria no modelo, mas não em todos os cenários, apenas nos cenários escada crescente, degrau não simétrico e no degrau positivo assumindo  $\lambda$  positivo. No entanto, o custo computacional para obter essa melhoria é relativamente alto. Logo, cabe ao pesquisador definir seus objetivos e tempo disponível para optar entre qual modelo ele utilizará.

O próximo capítulo apresenta um estudo de sensibilidade aplicado as séries do preço médio do Kg do tomate e dos retornos mensais do Ibovespa.

Neste capítulo são apresentadas e discutidas duas aplicações a dados reais dos modelos descritos nas Seções 2.4, 2.5. O objetivo é avaliar, na prática, qual modelo se adequa melhor a realidade dos dados através de um estudo de sensibilidade informal e o uso do *Deviance Information Criterion* (DIC). A primeira aplicação utiliza a série temporal mensal do preço médio do quilo do tomate no período que se estende de 1992 a 2009. A segunda aplicação refere-se ao retorno mensal do índice da bolsa de valores de São Paulo (Ibovespa) no período de abr/94 a jun/14.

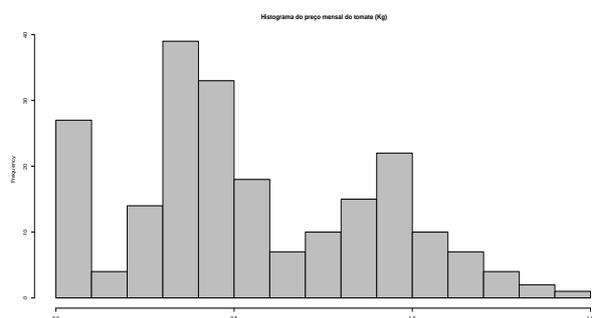
### 6.1 Série do Preço Médio do Tomate(Kg)

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.

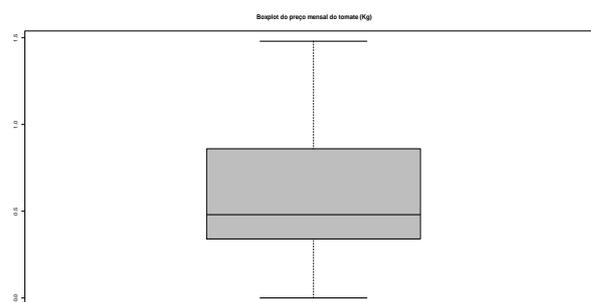
Optou-se por realizar essa aplicação a série do preço médio do tomate, pois segundo Perron (1990), identificar se a série possui mudança estrutural na média é importante, uma vez que a presença de quebra estrutural no nível da série pode levar a inferir erroneamente sobre a presença de uma raiz unitária, quando de fato a série em estudo não apresenta esta característica. A série utilizada contempla até 2009 por questões técnicas, uma vez que a série atualizada até 2014 não é mais

gratuita no site da FGVDados. A seguir serão apresentadas algumas estatísticas descritivas desse conjunto de dados.

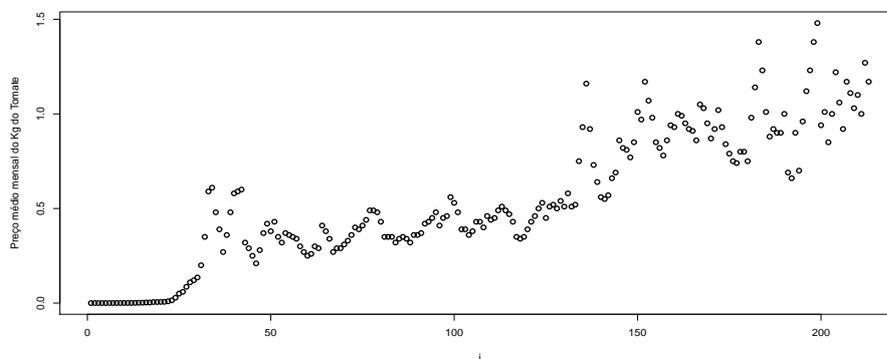
A Figura 6.1 mostra a distribuição dos dados. Analisando os gráficos percebe-se que a distribuição apresenta assimetria a direita, nenhum *outlier*, e multimodalidade. Tais características sugerem que uma mistura de distribuições assimétricas pode ser uma escolha razoável para modelar estes dados.



(a) Histograma do preço médio mensal do tomate



(b) Boxplot do preço médio mensal do tomate



(c) Série do preço médio mensal do tomate

Figura 6.1: Distribuição do preço mensal do Tomate (Kg)

A Tabela 6.1 apresenta as medidas resumo da série do tomate de onde percebe-se que a média é maior que a mediana e o desvio padrão é de 0.3508 o que resulta em um coeficiente de variação de 0.638.

Tabela 6.1: Estatísticas Descritivas - Série mensal do preço do tomate

Min	Q1	Mediana	Q3	Max	Média	Desvio padrão
0.0001	0.3400	0.4800	0.8600	1.480	0.5502	0.3508

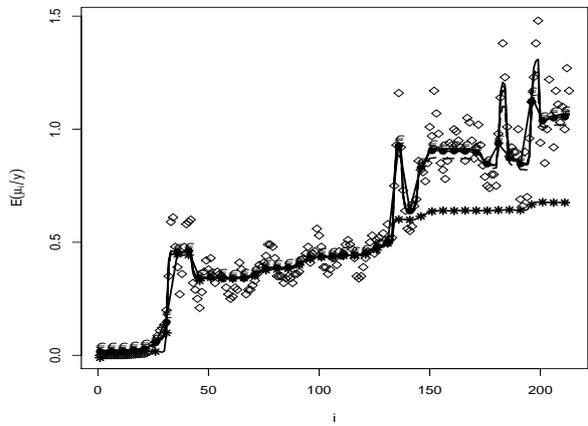
A seguir os dados serão modelados via Modelo I, assumindo assimetria da média, e pelo Modelo II, assumindo assimetria dos dados.

### 6.1.1 Aplicação - Modelo I - *priori* normal assimétrica

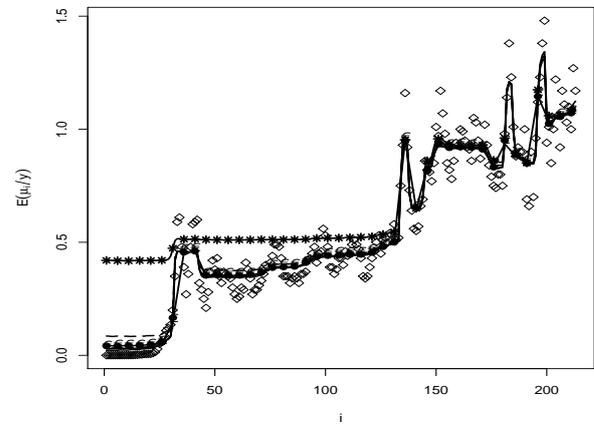
A análise realizada nesta seção pressupõe que o modelo correto assume assimetria *a priori* para a média, ou seja, considera-se que o preço do tomate no mês  $i$ , dado  $\mu_i$  segue uma distribuição normal com média  $\mu_i$  e variância  $\sigma_y^2 = 0.135$ , esse valor da variância foi estimado em Ferreira (2010). Assume-se que  $\mu_{i_j i_{j+1}} \sim \mathcal{SN}(\mu_0, \sigma_\mu^2, \lambda)$  e os valores de  $\lambda$  avaliados foram  $\lambda = (-10, -2, -1, -0.7, -0.5, 0, 0.5, 0.7, 1, 2, 10)$ . Os hiperparâmetros  $\mu_0$  e  $\sigma_\mu^2$  foram considerados fixos e iguais a média dos dados e 10, respectivamente. Assume-se *a priori*, que  $p \sim \text{Beta}(1.1, 1.1)$ . Para o MCMC o período de aquecimento da cadeia foi de 30.000, o lag foi de tamanho 10 e o tamanho da amostra de 1000.

A Figura 6.2 mostra as observações e as estimativas produto obtidas para os valores de  $\lambda = (-10, -2, -1, -0.7, -0.5, 0, 0.5, 0.7, 1, 2, 10)$ , os gráficos da esquerda ilustram os resultados para os valores de  $\lambda = (-10, -2, -1, -0.7, -0.5, 0)$  e os gráficos da direita mostram os resultados para  $\lambda = (0, 0.5, 0.7, 1, 2, 10)$ . Já os gráficos (c) e (d) da Figura 6.2 exibem as estimativas da variância *a posteriori* para cada  $\mu_i$ .

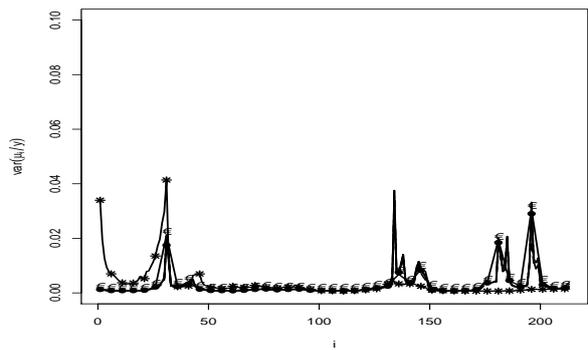
Da Figura 6.2 percebe-se que a escolha do valor de  $\lambda$  não influencia de forma expressiva nas estimativas *a posteriori*, apenas para  $\lambda = \pm 10$  notou-se uma diferença razoavelmente grande nas estimativas produto. Para as estimativas produto, nota-se também que o valor de  $\lambda$  não afeta a variância *a posteriori* dos  $\mu'_i$ s, ou seja, as variâncias *a posteriori* são muito similares para todos os valores de  $\lambda$  em todos os instantes e próximas de zero entre os instantes 50 – 130 e 150 – 175.



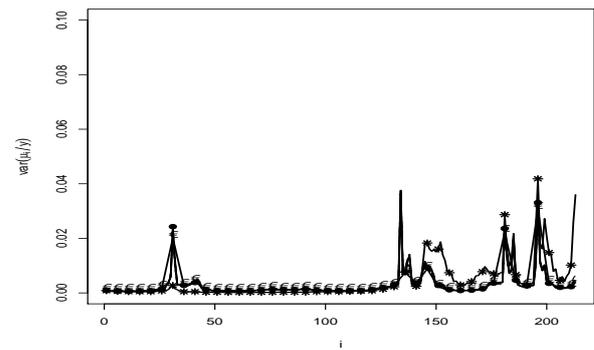
(a) Preço venda tomate(Kg) em reais -  $\lambda \leq 0$



(b) Preço venda tomate(Kg) em reais -  $\lambda \geq 0$



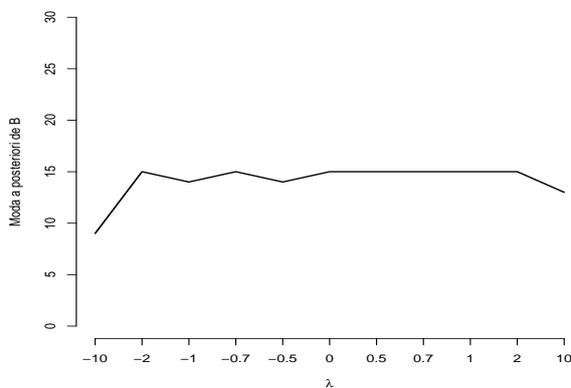
(c)  $\text{Var}(\mu_i) - \lambda \leq 0$



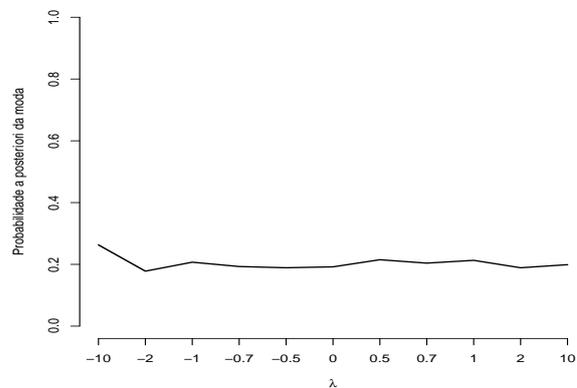
(d)  $\text{Var}(\mu_i) - \lambda \geq 0$

Figura 6.2: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Dados  $\diamond$ , estimativas para  $\lambda = \pm 10$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 1$  (linha com  $\bullet$ ),  $\lambda = \pm 0.7$  (linha pontilhada e tracejada),  $\lambda = \pm 0.5$  (linha com EUROSIMBOLO) e  $\lambda = 0$  (linha cheia).

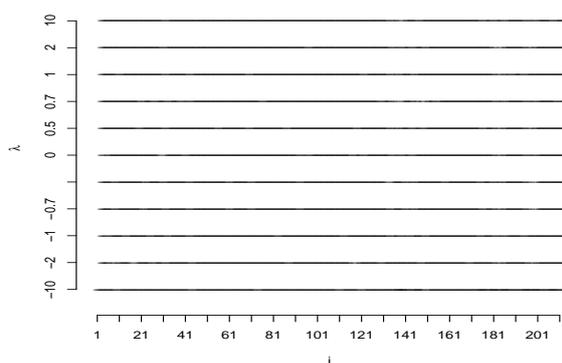
Objetiva-se ilustrar na Figura 6.3(a) o número de blocos estimados para cada valor de  $\lambda = (-10, -2, -1, -0.7, -0.5, 0, 0.5, 0.7, 1, 2, 10)$ , a probabilidade do modelo indicar  $B = b$  para cada  $\lambda$  é ilustrado na Figura 6.3(b). A partição mais provável para cada  $\lambda$  é indicada na Figura 6.3(c) e a probabilidade com que a partição mais provável foi indicada pelo modelo para cada  $\lambda$  é dada na Figura 6.3(d) e a Figura 6.4(a) mostra a probabilidade de cada instante ser um ponto de mudança para os valores de  $\lambda$  negativos e a Figura 6.4(b) ilustra para os valores de  $\lambda$  positivos.



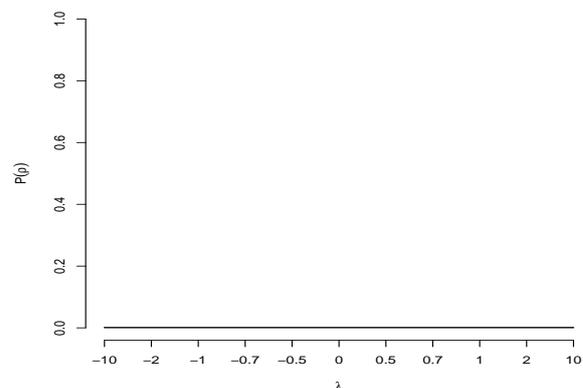
(a) Número de Blocos



(b)  $P(B = b)$



(c) Partição mais provável



(d)  $P(\rho)$

Figura 6.3: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ .

Note, na Figura 6.3, que o número de blocos foi próximo de 14 para os valores de  $\lambda$  avaliados, exceto para  $\lambda = -10$  onde a moda *a posteriori* de  $B$  foi 9 e todos os modelos indicaram probabilidade próxima de 0.21 para o número de blocos indicado. A partição mais provável foi indicada com probabilidade baixa, cerca de 0.001 para todos os modelos. Apesar do número de blocos indicado ser próximo de 14, nota-se na Figura 6.4 que cerca de 9 pontos foram indicados como pontos de mudança com uma probabilidade acima de 0.40 e 7 pontos acima de 0.5.

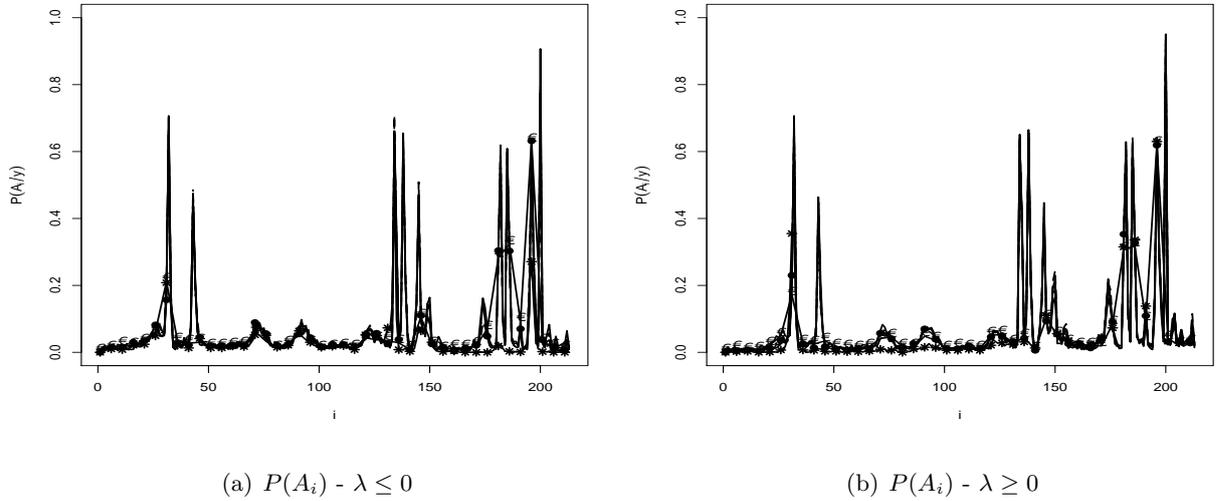


Figura 6.4: Probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 10$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 1$  (linha com •),  $\lambda = \pm 0.7$  (linha pontilhada e tracejada),  $\lambda = \pm 0.5$  (linha com EUROSIMBOLO) e  $\lambda = 0$  (linha cheia).

### DIC no Modelo Partição Produto

Para avaliar qual modelo se adequa melhor aos dados nesse estudo de aplicação, a medida de seleção de modelo Deviance Information Criterion (DIC) foi calculada. O DIC foi proposto por Spiegelhalter et al. (2002) e é definido como

$$\begin{aligned}
 DIC &= -2 \log p(y|E(\theta|y)) + 2(E(-2 \log p(y|\theta)|y) + 2 \log p(y|E(\theta|y))) \\
 &= 2(E(-2 \log p(y|\theta)|y) + \log p(y|E(\theta|y))).
 \end{aligned} \tag{6.1}$$

Assumindo que  $p(y|\theta)$  é a função de verossimilhança do modelo e que  $\theta = (\rho, \mu)$ ,  $E(-2 \log p(y|\rho, \mu)|y)$  pode ser reescrito como

$$\begin{aligned}
 E(-2 \log p(y|\rho, \mu)|y) &= \sum_{\rho} \int_{\Theta_{\mu}} -2 \log p(\mathbf{y}|\rho, \mu) \pi(\mu|\rho, \mathbf{y}) \pi(\rho|\mathbf{y}) d_{\mu} \\
 &= \sum_{\rho} \pi(\rho|\mathbf{y}) \left[ \int_{\Theta_{\mu}} -2 \log p(\mathbf{y}|\rho, \mu) \pi(\mu|\rho, \mathbf{y}) d_{\mu} \right] \\
 &= \sum_{\rho} \pi(\rho|\mathbf{y}) E(-2 \log p(\mathbf{y}|\rho, \mu))
 \end{aligned} \tag{6.2}$$

Desta forma, como no modelo partição produto não se tem uma única média e a soma sobre todas as partições ( $\sum_{\rho}$ ) deve ser calculada, faz-se necessário que a parcela  $E(-2 \log p(y|\rho, \mu)|y)$  seja calculada para cada amostra do algoritmo MCMC de forma que a soma para todas as partições seja incluída no cálculo do DIC.

Por outro lado, a parcela  $\log p(y|E(\rho, \mu|y))$  é mais simples, pois se refere a função de verossimilhança aplicada no ponto  $E(\rho, \mu|y)$  e da equação (2.6) para a média *a posteriori* do modelo partição produto, tem-se que

$$E(\mu_i|y_1, \dots, y_n) = \sum_{i_j=0}^{k-1} \sum_{i_{j+1}=i}^n r_{[i_j i_{j+1}]}^* E(\mu_i|y_{[i_j i_{j+1}]}) \tag{6.3}$$

que por sua vez já contempla a soma sobre todas as partições, finalizando o cálculo do DIC.

Os resultados do DIC para a aplicação a série do tomate encontram-se na Tabela 6.2.

Tabela 6.2: Valores do DIC para a Série mensal do preço do tomate

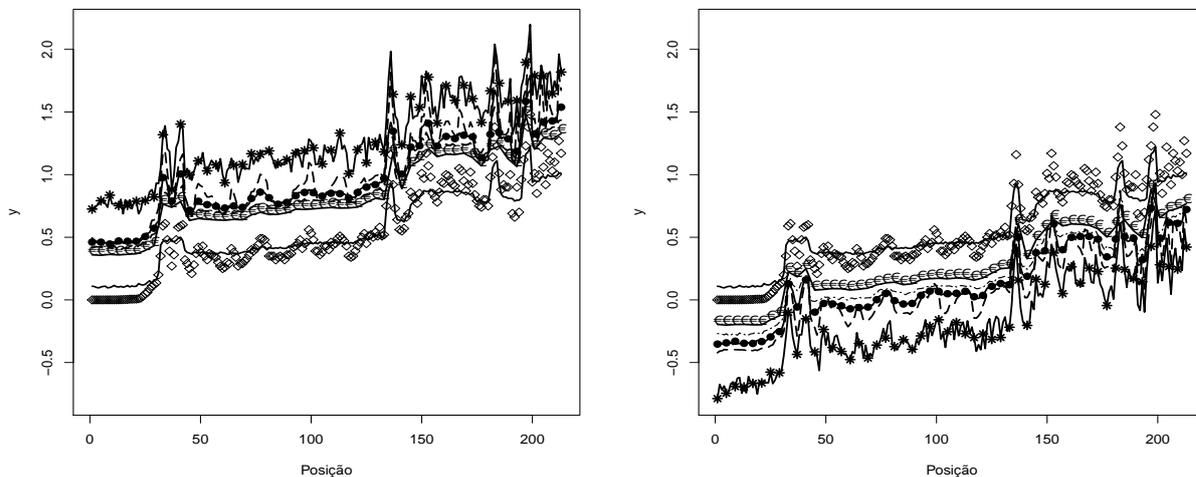
$\lambda \leq 0$					
-10	-2	-1	-0.7	-0.5	0
295.5	-332.6	-349.4	-349.7	-351.1	-352.4
$\lambda \geq 0$					
0.5	0.7	1	2	10	-
<b>-353.7</b>	-352.2	-351.5	-341.1	162.9729	-

Observa-se que os DICs são próximos, mas o modelo com  $\lambda = 0.5$  resultou no menor valor de DIC. No entanto, a partição mais provável foi indicada com probabilidade próxima de 0.001 para todos os valores de  $\lambda$ .

Apesar do DIC indicar  $\lambda = 0.5$ , as estimativas utilizando  $\lambda = 0$  são similares as estimativas assumindo  $\lambda \neq 0$ , exceto para  $\lambda = \pm 10$ , mas como visto no Capítulo 2 assumir assimetria para as médias torna o modelo mais complexo e, conseqüentemente, o custo computacional é mais alto. Assim sendo, do ponto de vista computacional, a utilização do modelo normal para identificação de pontos de mudança pode-se tornar uma opção mais vantajosa em uma aplicação. A seguir é apresentado o Modelo II ajustado a série do tomate.

### 6.1.2 Aplicação - Modelo II - dados normais assimétricos

A aplicação realizada nessa seção pressupõe que o modelo correto deve assumir assimetria nos dados. Os valores de  $\lambda$  avaliados foram os mesmo da seção anterior,  $\lambda = (-10, -2, -1, -0.7, -0.5, 0, 0.5, 0.7, 1, 2, 10)$ . O valor de  $\sigma_y^2$ ,  $\sigma_\mu^2$  e  $\mu_0$  foram os mesmos também. A distribuição *a priori* de  $\mu_{i_j i_{j+1}} \sim \mathcal{N}(\mu_0, \sigma_\mu^2)$  e a distribuição *a priori* para  $p$  foi mantida como sendo a mesma usada na Seção 6.1.1. O período de aquecimento da cadeia foi de 60.000, o lag foi de tamanho 10 e o tamanho da amostra de 1000. As Figuras 6.5 e 6.6 apresentam as estimativas produto e a variância estimada para cada  $\mu_i$  de acordo com os valores de  $\lambda$  positivos e negativos.



(a) Preço venda tomate(Kg) em reais -  $\lambda \leq 0$

(b) Preço venda tomate(Kg) em reais -  $\lambda \geq 0$

Figura 6.5: Estimativas produto. Dados  $\diamond$ , estimativas para  $\lambda = \pm 10$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 1$  (linha com  $\bullet$ ),  $\lambda = \pm 0.7$  (linha pontilhada e tracejada),  $\lambda = \pm 0.5$  (linha com EUROSIMBOLO) e  $\lambda = 0$  (linha cheia).

Percebe-se, na Figura 6.5, que as estimativas de  $\mu_i$  tendem a ser menores para valores de  $\lambda$  positivos e maiores para valores de  $\lambda$  negativos. Para  $\lambda = 0$  (linha cheia) as estimativas tendem a acompanhar melhor o comportamento dos dados. Nota-se também na Figura 6.6, que houve uma maior discrepância das estimativas das variâncias de  $\mu_i$  para  $\lambda = \pm 10$ .

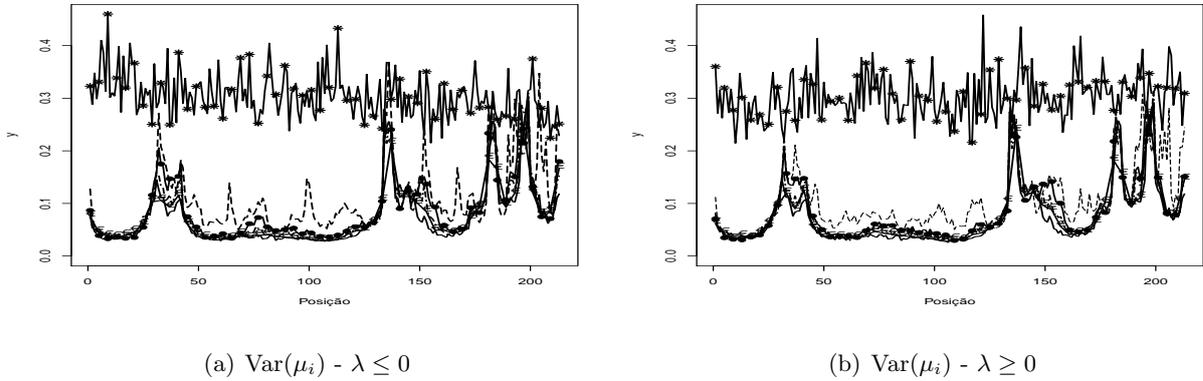


Figura 6.6: Estimativas da variância *a posteriori* para cada  $\mu_i$ . Dados  $\diamond$ , estimativas para  $\lambda = \pm 10$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 1$  (linha com  $\bullet$ ),  $\lambda = \pm 0.7$  (linha pontilhada e tracejada),  $\lambda = \pm 0.5$  (linha com EUROSIMBOLO) e  $\lambda = 0$  (linha cheia).

A Figura 6.7 ilustra a probabilidade de mudança em cada instante e nota-se que essa probabilidade tende a indicar todos os pontos como pontos de mudança para valores de  $\lambda = \pm 100$ . Para  $\lambda = 0$ , cerca de 14 pontos são indicados como pontos de mudança com probabilidade acima de 0.4, mas para  $\lambda = \pm 2$  o modelo tende a indicar cerca de 26 pontos como pontos de mudança com probabilidade acima de 0.4.

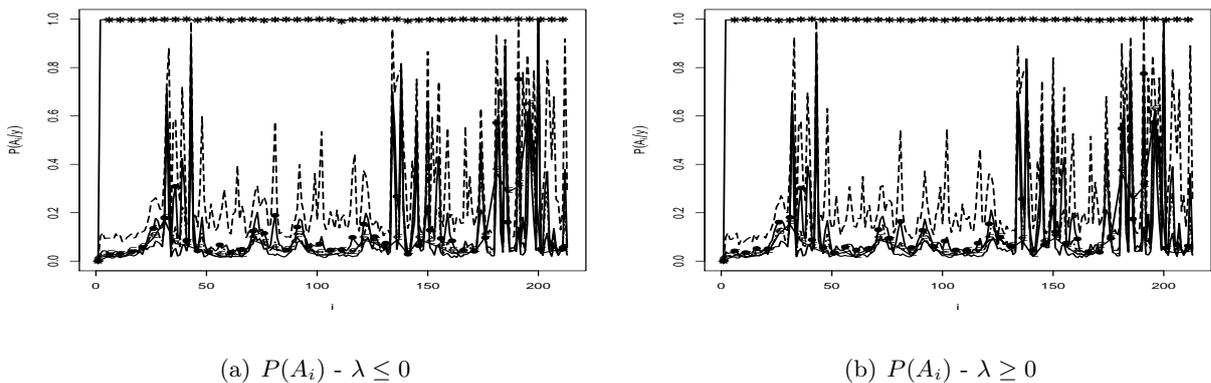
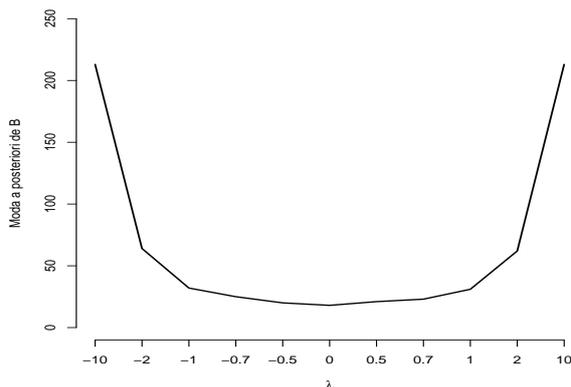


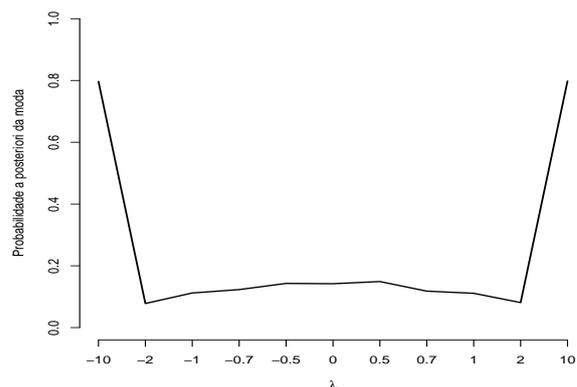
Figura 6.7: Probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 10$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 1$  (linha com  $\bullet$ ),  $\lambda = \pm 0.7$  (linha pontilhada e tracejada),  $\lambda = \pm 0.5$  (linha com EUROSIMBOLO) e  $\lambda = 0$  (linha cheia).

Em relação ao número de blocos, nota-se um número maior de blocos para valores de  $\lambda = (\pm 10$

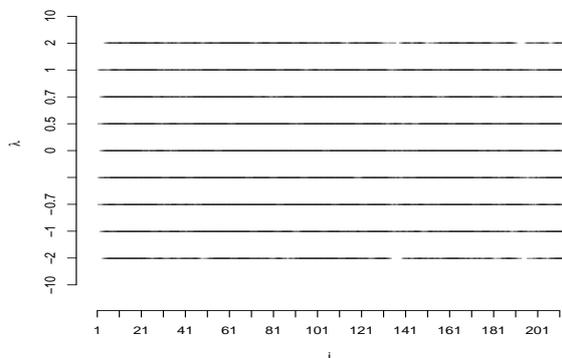
e  $\pm 2$ ), Figura 6.8, o menor número de blocos é indicado com maior probabilidade quando  $\lambda = 0$ . As partições indicadas como mais prováveis para os valores de  $\lambda = \pm 10$  indicam que todos os pontos são pontos de mudança, os demais valores de  $\lambda$  sugerem com baixa probabilidade, 0.001, as partições mais prováveis, como é possível observar na Figura 6.8(d).



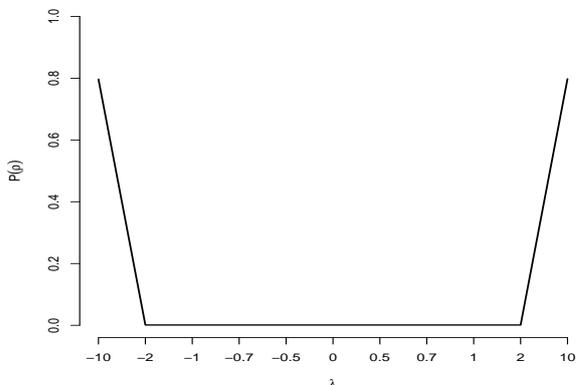
(a) Número de Blocos



(b)  $P(B = b)$



(c) Partição mais provável



(d)  $P(\rho)$

Figura 6.8: Moda a *posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável a *posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ .

A Tabela 6.3 apresenta os valores dos DICs obtidos para cada modelo. Observa-se que o menor DIC ocorre quando  $\lambda = 2$  indicando que o modelo correto deve assumir que os dados são assimétricos à direita.

Tabela 6.3: Valores do DIC para a Série mensal do preço do tomate

$\lambda \leq 0$					
-10	-2	-1	-0.7	-0.5	0
15944.5	-443744.8	-286601.9	-188992.7	-136366	-62815.97
$\lambda \geq 0$					
0.5	0.7	1	2	10	-
-179285.5	-221001.6	-293914.2	<b>-492624</b>	15093.4	-

## 6.2 Série de retorno mensal do Ibovespa

O Ibovespa, índice da bolsa de valores de São Paulo, é o resultado de uma carteira teórica de ativos, elaborada de acordo com os critérios estabelecidos em sua metodologia, cujo objetivo, é ser o indicador do desempenho médio das cotações dos ativos de maior negociabilidade e representatividade do mercado de ações brasileiro (BM&FBovespa, 2014).

Os dados considerados neste estudo de caso são os retornos mensais, no período de abr/94 a jun/14, do Ibovespa. Estes retornos são obtidos da seguinte forma

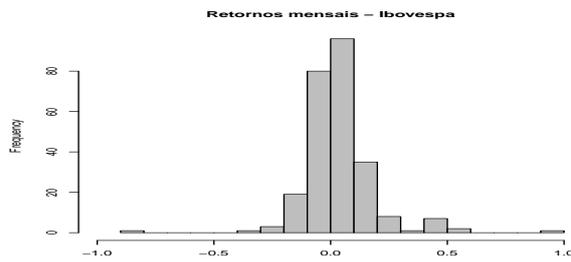
$$R_i = \frac{P_i - P_{i-1}}{P_{i-1}}.$$

A Tabela 6.4 mostra as estatísticas descritivas dos retornos mensais, de onde se observa que a média é maior que a mediana e o primeiro quartil é negativo indicando uma leve assimetria a direita. O coeficiente de variação é de 4.25.

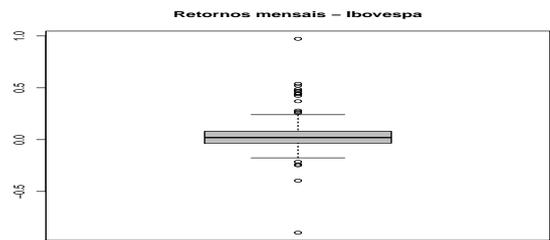
Tabela 6.4: Estatísticas Descritivas - Série dos retornos mensais do Ibovespa

Min	Q1	Mediana	Q3	Max	Média	Desvio padrão
-0.8976	-0.0341	0.0199	0.0796	0.9723	0.0353	0.1500

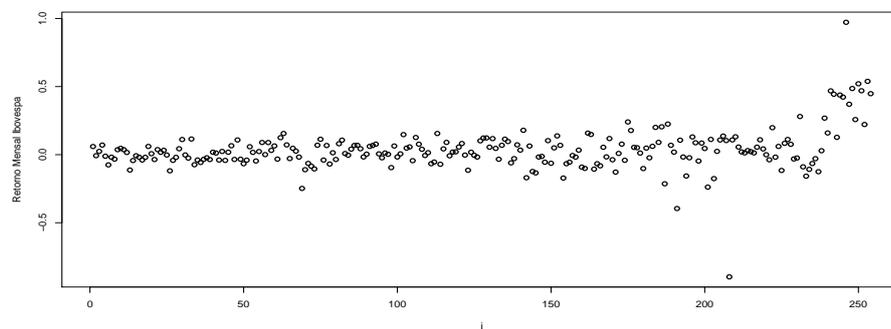
A Figura 6.9 mostra o boxplot e o histograma da série do Ibovespa. A distribuição empírica dos dados revela calda pesada em ambos os lados da distribuição e bimodalidade com a primeira moda sendo relativamente mais importante que a segunda.



(a) Histograma do Retorno Mensal do Ibovespa



(b) Boxplot do Retorno Mensal do Ibovespa



(c) Série do Retorno Mensal do Ibovespa

Figura 6.9: Distribuição do Retorno Mensal do Ibovespa

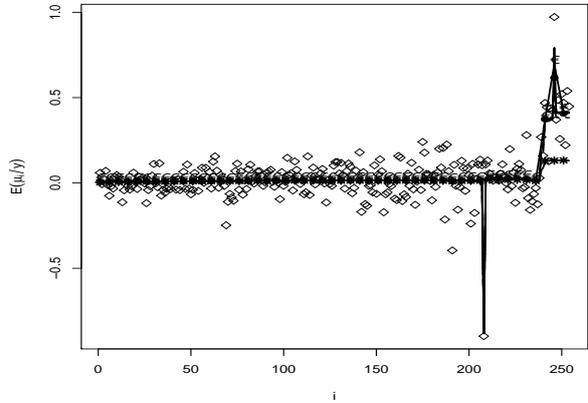
Nas próximas seções são apresentados os resultados para a série de retornos utilizando o Modelo I e o Modelo II.

### 6.2.1 Aplicação - Modelo I - *priori* normal assimétrica

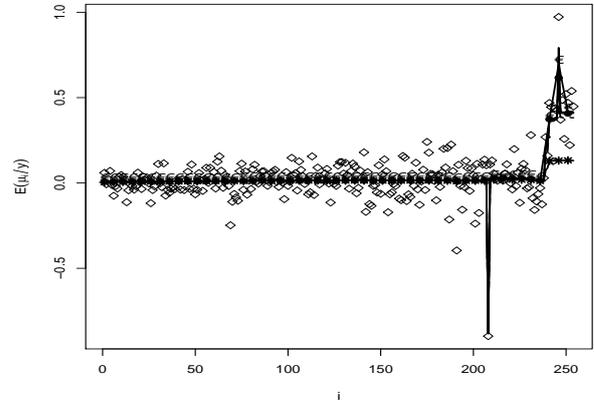
Nesta seção pressupõe-se que o modelo correto deve assumir assimetria *a priori* para as médias. Os valores de  $\lambda$  considerados no ajuste do modelo e as distribuições *a priori* consideradas nesse estudo são os mesmos utilizados na aplicação para a série do tomate no Modelo I discutido na Seção 6.1.1.

Na Figura 6.10 encontram-se as estimativas produto e as variâncias *a posteriori* para cada  $\mu_i$ . Os gráficos posicionados na coluna da direita mostram os resultados obtidos quando ajusta-se um modelo com  $\lambda$  positivo e na coluna da esquerda encontram-se os resultados quando valores negativos

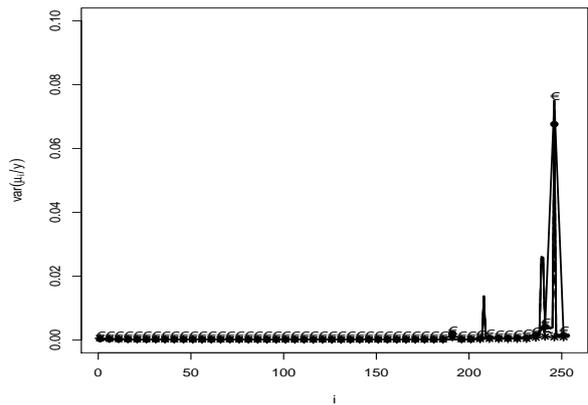
de  $\lambda$  foram considerados.



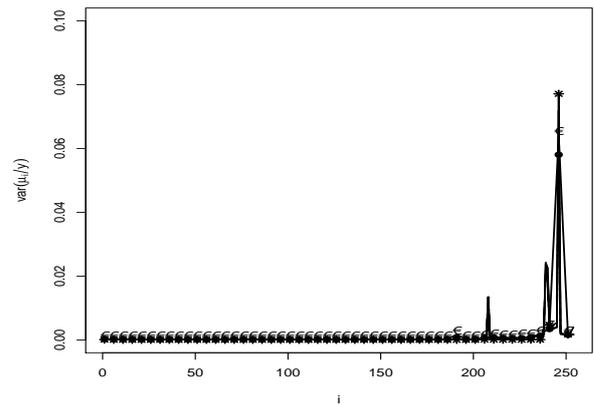
(a) Retorno mensal Ibovespa -  $\lambda \leq 0$



(b) Retorno mensal Ibovespa -  $\lambda \geq 0$



(c)  $Var(\mu_i)$  -  $\lambda \leq 0$



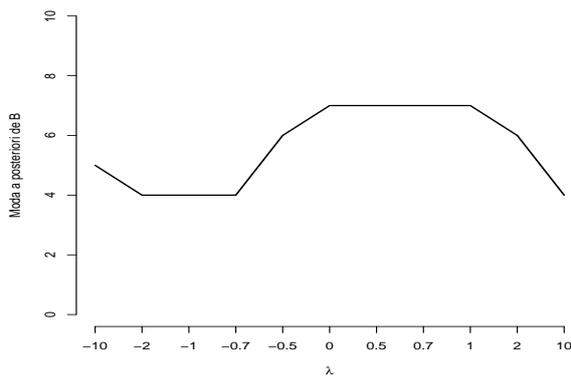
(d)  $Var(\mu_i)$  -  $\lambda \geq 0$

Figura 6.10: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Dados  $\diamond$ , estimativas para  $\lambda = \pm 10$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 1$  (linha com  $\bullet$ ),  $\lambda = \pm 0.7$  (linha pontilhada e tracejada),  $\lambda = \pm 0.5$  (linha com EUROSIMBOLO) e  $\lambda = 0$  (linha cheia).

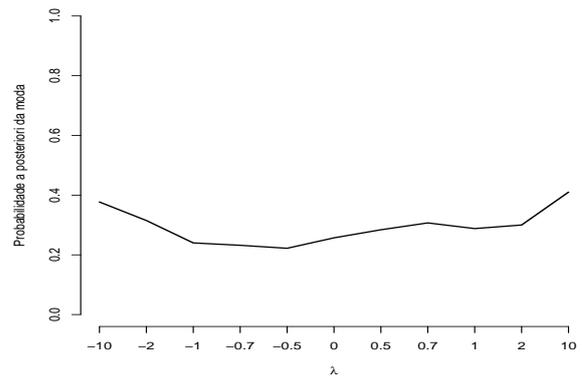
A Figura 6.10 mostra que assumir valores de  $\lambda \neq 0$ , mas próximos de zero não conduzem a estimativas produto para  $\mu_i$  que diferenciam substancialmente. No entanto, as estimativas produto para  $\mu_i$  no final da série tendem a diferir das demais quando  $\lambda = \pm 10$ . Também percebe-se que as incertezas *a posteriori* sobre cada  $\mu_i$  é muita baixa, exceto no final da série, para todos os  $\lambda$ . Nota-se da Figura 6.10(c) e (d) que a variância *a posteriori* foi próxima para todos os valores de

$\lambda$  considerados. Indicando que assumir distribuições *a priori* para  $\mu_i$  com assimetria entre  $-2$  e  $2$  não interfere de forma expressiva nas estimativas.

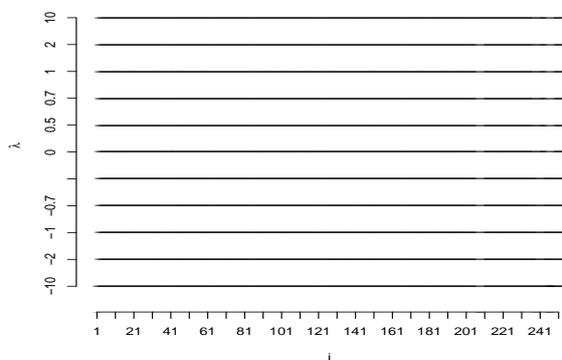
As Figuras 6.11 e 6.12 apresentam as estimativas para o número de blocos, partição mais prováveis e probabilidade de cada instante ser um ponto de mudança.



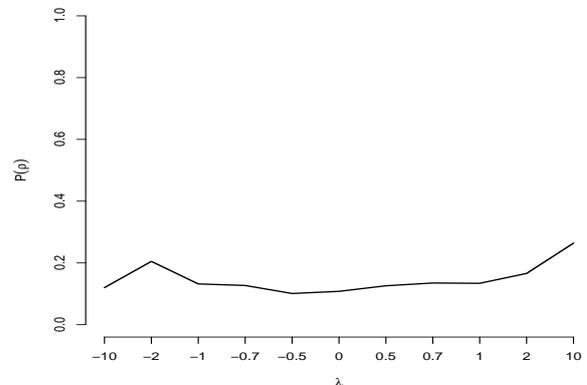
(a) Número de Blocos



(b)  $P(B = b)$



(c) Partição mais provável



(d)  $P(\rho)$

Figura 6.11: Moda a *posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável a *posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ .

Da Figura 6.11 nota-se que para valores de  $\lambda$  entre 0 e 1 a moda *a posteriori* para  $B$  foi 7 e para os outros valores de  $\lambda$  está moda tende a ser menor. No entanto, observa-se na Figura 6.12 que

apenas 4 pontos apresentam probabilidade de mudança acima de 0.5. Indicando que o número de blocos deve ser baixo.

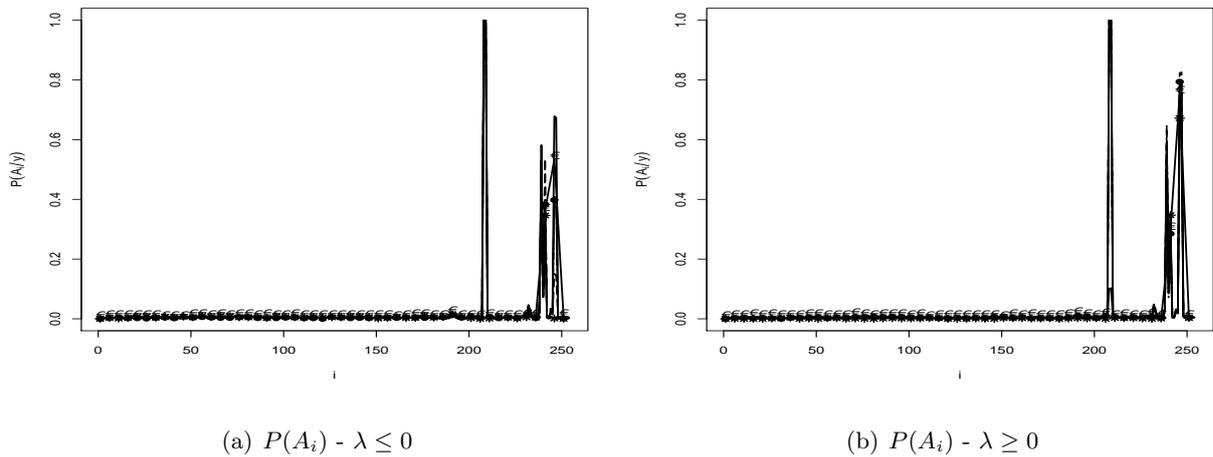


Figura 6.12: Probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 10$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 1$  (linha com ●),  $\lambda = \pm 0.7$  (linha pontilhada e tracejada),  $\lambda = \pm 0.5$  (linha com EUROSIMBOLO) e  $\lambda = 0$  (linha cheia).

A medida de seleção de modelo (DIC) foi calculada e os resultados encontram-se na Tabela 6.5. Avaliando o DIC o melhor modelo ajustado aos dados foi observado para  $\lambda = 2$ .

Tabela 6.5: Valores do DIC para a Série dos retornos mensais do Ibovespa

$\lambda \leq 0$					
-10	-2	-1	-0.7	-0.5	0
-340.5	-432.4	-436.6	-437.2	-439.1	-443.1
$\lambda \geq 0$					
0.5	0.7	1	2	10	-
-445.4	-446.1	<b>-448.1</b>	-447.5	-380.6	-

### 6.2.2 Aplicação - Modelo II - dados normais assimétricos

Como visto na Figura 6.9, os retornos do Ibovespa apresentam certo grau de assimetria. Nesta seção pressupõe-se então, que o modelo correto deve assumir assimetria na distribuição dos dados. Os

valores de  $\lambda$  e as distribuições *a priori* foram as mesmas discutidas na Seção 6.1.1 para a aplicação da série do preço do tomate. Para o MCMC considerou-se que o período de aquecimento da cadeia foi de 60.000, o lag foi de tamanho 10 e o tamanho da amostra de 1000.

As Figuras 6.13 e 6.14 apresentam as estimativas produto e variância *a posteriori* para cada  $\mu_i$ .

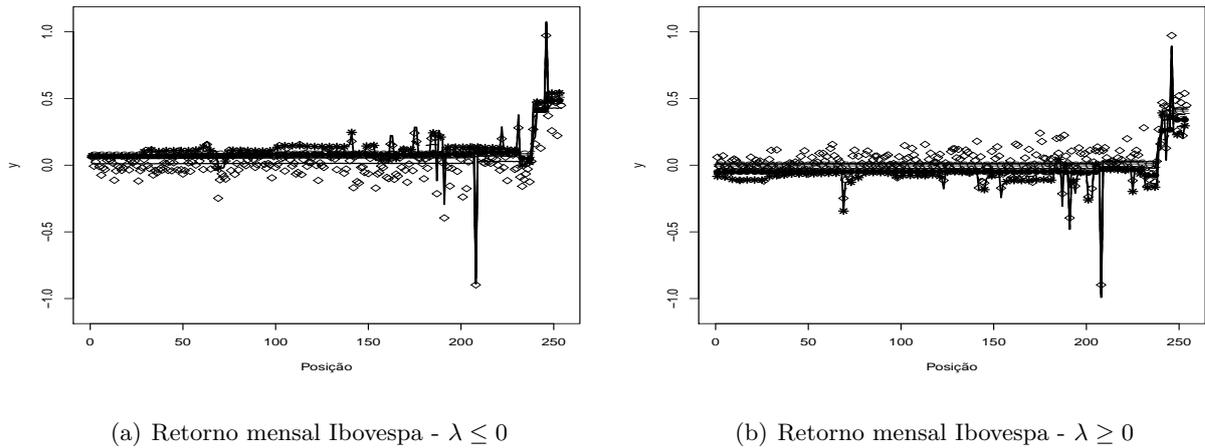


Figura 6.13: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Dados  $\circ$ , estimativas para  $\lambda = \pm 2$  (linha com \*),  $\lambda = \pm 1$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ),  $\lambda = \pm 0.5$  (linha pontilhada e tracejada) e  $\lambda = 0$  (linha cheia).

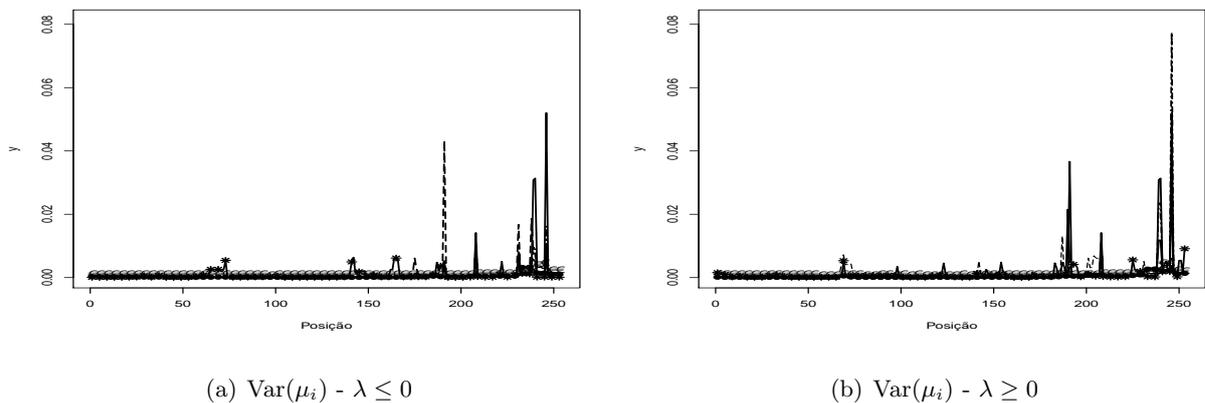
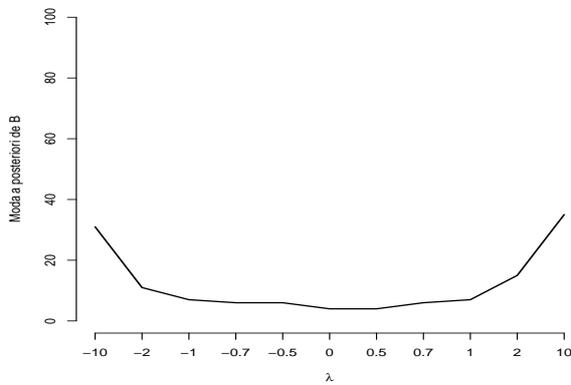


Figura 6.14: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Dados  $\circ$ , estimativas para  $\lambda = \pm 2$  (linha com \*),  $\lambda = \pm 1$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $\bullet$ ),  $\lambda = \pm 0.5$  (linha pontilhada e tracejada) e  $\lambda = 0$  (linha cheia).

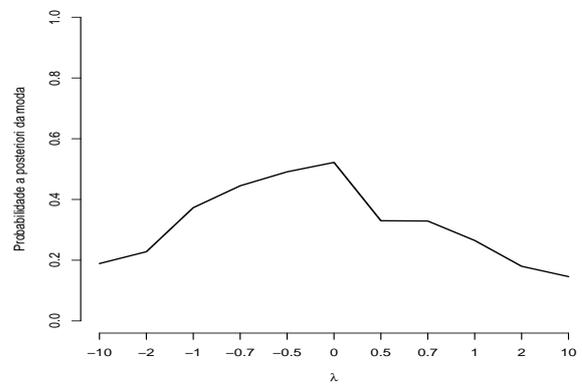
Observa-se nas Figuras 6.14 e 6.13 que as estimativas produto tendem a ser melhores quando val-

ores de  $\lambda$  entre  $-2$  e  $2$  são utilizados. Comparado ao modelo normal em que  $\lambda = 0$  nota-se que para  $\lambda$  negativo as estimativas são maiores que para  $\lambda = 0$ , enquanto que para  $\lambda = 10$  as estimativas produto são menores. Percebe-se também que a incerteza *a posteriori* sobre  $\mu_i$  tende a ser maior em alguns instantes no final da série.

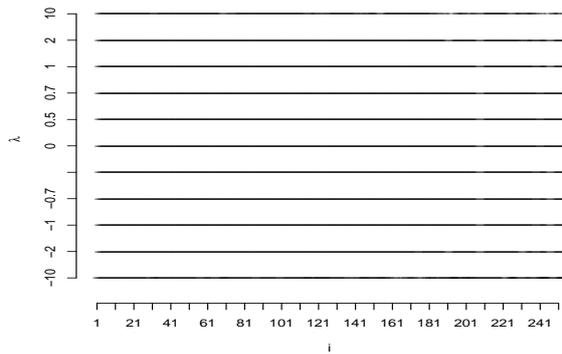
As estimativas para o número de blocos e partição mais prováveis *a posteriori* são mostradas na Figura 6.15.



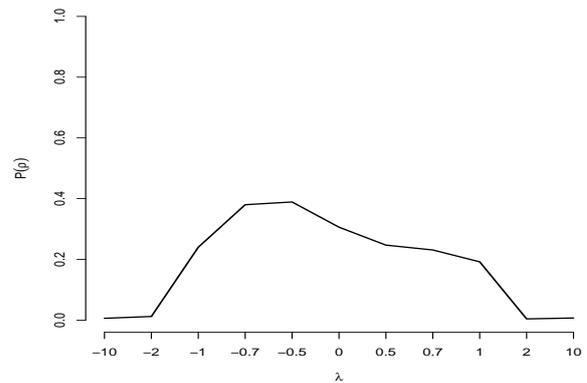
(a) Número de Blocos



(b)  $P(B = b)$



(c) Partição mais provável



(d)  $P(\rho)$

Figura 6.15: Moda *a posteriori* de  $B$ (a) e sua probabilidade(b). Partição mais provável *a posteriori*(c) e sua probabilidade(d), para cada  $\lambda$ .

A moda do número de blocos *a posteriori*, tende a ser maior para valores de  $\lambda = \pm 10$  e esta corresponde a 38 blocos para  $\lambda = -10$  e 41 para  $\lambda = 10$ . A moda *a posteriori* ocorre com maior probabilidade quando  $\lambda = 0$  e corresponde a 4 blocos. A partição mais provável *a posteriori* atinge maior probabilidade quando  $\lambda = -0.5$  e  $-0.7$  e a probabilidade é aproximadamente 0.40.

A Figura 6.16 mostra a probabilidade de cada instante ser um ponto de mudança.

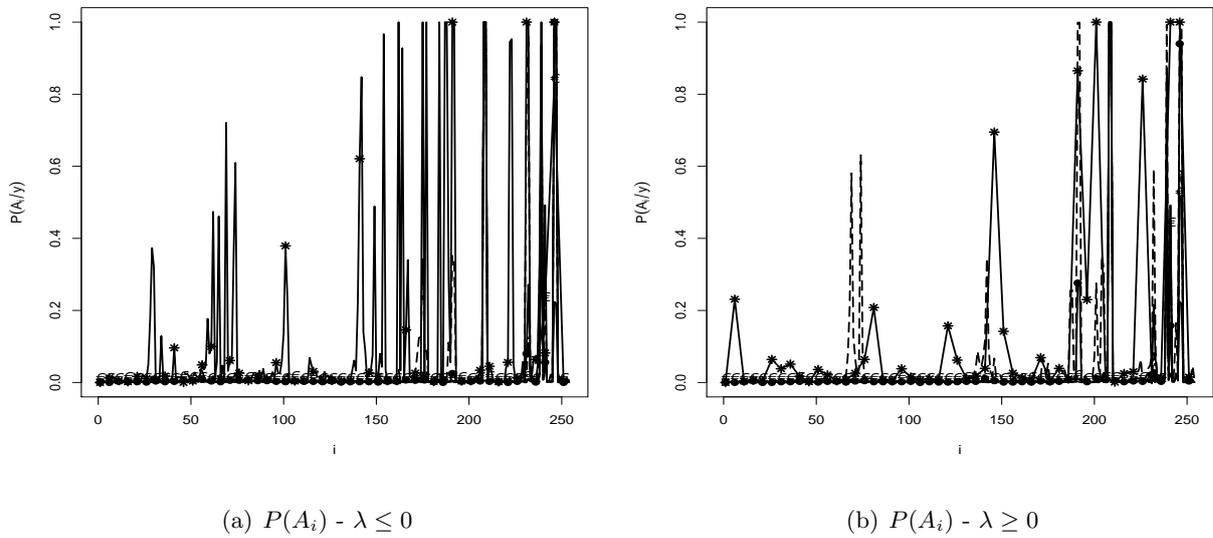


Figura 6.16: Probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 10$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 1$  (linha com ●),  $\lambda = \pm 0.7$  (linha pontilhada e tracejada),  $\lambda = \pm 0.5$  (linha com EUROSIMBOLO) e  $\lambda = 0$  (linha cheia).

A probabilidade de mudança em cada instante tende a ser maior e indicar muitos pontos como mudança para  $\lambda = \pm 10$ . Valores de  $\lambda$  mais próximos de zero resultam em apenas cinco pontos com probabilidade acima de 0.4 de serem um mudança, ver Figura 6.15.

A Tabela 6.6 mostra o DIC calculado para todos os valores de  $\lambda$  e nota-se que o menor valor ocorre para  $\lambda = 0.7$  o que corrobora com a discussão de que os dados possuem uma leve assimetria a direita.

Tabela 6.6: Valore do DIC para a Série dos retornos mensais do Ibovespa

$\lambda \leq 0$					
-10	-2	-1	-0.7	-0.5	0
1829.4	5289.1	3841.9	2888.5	1262.5	1221.0
$\lambda \geq 0$					
0.5	0.7	1	2	10	-
-163.6437	<b>-762.6987</b>	-255.7	4366.1	1599.3	-

---

## Considerações Finais e Trabalhos Futuros

---

O presente trabalho solidificou o estudo do modelo partição produto com correlação entre blocos proposto por Ferreira (2010) por meio de um estudo Monte Carlo e do estudo da correlação imposta pelo modelo proposto. Outra contribuição deste trabalho é propor medidas para investigar a robustez bayesiana no modelo partição produto, no contexto em que a distribuição amostral ou a distribuição *a priori* pertencem a uma classe de distribuições normais assimétricas, além de realizar um estudo de sensibilidade informal e a aplicação dos modelos avaliados a dados reais.

A luz dos resultados obtidos no estudo de robustez global, observa-se que a contaminação na função de verossimilhança afeta de forma mais efetiva as estimativas *a posteriori* que a contaminação na distribuição *a priori* para a grade de  $\lambda$  avaliada. No entanto, na análise de sensibilidade informal assumindo  $\lambda = 100$ , ambos os modelos são afetados de forma significativa, mas na prática  $\lambda = 100$  não é um valor comumente utilizado.

Quando há contaminação da função de verossimilhança nota-se uma perda nas estimativas produto, no entanto há um ganho na estimativa da partição de interesse, com exceção dos cenários epsilon e tendência. Mas a diminuição da variância da distribuição condicional completa de  $\mu_i$  pode induzir

o modelo a identificar com alta probabilidade uma partição próxima da partição de interesse, mas não exatamente a partição correta, como aconteceu no cenário degrau positivo. Todavia, ao assumir  $\lambda = 0$  em nenhum momento o modelo estimou erroneamente a partição de interesse.

As medidas de robustez global de  $\mu_i$ , da  $var(\mu_i)$  e de  $P(A_l)$  apresentaram maiores magnitudes próximas aos pontos de mudança. As menores diferenças observadas ocorreram para o cenário degrau não simétrico, cujo o tamanho do salto entre os blocos é maior que nos demais cenários avaliados.

Em relação as variações das médias amostrais propostas para cada bloco,  $\epsilon = (0, 2, 5, \text{ e } 50)$ , conclui-se, em virtude dos resultados obtidos, que para o estudo de robustez global da distribuição *a priori* não houve diferença para os valores de  $\epsilon$ . Para o caso do estudo de robustez quando a contaminação é na função de verossimilhança, houve uma diferença maior para a quantidade  $\epsilon$ , principalmente, nos cenários tendência e epsilon. Mas os cenários epsilon e tendência são cenários em que normalmente o Modelo Partição Produto apresenta dificuldades na estimação dos parâmetros de interesse.

Em suma, observa-se que considerar assimetria na função de verossimilhança pode trazer ganhos em termos de estimação do número de blocos e partição de interesse para os cenários escada crescente, degrau não simétrico e degrau positivo, mas o custo computacional é alto e as estimativas produto poderão ser levemente viciadas. No entanto, se o objetivo é considerar assimetria na distribuição *a priori* não foram evidenciados ganhos, apenas a indicação que o Modelo I não é robusto em relação ao número de blocos e partição de interesse.

O número de pontos de mudança para a aplicação da série de retorno mensal do Ibovespa assumindo que o parâmetro de assimetria no Modelo II é  $\lambda = -10$  resultou em um maior número de pontos de mudança, como também ocorreu na análise de sensibilidade para  $\lambda = \pm 100$ .

O MPP foi utilizado para identificar “*clusters*” espaciais por Hegarty & Barry (2008). Neste caso, os autores utilizam dados de área e tem como meta estimar o risco relativo nas áreas assumindo a distribuição de Poisson para modelar o comportamento dos dados. Como uma proposta de estudo

futuro, pretende-se estender os resultados de Hegarty & Barry (2008) considerando a classe de distribuições log-normais assimétricas, Marchenko & Genton (2009), para modelar o comportamento do risco relativo. Como esta classe se encaixa na classe de distribuições contaminada de forma multiplicativa, um estudo de robustez para o modelo pode ser realizado.

Este apêndice contém o artigo publicado durante o doutorado.

**A.1 Artigo: Detecting changes in time series: A product partition model with across-cluster correlation**

---

Demonstrações da Proposição 1 - Modelo I

---

Aqui é apresentada a demonstração da Proposição 1 da Seção 2.4. Para este fim, o Lema 2 a seguir, encontrado, por exemplo, em Arellano et al. (2013) será utilizado. Também considerar-se-á a distribuição normal assimétrica multivariada generalizada (Grupta et. al 2004) apresentada na Definição 3, a seguir.

**Lema 2.** *Se  $\mathbf{U} \sim N_n(\mathbf{c}, \mathbf{C})$  é um vetor aleatório não-singular e  $n$ -dimensional, então para qualquer vetor  $\mathbf{a}$  de ordem  $m \times 1$  e qualquer matriz  $\mathbf{A}$  de ordem  $m \times n$ , teremos*

$$E[\Phi_m(\mathbf{A}\mathbf{U} + \mathbf{a}|\mathbf{b}, \mathbf{B})] = \Phi_m(\mathbf{A}\mathbf{c} + \mathbf{a}|\mathbf{b}, \mathbf{B} + \mathbf{A}\mathbf{C}\mathbf{A}^T).$$

A definição a seguir, apresenta a família de distribuições normal assimétrica multivariada generalizada definida por Grupta et. al (2004). Esta família de distribuições será útil na prova das Proposições 1 e 2.

**Definição 3.** Um vetor aleatório  $\mathbf{Y} \sim \mathcal{GMSN}_{p,q}(\xi, \Sigma, \mathbf{D}, \nu, \Delta)$ , se a função densidade de probabilidade de  $\mathbf{Y}$  é dada por

$$f(\mathbf{y}|\xi, \Sigma, \mathbf{D}, \nu, \Delta) = \phi_p(\mathbf{y}|\xi, \Sigma) \frac{\Phi_q(\mathbf{D}\mathbf{y}|\nu, \Delta)}{\Phi_q(\mathbf{D}\xi|\nu, \Delta + \mathbf{D}\Sigma\mathbf{D}')}$$

e a função geradora de momentos de  $\mathbf{Y}$  é dada por

$$M_Y(t) = \frac{\Phi_q[\mathbf{D}(\xi + \mathbf{\Sigma}t) | \nu, \mathbf{\Delta} + \mathbf{D}\mathbf{\Sigma}\mathbf{D}']}{\Phi_q(\mathbf{D}\xi | \nu, \mathbf{\Delta} + \mathbf{D}\mathbf{\Sigma}\mathbf{D}')} \exp^{\xi't + \frac{1/2}{t'\mathbf{\Sigma}t}},$$

onde  $\xi \in \mathbb{R}^p$ ,  $\nu \in \mathbb{R}^q$ ,  $\mathbf{\Sigma}(p \times q)$  e  $\mathbf{\Delta}(q \times q)$  são duas matrizes de covariância e  $\mathbf{D}(q \times p)$  é uma matriz arbitrária.

Logo, o primeiro momento é dado por

$$\xi + \frac{\mathbf{D}\mathbf{\Sigma}\phi(\mathbf{D}\xi | \nu, \mathbf{\Delta} + \mathbf{D}\mathbf{\Sigma}\mathbf{D}')}{\Phi(\mathbf{D}\xi | \nu, \mathbf{\Delta} + \mathbf{D}\mathbf{\Sigma}\mathbf{D}')}.$$

## B.1 Demonstração da Proposição 1

i) Assuma a função de verossimilhança em (2.19) para o bloco  $[i_j i_{j+1}]$  e as distribuições *a priori* para  $\mu_{i_j i_{j+1}}$ , com  $j = 0, \dots, b-1$ ,  $p$  e  $\rho$  eliciadas em (2.2), (2.20) e (2.21). Como consequência, segue que

$$f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]}, \mu_{[i_j i_{j+1}]}) = \prod_{k=i_j}^{i_{j+1}} \frac{1}{\sigma_y} \phi\left(\frac{y_k - \mu_{[i_j i_{j+1}]}}{\sigma_y}\right) \frac{2}{\sigma_\mu} \phi\left(\frac{\mu_{[i_j i_{j+1}]} - \mu_0}{\sigma_\mu}\right) \Phi\left(\lambda \left(\frac{\mu_{[i_j i_{j+1}]} - \mu_0}{\sigma_\mu}\right)\right) \quad (\text{B.1})$$

Para facilitar a notação, assuma que  $n_j = i_{j+1} - i_j + 1$ . Desta forma, a distribuição preditiva para o bloco  $Y_{[i_j i_{j+1}]}$  é dada por

$$\begin{aligned} f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]}) &= \int_{\mu_{[i_j i_{j+1}]}} \prod_{k=i_j}^{i_{j+1}} \frac{1}{\sigma_y} \phi\left(\frac{y_k - \mu_{[i_j i_{j+1}]}}{\sigma_y}\right) \frac{2}{\sigma_\mu} \phi\left(\frac{\mu_{[i_j i_{j+1}]} - \mu_0}{\sigma_\mu}\right) \\ &\quad \times \Phi\left(\lambda \left(\mu_{[i_j i_{j+1}]} \mid \lambda\mu_0, \sigma_\mu\right)\right) d\mu_{[i_j i_{j+1}]} \\ &= \int_{\mu_{[i_j i_{j+1}]}} \left(\frac{1}{\sqrt{2\pi\sigma_y^2}}\right)^{n_j} \exp\left\{-\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \mu_{[i_j i_{j+1}]})^2\right\} \\ &\quad \times \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \exp\left\{-\frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2\right\} \Phi\left(\lambda \left(\mu_{[i_j i_{j+1}]} \mid \lambda\mu_0, \sigma_\mu\right)\right) d\mu_{[i_j i_{j+1}]} \end{aligned} \quad (\text{B.2})$$

$$\begin{aligned}
f_{[i_j i_{j+1}]}(Y_{[i_j i_{j+1}]}) &= \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \int_{\mu_{[i_j i_{j+1}]}} \exp \left\{ -\frac{1}{2\sigma_y^2} \left[ \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 + n_j (\mu_{[i_j i_{j+1}]} - \bar{y})^2 \right] \right\} \\
&\times \exp \left\{ -\frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\} \Phi \left( \lambda \mu_{[i_j i_{j+1}]} \mid \lambda \mu_0, \sigma_\mu \right) d\mu_{[i_j i_{j+1}]} \\
&= \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \\
&\times \int_{\mu_{[i_j i_{j+1}]}} \exp \left\{ -\frac{n_j}{2\sigma_y^2} (\mu_{[i_j i_{j+1}]} - \bar{y})^2 \right\} \exp \left\{ -\frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\} \\
&\times \Phi \left( \lambda \mu_{[i_j i_{j+1}]} \mid \lambda \mu_0, \sigma_\mu \right) d\mu_{[i_j i_{j+1}]} \\
&= \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \\
&\times \int_{\mu_{[i_j i_{j+1}]}} \exp \left\{ -\frac{1}{2} \left[ \left( \frac{n_j}{\sigma_y^2} + \frac{1}{\sigma_\mu^2} \right) \left( \mu_{[i_j i_{j+1}]} - \frac{\frac{n_j \bar{y}}{\sigma_y^2} + \frac{\mu_0}{\sigma_\mu^2}}{\frac{n_j}{\sigma_y^2} + \frac{1}{\sigma_\mu^2}} \right) \right] \right\} \\
&\times \exp \left\{ -\frac{1}{2} \left( \frac{\frac{n_j}{\sigma_y^2} \frac{1}{\sigma_\mu^2}}{\frac{n_j}{\sigma_y^2} + \frac{1}{\sigma_\mu^2}} \right) (\bar{y} - \mu_0)^2 \right\} \Phi \left( \lambda \mu_{[i_j i_{j+1}]} \mid \lambda \mu_0, \sigma_\mu \right) d\mu_{[i_j i_{j+1}]} \\
&= \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \times \exp \left\{ -\frac{1}{\frac{2(n_j\sigma_\mu^2 + \sigma_y^2)}{n_j}} (\bar{y} - \mu_0)^2 \right\} \\
&\times \int_{\mu_{[i_j i_{j+1}]}} \exp \left\{ \frac{1}{\frac{2(\sigma_\mu^2 \sigma_y^2)}{n_j \sigma_\mu^2 + \sigma_y^2}} \left( \mu_{[i_j i_{j+1}]} - \frac{\sigma_\mu^2 n_j \bar{y} + \mu_0 \sigma_y^2}{n_j \sigma_\mu^2 + \sigma_y^2} \right)^2 \right\} \\
&\times \Phi \left( \lambda \mu_{[i_j i_{j+1}]} \mid \lambda \mu_0, \sigma_\mu \right) d\mu_{[i_j i_{j+1}]} \\
&= \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \\
&\exp \left\{ -\frac{1}{\frac{2(n_j\sigma_\mu^2 + \sigma_y^2)}{n_j}} (\bar{y} - \mu_0)^2 \right\} \frac{1}{\sqrt{\frac{2\pi(\sigma_\mu^2 \sigma_y^2)}{n_j \sigma_\mu^2 + \sigma_y^2}}} \\
&\times \int_{\mu_{[i_j i_{j+1}]}} \frac{1}{\sqrt{\frac{2\pi(\sigma_\mu^2 \sigma_y^2)}{n_j \sigma_\mu^2 + \sigma_y^2}}} \exp \left\{ \frac{1}{\frac{2(\sigma_\mu^2 \sigma_y^2)}{n_j \sigma_\mu^2 + \sigma_y^2}} \left( \mu_{[i_j i_{j+1}]} - \frac{\sigma_\mu^2 n_j \bar{y} + \mu_0 \sigma_y^2}{n_j \sigma_\mu^2 + \sigma_y^2} \right)^2 \right\} \\
&\times \Phi \left( \lambda \mu_{[i_j i_{j+1}]} \mid \lambda \mu_0, \sigma_\mu \right) d\mu_{[i_j i_{j+1}]}
\end{aligned}$$

(B.3)

Note que a expressão dentro da integral em (B.3) é a esperança de  $\Phi(\lambda\mu_{[i_j i_{j+1}]} | \lambda\mu_0, \sigma_\mu^2)$  com respeito à distribuição de uma variável aleatória que tem distribuição normal com média  $m^* = \frac{\sigma_\mu^2 n_j \bar{y} + \mu_0 \sigma_y^2}{\sigma_\mu n_j + \sigma_y^2}$  e variância  $V^* = \frac{\sigma_y^2 \sigma_\mu^2}{n_j \sigma_\mu^2 + \sigma_y^2}$ . Logo, utilizando o Lema 2 tem-se que

$$E \left( \Phi \left( \lambda\mu_{[i_j i_{j+1}]} | \lambda\mu_0, \sigma_\mu^2 \right) \right) = \Phi \left( \lambda m^* | \lambda\mu_0, \sigma_\mu^2 + \lambda^2 V^* \right). \quad (\text{B.4})$$

Assim, substituindo (B.4) em (B.3) segue que a distribuição preditiva por bloco possui a seguinte expressão

$$\begin{aligned} f_{[i_j i_{j+1}]}(\mathbf{Y}_{[i_j i_{j+1}]}) &= \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \\ &\times \frac{1}{\sqrt{\frac{2\pi(\sigma_y^2 \sigma_\mu^2)}{n_j \sigma_\mu^2 + \sigma_y^2}}} \exp \left\{ -\frac{1}{\frac{2(n_j \sigma_\mu^2 + \sigma_y^2)}{n_j}} (\bar{y} - \mu_0)^2 \right\} \Phi \left( \lambda m^* | \lambda\mu_0, \sigma_\mu^2 + \lambda^2 V^* \right) \\ &= \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \\ &\times \sqrt{\sigma_y^2 \sigma_\mu^2} \frac{1}{\sqrt{n_j}} \frac{1}{\sqrt{\frac{2\pi(n_j \sigma_\mu^2 + \sigma_y^2)}{n_j}}} \exp \left\{ -\frac{1}{\frac{2(n_j \sigma_\mu^2 + \sigma_y^2)}{n_j}} (\bar{y} - \mu_0)^2 \right\} \Phi \left( \lambda m^* | \lambda\mu_0, \sigma_\mu^2 + \lambda^2 V^* \right) \\ &= 2\sqrt{\frac{\sigma_y^2}{2\pi n_j}} \phi_{n_j} \left( y_k | \bar{y}, \sigma_y^2 \right) \phi \left( \bar{y} | \mu_0, \frac{n_j \sigma_\mu^2 + \sigma_y^2}{n_j} \right) \Phi \left( \lambda m^* | \lambda\mu_0, \sigma_\mu^2 + \lambda^2 V^* \right). \end{aligned} \quad (\text{B.5})$$

ii) A prova do item (ii) é verificada usando o teorema de Bayes como segue.

$$\begin{aligned} f(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) &\propto \prod_{k=i_j}^{i_{j+1}} \frac{1}{\sigma_y} \phi \left( \frac{y_k - \mu_{[i_j i_{j+1}]}}{\sigma_y} \right) \frac{2}{\sigma_\mu} \phi \left( \frac{\mu_{[i_j i_{j+1}]} - \mu_0}{\sigma_\mu} \right) \Phi \left( \lambda \left( \mu_{[i_j i_{j+1}]} | \lambda\mu_0, \sigma_\mu \right) \right) \\ &\propto \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \mu_{[i_j i_{j+1}]})^2 \right\} \\ &\times \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \exp \left\{ -\frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\} \times \Phi \left( \lambda \left( \mu_{[i_j i_{j+1}]} | \lambda\mu_0, \sigma_\mu \right) \right) \end{aligned} \quad (\text{B.6})$$

$$\begin{aligned}
f(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) &\propto \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \left[ \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 + n_j (\mu_{[i_j i_{j+1}]} - \bar{y})^2 \right] \right\} \\
&\times \exp \left\{ -\frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\} \Phi \left( \lambda \mu_{[i_j i_{j+1}]} | \lambda \mu_0, \sigma_\mu \right) \\
&\propto \frac{2}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \\
&\times \exp \left\{ -\frac{n_j}{2\sigma_y^2} (\mu_{[i_j i_{j+1}]} - \bar{y})^2 \right\} \exp \left\{ -\frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\} \Phi \left( \lambda \mu_{[i_j i_{j+1}]} | \lambda \mu_0, \sigma_\mu \right) \\
&\propto \sqrt{\frac{\sigma_y^2 \sigma_\mu^2}{2\pi(n_j \sigma_\mu^2 + \sigma_y^2)}} \exp \left\{ -\frac{\sigma_\mu^2 n_j + \sigma_y^2}{2\sigma_y^2 \sigma_\mu^2} \left[ \left( \mu_{[i_j i_{j+1}]} - \frac{\sigma_\mu^2 n_j \bar{y} + \sigma_y^2 \mu_0}{n_j \sigma_\mu^2 + \sigma_y^2} \right) \right]^2 \right\} \\
&\times \Phi \left( \lambda \mu_{[i_j i_{j+1}]} | \lambda \mu_0, \sigma_\mu \right) \\
&\propto \phi \left( \mu_{[i_j i_{j+1}]} | m^*, v^* \right) \Phi \left( \lambda \mu_{[i_j i_{j+1}]} | \lambda \mu_0, \sigma_\mu \right)
\end{aligned} \tag{B.7}$$

Logo, dividindo pela distribuição preditiva dada em (i) tem-se que a distribuição *a posteriori* por bloco para  $\mu_{[i_j i_{j+1}]}$ , dada as observações do bloco  $[i_j i_{j+1}]$ , é

$$f(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) = \phi \left( \mu_{[i_j i_{j+1}]} | m^*, v^* \right) \frac{\Phi \left( \lambda \mu_{[i_j i_{j+1}]} | \lambda \mu_0, \sigma_\mu \right)}{\Phi \left( \lambda m^* | \lambda \mu_0, \sigma_\mu^2 + \lambda^2 V^* \right)}. \tag{B.8}$$

Pode-se notar da Definição 3 que a distribuição em (B.8) pertence a classe de distribuições definida por Gupta et al. (2004), que é denotada por

$$\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]} \sim \mathcal{GMSN}_{1,1}(\xi_{[i_j i_{j+1}]}, \Sigma, D, v, \Delta),$$

onde  $\xi_{[i_j i_{j+1}]} = m^*$ ,  $\Sigma = V^*$ ,  $D = \lambda$ ,  $v = \lambda \mu_0$  e  $\Delta = \sigma_\mu$ ,  $m^*$  e  $V^*$  são como definidas em (i).

iii) Segue então do resultado dado em Gupta et al. (2004) que o valor esperado de  $\mu_{[i_j i_{j+1}]}$ , dado  $\mathbf{y}_{[i_j i_{j+1}]}$ , é

$$E(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) = m^* + \frac{\lambda V^* \phi(\lambda m^* | \lambda \mu_0, \sigma_\mu + \lambda^2 V^*)}{\Phi(\lambda m^* | \lambda \mu_0, \sigma_\mu + \lambda^2 V^*)}.$$

■

---

Demonstração da Proposição 2 - Modelo II

---

### C.1 Demonstração da Proposição 2

*i)* Assuma que  $Y_i \mid \mu_i, \sigma_y^2, \lambda \sim \mathcal{SN}(\mu_i, \sigma_y^2, \lambda)$  e considere as distribuições *a priori* para  $\mu_{[i_j i_{j+1}]}$ ,  $\rho$  e  $p$  assumidas na Proposição 2. Segue que para as observações no bloco  $[i_j i_{j+1}]$ , isto é, para  $y_{i_j+1}, \dots, y_{i_{j+1}}$  tem-se que

$$f(\mathbf{y}_{[i_j i_{j+1}]}, \mu_{[i_j i_{j+1}]}) = \prod_{k=i_j}^{i_{j+1}} \frac{2}{\sigma_y} \phi\left(\frac{y_k - \mu_{[i_j i_{j+1}]}}{\sigma_y}\right) \Phi\left(\lambda \left(\frac{y_k - \mu_{[i_j i_{j+1}]}}{\sigma_y}\right)\right) \frac{1}{\sigma_\mu} \phi\left(\frac{\mu_{[i_j i_{j+1}]} - \mu_0}{\sigma_\mu}\right) \quad (\text{C.1})$$

Como antes, assuma que  $n_j = i_{j+1} - i_j + 1$ . Desta forma, a distribuição preditiva *a priori* para o bloco  $\mathbf{y}_{i_j i_{j+1}}$  é dada por

$$f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) = \int_{\mu_{[i_j i_{j+1}]}} \prod_{k=i_j}^{i_{j+1}} \frac{2}{\sigma_y} \phi\left(\frac{y_k - \mu_{[i_j i_{j+1}]}}{\sigma_y}\right) \Phi\left(\lambda \left(\frac{y_k - \mu_{[i_j i_{j+1}]}}{\sigma_y}\right)\right) \frac{1}{\sigma_\mu} \times \phi\left(\frac{\mu_{[i_j i_{j+1}]} - \mu_0}{\sigma_\mu}\right) d\mu_{[i_j i_{j+1}]} \quad (\text{C.2})$$

$$\begin{aligned}
f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) &= \int_{\mu_{[i_j i_{j+1}]}} \left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \mu_{[i_j i_{j+1}]})^2 \right\} \\
&\times \Phi_{n_j} \left( \lambda \mathbf{y}_{[i_j i_{j+1}]} | \lambda \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right) \\
&\times \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \exp \left\{ -\frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\} d\mu_{[i_j i_{j+1}]} \\
&= \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \int_{\mu_{[i_j i_{j+1}]}} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \mu_{[i_j i_{j+1}]})^2 - \frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\} \\
&\times \Phi_{n_j} \left( \lambda \mathbf{y}_{[i_j i_{j+1}]} | \lambda \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right) d\mu_{[i_j i_{j+1}]} \\
&= \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \\
&\times \exp \left\{ -\frac{1}{2\sigma_y^2} \left[ \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 + \sum_{k=i_j}^{i_{j+1}} (\mu_{[i_j i_{j+1}]} - \bar{y})^2 \right] - \frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\} \\
&= \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \\
&\times \int_{\mu_{[i_j i_{j+1}]}} \exp \left\{ -\frac{n_j}{2\sigma_y^2} (\mu_{[i_j i_{j+1}]} - \bar{y})^2 - \frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - \mu_0)^2 \right\} \\
&\times \Phi_{n_j} \left( \lambda \mathbf{y}_{[i_j i_{j+1}]} | \lambda \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right) d\mu_{[i_j i_{j+1}]} \\
&= \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \exp \left\{ -\frac{n_j}{2(\sigma_y^2 + \sigma_\mu^2 n_j)} (\bar{y} - \mu_0)^2 \right\} \\
&\times \sqrt{\frac{2\pi\sigma_y^2}{\sigma_\mu^2 n_j + \sigma_y^2}} \int_{\mu_{[i_j i_{j+1}]}} \frac{1}{\sqrt{\frac{2\pi\sigma_y^2\sigma_\mu^2}{\sigma_\mu^2 n_j + \sigma_y^2}}} \exp \left\{ -\frac{1}{2} \frac{\sigma_\mu^2 n_j + \sigma_y^2}{\sigma_y^2 \sigma_\mu^2} \left( \mu_{[i_j i_{j+1}]} - \frac{\sigma_y^2 \mu_0 + \sigma_\mu^2 n_j \bar{y}}{\sigma_y^2 + \sigma_\mu^2 n_j} \right)^2 \right\} \\
&\times \Phi_{n_j} \left( \lambda \mathbf{y}_{[i_j i_{j+1}]} | \lambda \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right) d\mu_{[i_j i_{j+1}]} \cdot
\end{aligned} \tag{C.3}$$

Note que a expressão dentro da integral na expressão (C.3) é a esperança de  $\Phi_{n_j} \left( -\lambda \mu_{[i_j i_{j+1}]} | -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right)$  com respeito a função acumulada da distribuição normal com média  $m^* = \frac{\sigma_\mu^2 n_j \bar{y} + \mu_0 \sigma_y^2}{\sigma_\mu^2 n_j + \sigma_y^2}$  e variância  $V^* = \frac{\sigma_y^2 \sigma_\mu^2}{n_j \sigma_\mu^2 + \sigma_y^2}$ . Logo, utilizando o Lema 2 tem-se que:

$$E \left( \Phi_{n_j} \left( -\lambda \mu_{[i_j i_{j+1}]} \mid -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right) \right) = \Phi_{n_j} \left( -\lambda m^* \mathbf{1}_{n_j} \mid -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j} \right). \quad (\text{C.4})$$

Logo, segue de (C.3) e (C.4) que a distribuição preditiva a priori por bloco é dada por

$$\begin{aligned} f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]}) &= \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \sqrt{\frac{2\pi\sigma_y^2\sigma_\mu^2}{\sigma_\mu^2 n_j + \sigma_y^2}} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \\ &\times \exp \left\{ -\frac{n_j}{2(\sigma_y^2 + \sigma_\mu^2 n_j)} (\bar{y} - \mu_0)^2 \right\} \Phi_{n_j} \left( -\lambda m^* \mathbf{1}_{n_j} \mid -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j} \right) \\ &= 2^{n_j} \sqrt{\frac{2\pi\sigma_y^2}{n_j}} \left( \frac{1}{\sqrt{\frac{\sigma_\mu^2 n_j + \sigma_y^2}{n_j}}} \right) \phi \left( \frac{\bar{y} - \mu_0}{\sqrt{\frac{\sigma_\mu^2 n_j + \sigma_y^2}{n_j}}} \right) \\ &\times \left( \frac{1}{\sigma_y} \right)^{n_j} \phi_{n_j} \left( \frac{y_k - \bar{y}}{\sigma_y} \right) \Phi_{n_j} \left( -\lambda m^* \mathbf{1}_{n_j} \mid -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j} \right) \\ &= 2^{n_j} \sqrt{\frac{2\pi\sigma_y^2}{n_j}} \phi_{n_j} (y_k \mid \bar{y}, \sigma_y^2) \\ &\times \phi \left( \bar{y} \mid \mu_0, \frac{\sigma_\mu^2 n_j + \sigma_y^2}{n_j} \right) \Phi_{n_j} \left( -\lambda m^* \mathbf{1}_{n_j} \mid -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^* \mathbf{1}_{n_j} \mathbf{1}'_{n_j} \right) \end{aligned} \quad (\text{C.5})$$

ii) Para provar o item (ii) note que a distribuição *a posteriori* para a média comum do bloco  $[i_j, i_{j+1}]$  é dada por

$$f(\mu_{[i_j i_{j+1}]} \mid \mathbf{y}_{[i_j i_{j+1}]}) = \frac{f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]} \mid \mu_{[i_j i_{j+1}]}) \pi(\mu_{[i_j i_{j+1}]})}{f_{[i_j i_{j+1}]}(\mathbf{y}_{[i_j i_{j+1}]})} \quad (\text{C.6})$$

Resolvendo o numerador da razão em C.6, tem-se que

$$\begin{aligned} f(\mathbf{y}_{[i_j i_{j+1}]} \mid \mu_{[i_j i_{j+1}]}) \pi(\mu_{[i_j i_{j+1}]}) &= \prod_{k=i_j}^{i_{j+1}} f(y_k \mid \mu_{[i_j i_{j+1}]}) \pi(\mu_{[i_j i_{j+1}]}) \\ &= \phi_{n_j} \left( \mathbf{y}_{[i_j i_{j+1}]} \mid \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right) \Phi_{n_j} \left( \lambda \mathbf{y}_{[i_j i_{j+1}]} \mid \lambda \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right) \\ &\times \phi \left( \mu_{[i_j i_{j+1}]} \mid \mu_0, \sigma_\mu^2 \right) \end{aligned} \quad (\text{C.7})$$

$$\begin{aligned}
f(\mathbf{y}_{[i_j i_{j+1}]} | \mu_{[i_j i_{j+1}]}) \pi(\mu_{[i_j i_{j+1}]}) &= \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \\
&\times \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k^2 - \mu_{[i_j i_{j+1}]})^2 - \frac{1}{\sigma_\mu^2} (\mu_{[i_j i_{j+1}]}^2 - \mu_0)^2 \right\}
\end{aligned} \tag{C.8}$$

Da demonstração da parte (i) da Proposição 2 a expressão pode ser reescrita como

$$\begin{aligned}
f(\mathbf{y}_{[i_j i_{j+1}]} | \mu_{[i_j i_{j+1}]}) \pi(\mu_{[i_j i_{j+1}]}) &= \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\} \\
&\times \exp \left\{ -\frac{n_j}{2(n_j\sigma_\mu^2 + \sigma_y^2)} \right\} \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \\
&\times \exp \left\{ -\frac{1}{2V^\star} (\mu_{[i_j i_{j+1}]} - m^\star)^2 \right\} \Phi_{n_j} \left( \lambda \mathbf{y}_{[i_j i_{j+1}]} | \lambda \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right)
\end{aligned} \tag{C.9}$$

Daí, substituindo C.9 e C.5 em C.6 tem-se que

$$\begin{aligned}
f(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) &= \frac{\left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\}}{\left( \frac{2}{\sqrt{2\pi\sigma_y^2}} \right)^{n_j} \exp \left\{ -\frac{1}{2\sigma_y^2} \sum_{k=i_j}^{i_{j+1}} (y_k - \bar{y})^2 \right\}} \\
&\times \frac{\frac{1}{\sqrt{2\pi\sigma_\mu^2}} \exp \left\{ -\frac{n_j}{2(n_j\sigma_\mu^2 + \sigma_y^2)} (\bar{y} - \mu_0)^2 \right\}}{\left( \frac{1}{\sqrt{\frac{2\pi(n_j\sigma_\mu^2 + \sigma_y^2)}{n_j}}} \right) \exp \left\{ -\frac{n_j}{2(n_j\sigma_\mu^2 + \sigma_y^2)} (\bar{y} - \mu_0)^2 \right\}} \\
&\times \frac{\exp \left\{ -\frac{1}{2} \frac{n_j\sigma_\mu^2 + \sigma_y^2}{\sigma_y^2\sigma_\mu^2} (\mu_{[i_j i_{j+1}]} - m^\star)^2 \right\}}{\frac{\sqrt{2\pi\sigma_y^2}}{\sqrt{n_j}} \frac{1}{\sqrt{\frac{2\pi\sigma_\mu^2\sigma_y^2}{n_j\sigma_\mu^2 + \sigma_y^2}}}} \\
&\times \frac{\Phi_{n_j} \left( \lambda y_k | \lambda \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right)}{\Phi_{n_j} \left( -\lambda m^\star \mathbf{1}_{n_j} | -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^\star \mathbf{1}_{n_j} \mathbf{1}'_{n_j} \right)}
\end{aligned} \tag{C.10}$$

$$\begin{aligned}
f(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) &= \frac{\frac{1}{\sqrt{2\pi\sigma_\mu^2}} \exp \left\{ -\frac{1}{2} \frac{n_j \sigma_\mu^2 + \sigma_y^2}{\sigma_y^2 \sigma_\mu^2} \left( \mu_{[i_j i_{j+1}]} - m^\star \right)^2 \right\}}{\frac{\sqrt{2\pi\sigma_y^2}}{\sqrt{n_j}} \frac{1}{\sqrt{\frac{2\pi\sigma_\mu^2 \sigma_y^2}{n_j \sigma_\mu^2 + \sigma_y^2}}} \left( \frac{1}{\sqrt{\frac{2\pi(n_j \sigma_\mu^2 + \sigma_y^2)}{n_j}}} \right)} \\
&\times \frac{\Phi_{n_j} \left( \lambda y_k | \lambda \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right)}{\Phi_{n_j} \left( -\lambda m^\star \mathbf{1}_{n_j} | -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^\star \mathbf{1}_{n_j} \mathbf{1}'_{n_j} \right)} \quad (\text{C.11}) \\
&= \frac{1}{\sqrt{\frac{2\pi\sigma_y^2 \sigma_\mu^2}{\sigma_\mu^2 n_j + \sigma_y^2}}} \exp \left\{ -\frac{1}{2} \frac{n_j \sigma_\mu^2 + \sigma_y^2}{\sigma_y^2 \sigma_\mu^2} \left( \mu_{[i_j i_{j+1}]} - m^\star \right)^2 \right\} \\
&\times \frac{\Phi_{n_j} \left( \lambda y_k | \lambda \mu_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} \right)}{\Phi_{n_j} \left( -\lambda m^\star \mathbf{1}_{n_j} | -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^\star \mathbf{1}_{n_j} \mathbf{1}'_{n_j} \right)}
\end{aligned}$$

A função de densidade de probabilidade em C.11 pode ser reescrita da seguinte forma

$$f(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) = \frac{\phi \left( \mu_{[i_j i_{j+1}]} | m^\star, V^\star \right) \Phi_{n_j} \left( -\lambda \mu_{[i_j i_{j+1}]} | -\lambda y_k, \sigma_y^2 \mathbf{I}_{n_j} \right)}{\Phi_{n_j} \left( -\lambda m^\star \mathbf{1}_{n_j} | -\lambda \mathbf{y}_{[i_j i_{j+1}]}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^\star \mathbf{1}_{n_j} \mathbf{1}'_{n_j} \right)}. \quad (\text{C.12})$$

Note da Definição 3 que a distribuição em C.12 pertence a classe de distribuições normal assimétrica multivariada generalizada denotada por

$$\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]} \sim \mathcal{GMSN}_{1, n_j}(\xi_{[i_j i_{j+1}]}, \Sigma, D, v, \Delta),$$

onde  $\xi_{[i_j i_{j+1}]} = m^\star$ ,  $\Sigma = V^\star$ ,  $D = -\lambda \mathbf{1}_{n_j}$ ,  $v = -\lambda \mathbf{y}_{[i_j i_{j+1}]}$  e  $\Delta = \sigma_y^2 \mathbf{I}_{n_j}$ .

iii) Consequentemente, o valor esperado de  $\mu_{i_j i_{j+1}}$ , dado  $y_{i_j i_{j+1}}$  é

$$E(\mu_{[i_j i_{j+1}]} | \mathbf{y}_{[i_j i_{j+1}]}) = m^\star + \frac{\lambda V^\star \phi_{n_j}(\lambda m^\star \mathbf{1}_{n_j} | \lambda \mathbf{y}_{i_j i_{j+1}}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^\star \mathbf{1}_{n_j} \mathbf{1}'_{n_j})}{\Phi_{n_j}(\lambda m^\star \mathbf{1}_{n_j} | \lambda \mathbf{y}_{i_j i_{j+1}}, \sigma_y^2 \mathbf{I}_{n_j} + \lambda^2 V^\star \mathbf{1}_{n_j} \mathbf{1}'_{n_j})},$$

concluindo a prova. ■

---

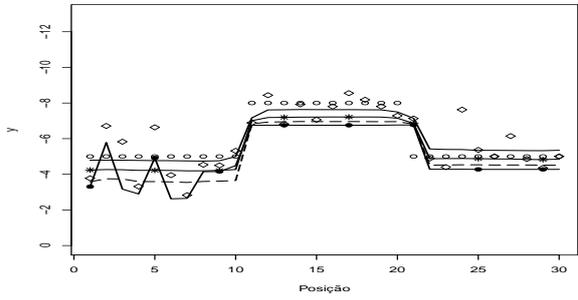
### Gráficos da sensibilidade informal na verossimilhança

---

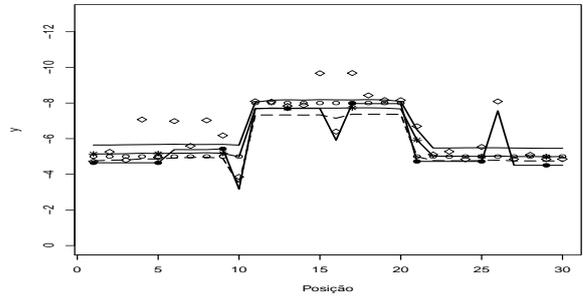
Este apêndice contém os demais gráficos de sensibilidade informal na verossimilhança omitidos ao longo do texto por apresentarem resultados semelhantes aos que foram discutidos e analisados neste trabalho.

#### **D.1 Sensibilidade informal nas Estimativas produto**

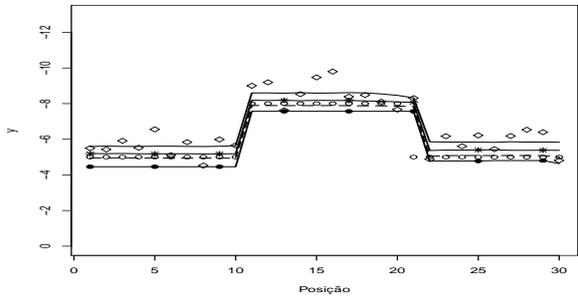
##### **D.1.1 A influência de $\lambda$ e $n$ nas estimativas produto e na variância *a posteriori* das médias**



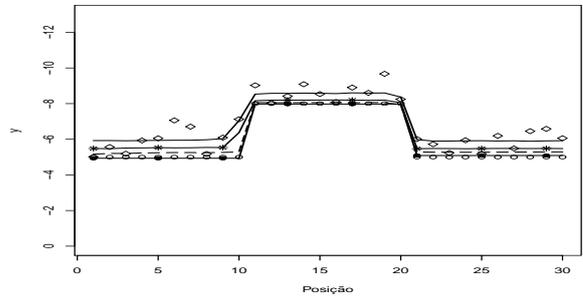
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



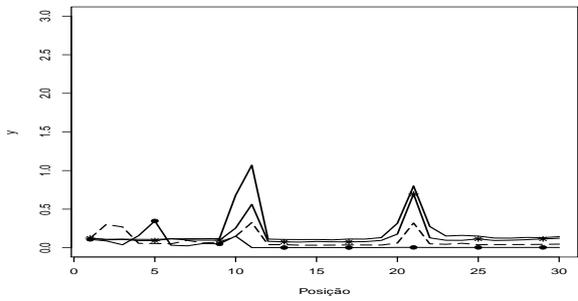
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



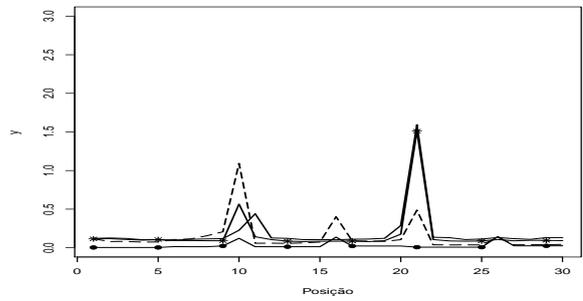
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



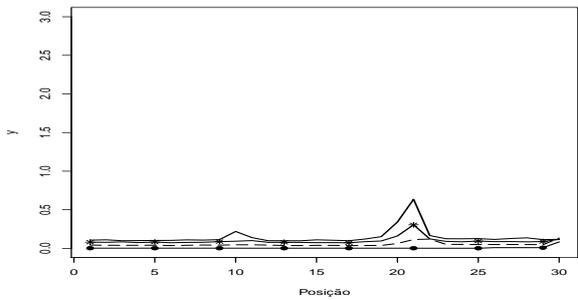
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



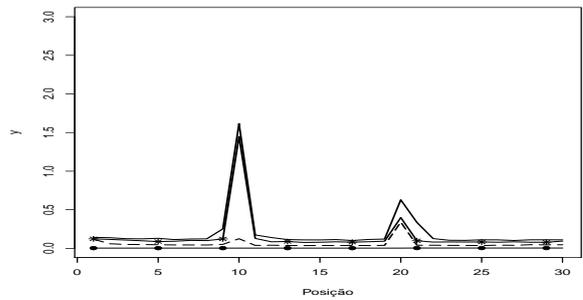
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$

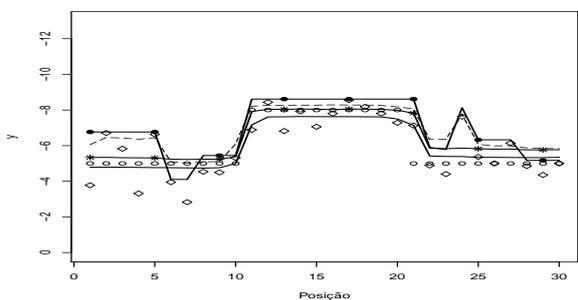


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$

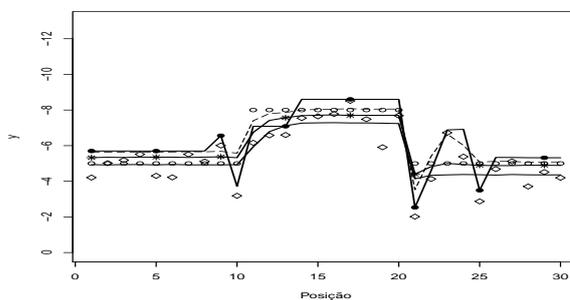


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

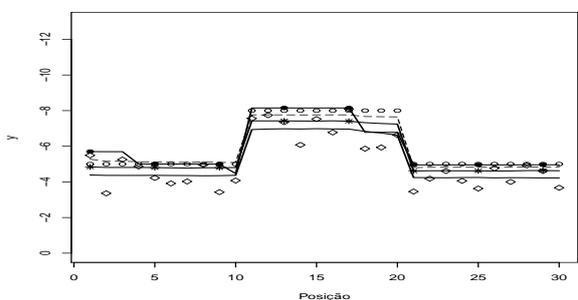
Figura D.1: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário de grau negativo,  $n = 30$



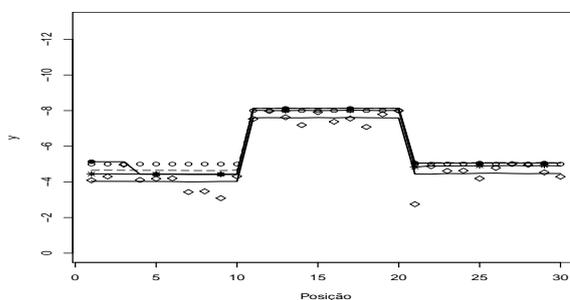
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



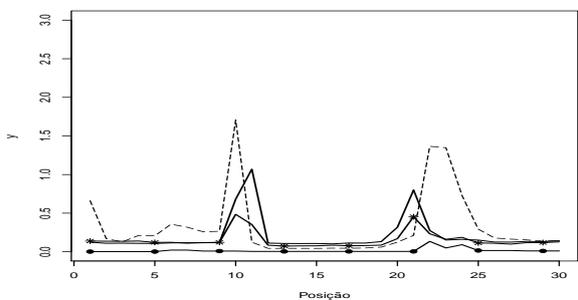
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



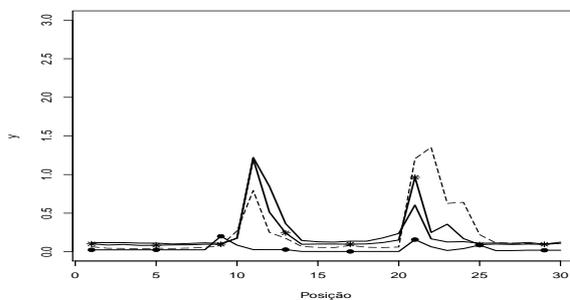
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$



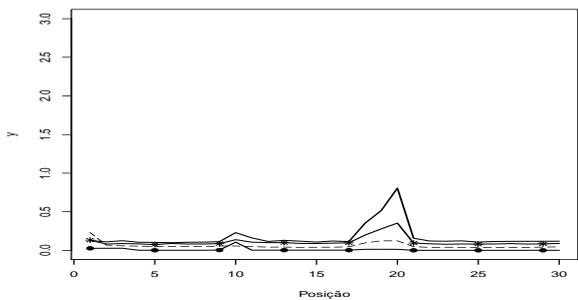
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$



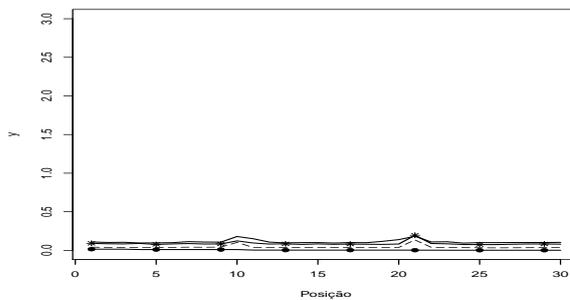
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$

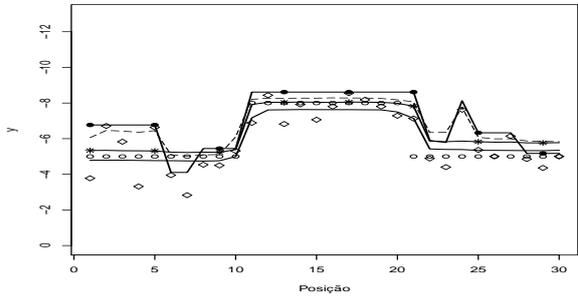


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

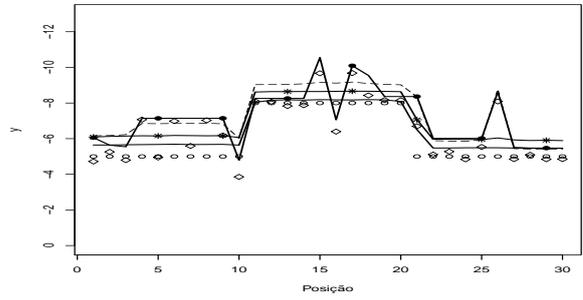


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

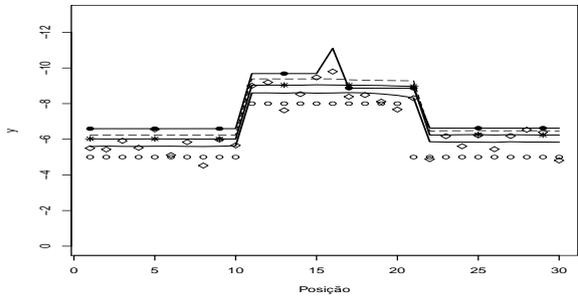
Figura D.2: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário de grau negativo,  $n = 30$ .



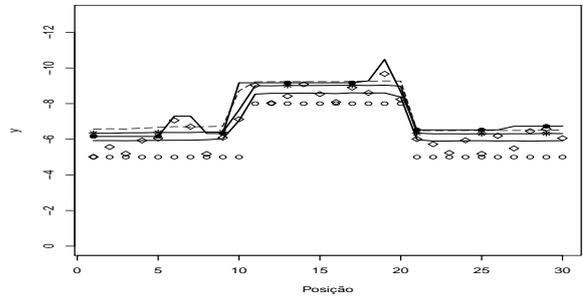
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



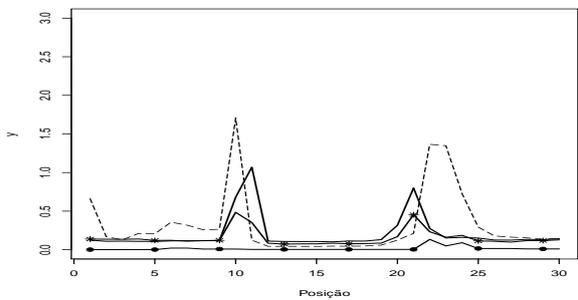
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



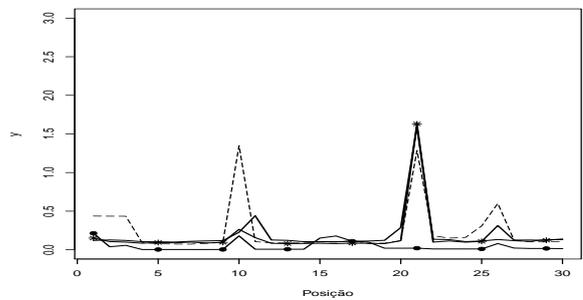
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$



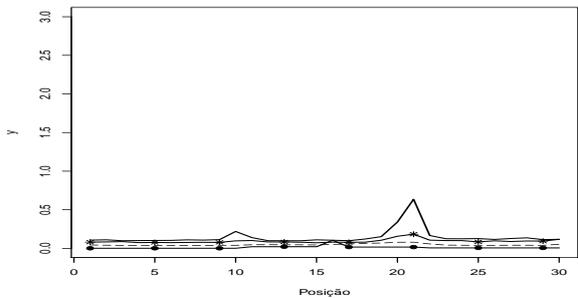
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$



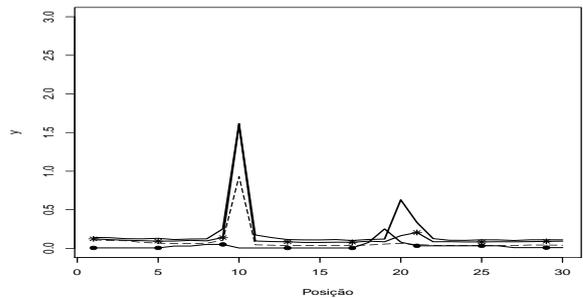
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$

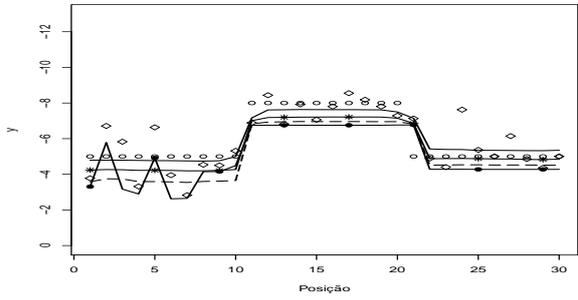


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

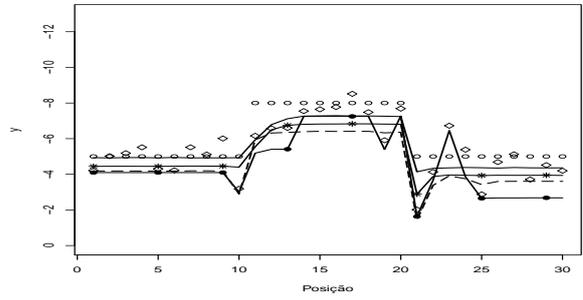


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

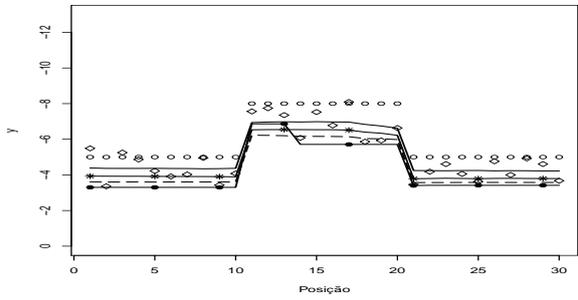
Figura D.3: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 30$ .



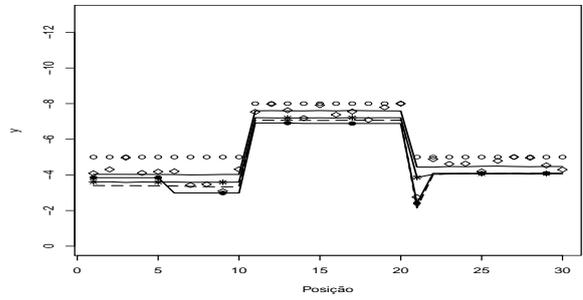
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



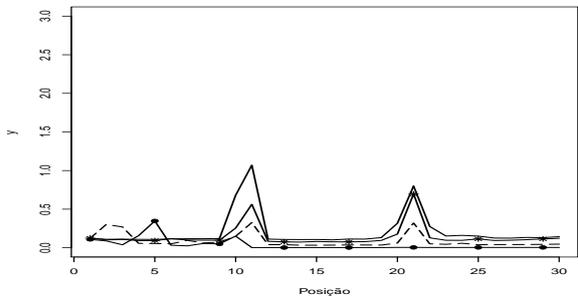
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



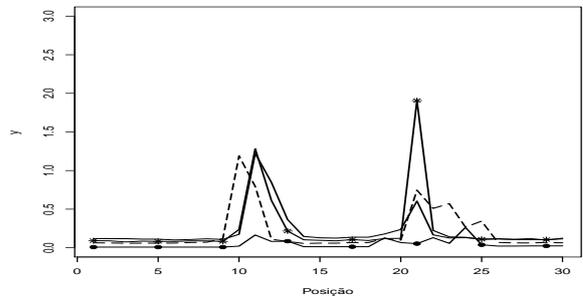
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



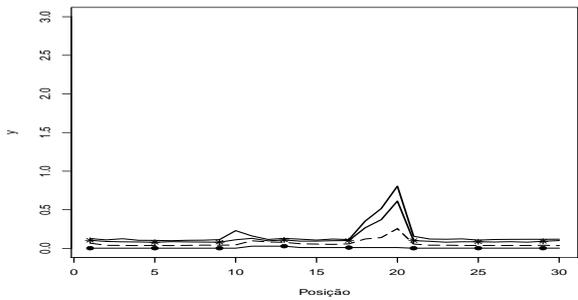
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$



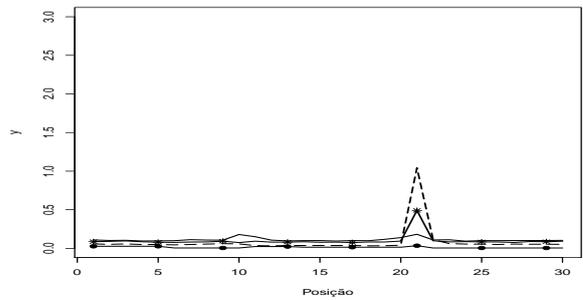
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$

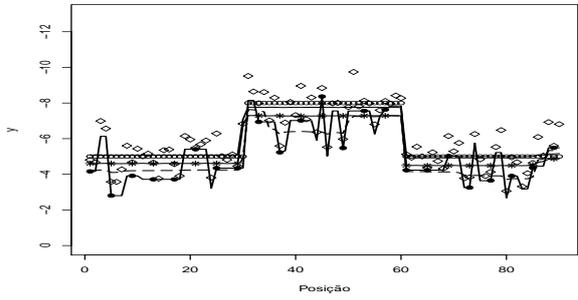


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$

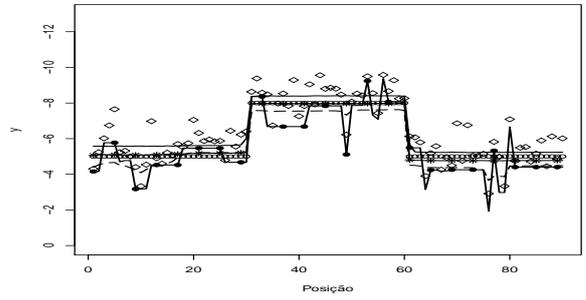


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

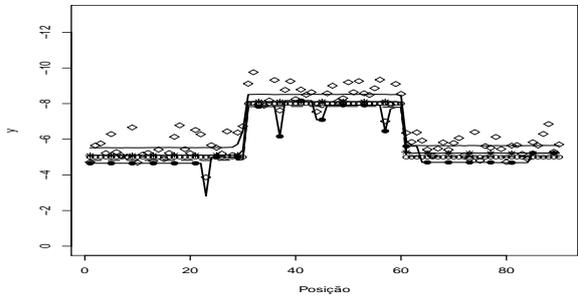
Figura D.4: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 30$ .



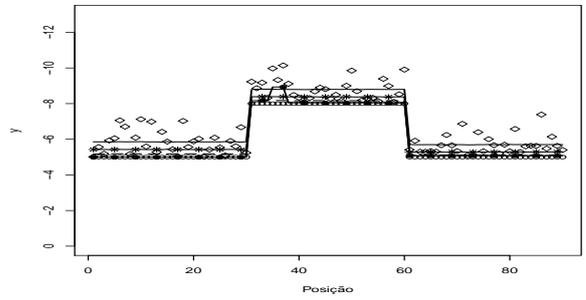
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



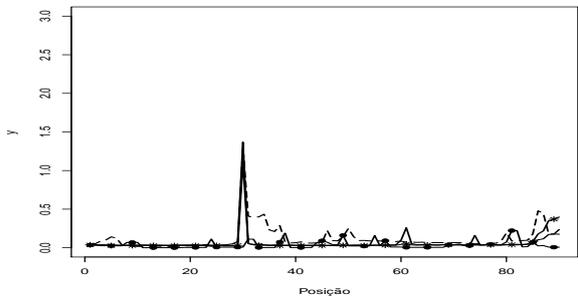
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



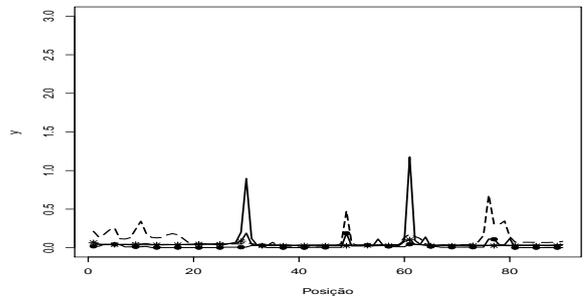
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



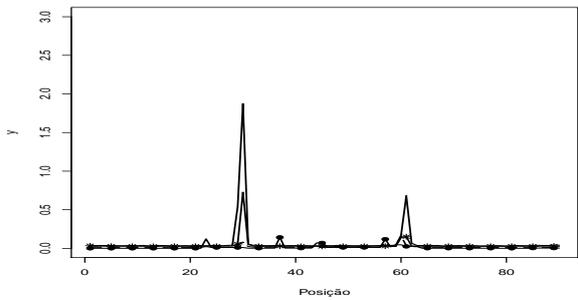
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



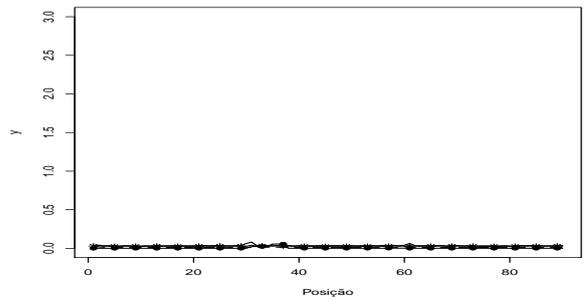
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$

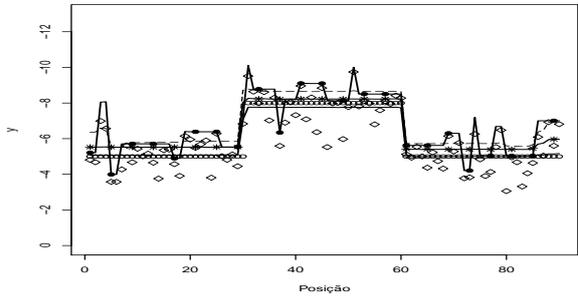


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$

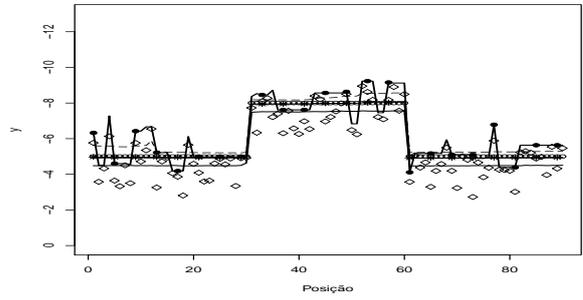


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

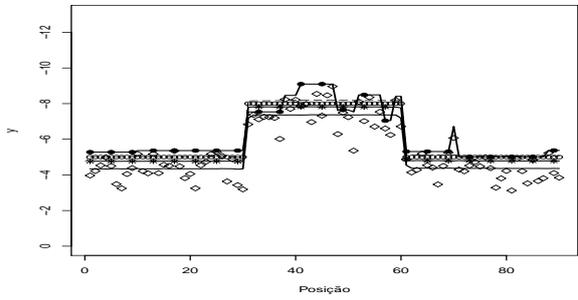
Figura D.5: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 90$ .



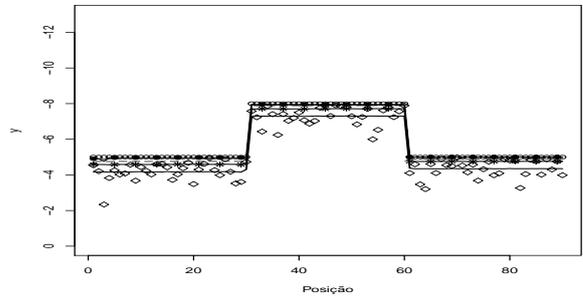
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



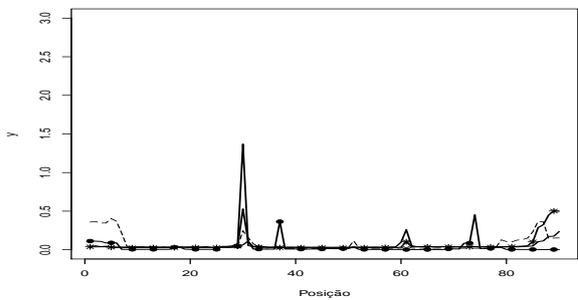
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



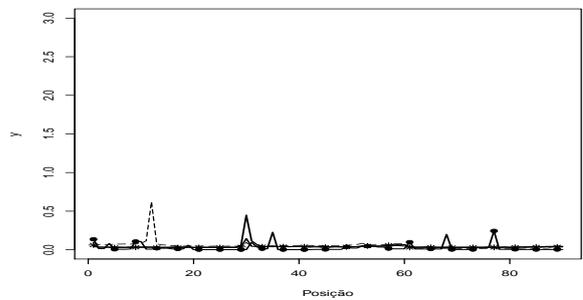
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$



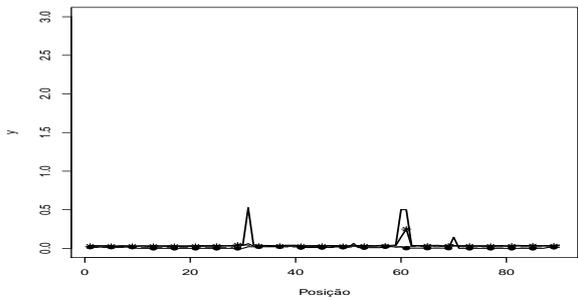
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$



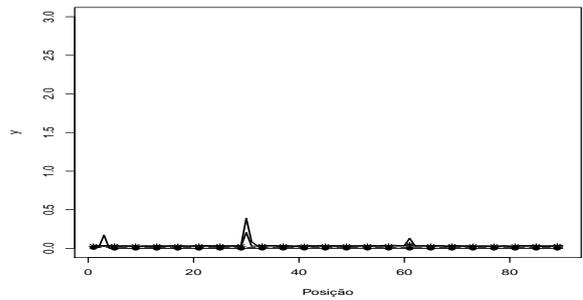
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$

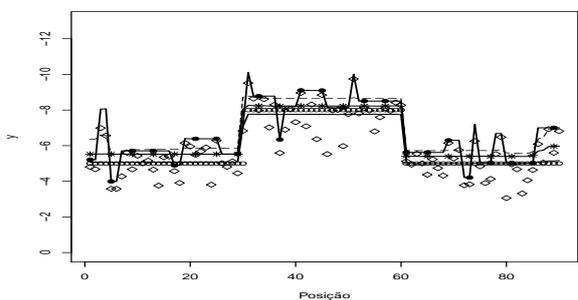


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

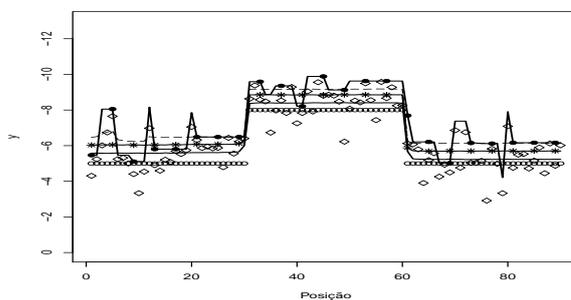


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

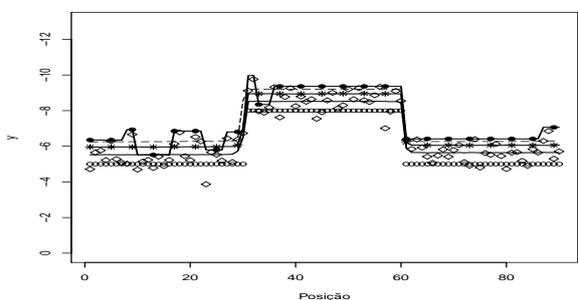
Figura D.6: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 90$ .



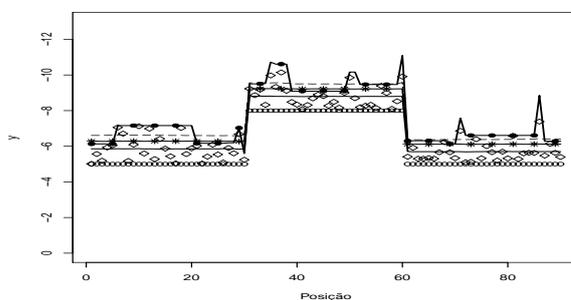
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



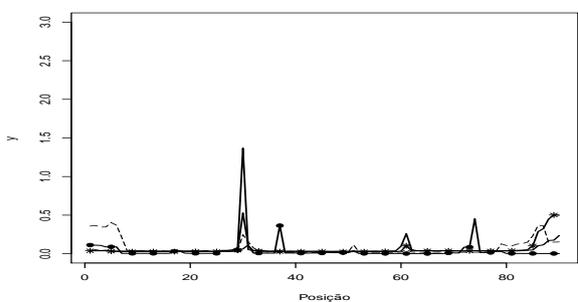
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



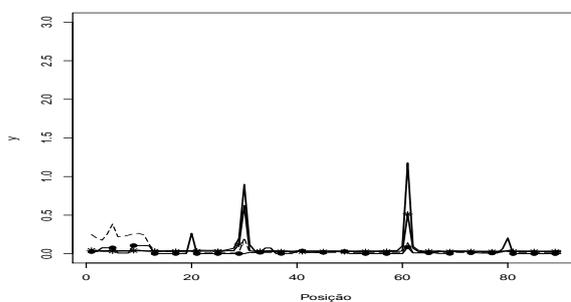
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$



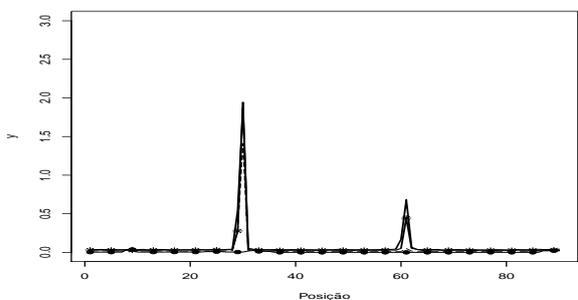
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$



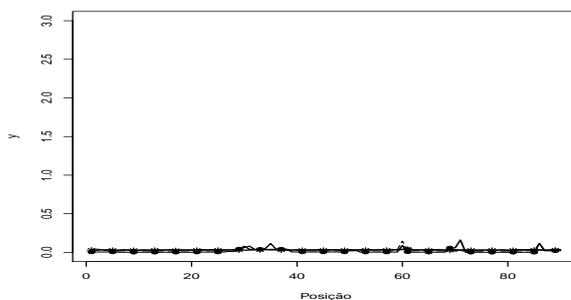
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$

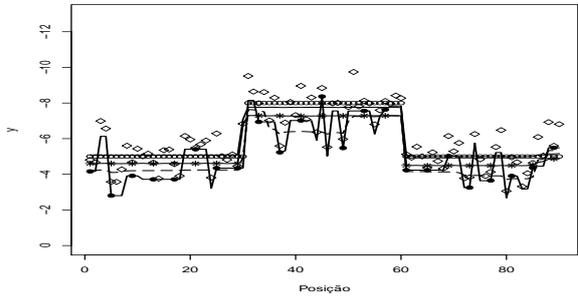


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

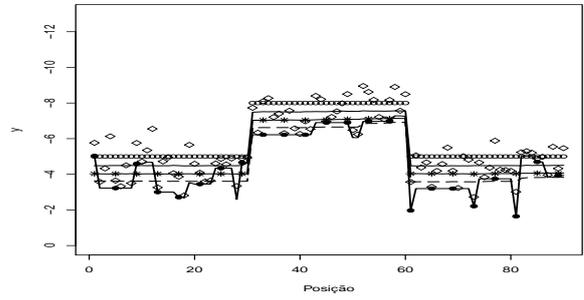


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

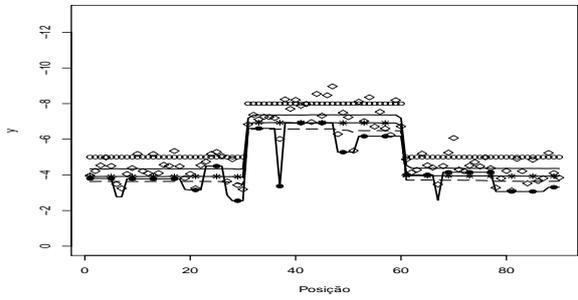
Figura D.7: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário de grau negativo,  $n = 90$ .



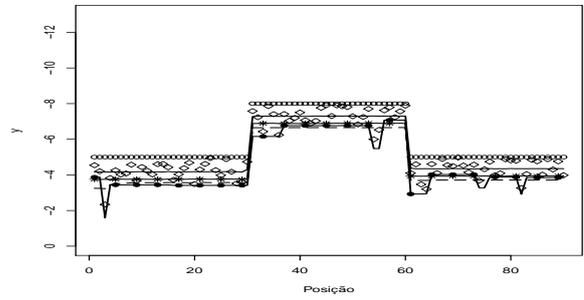
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



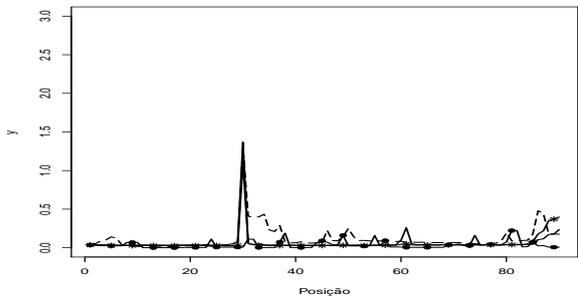
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



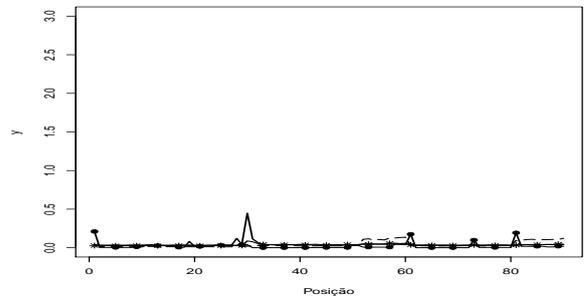
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



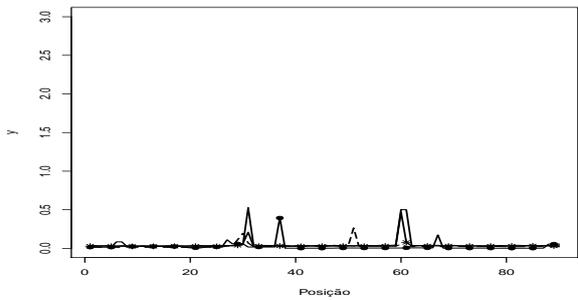
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$



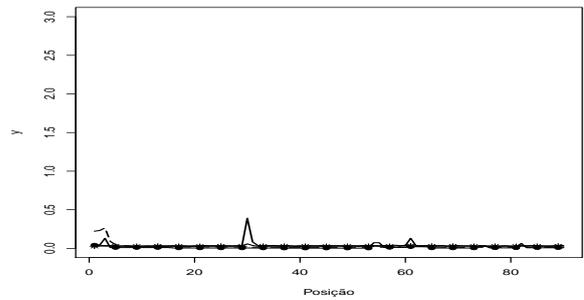
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$

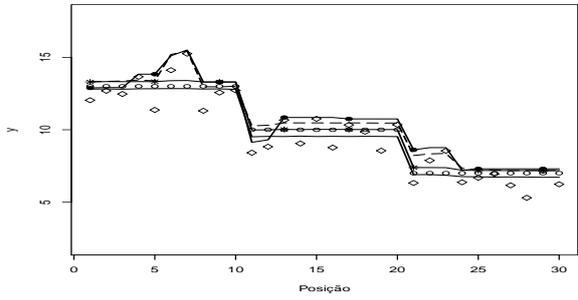


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$

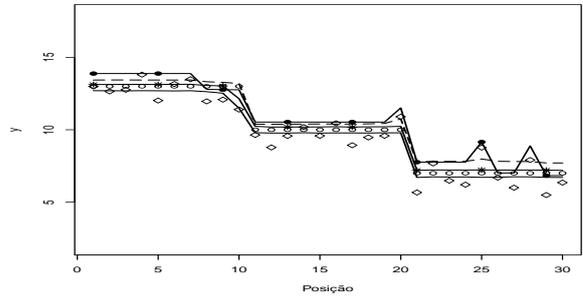


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

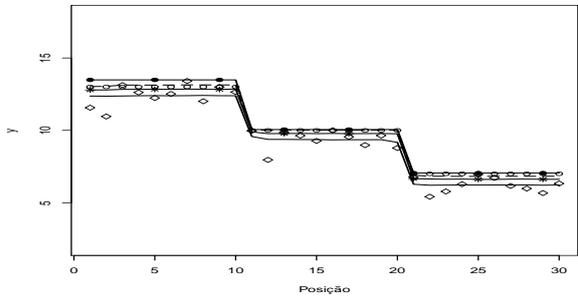
Figura D.8: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 90$ .



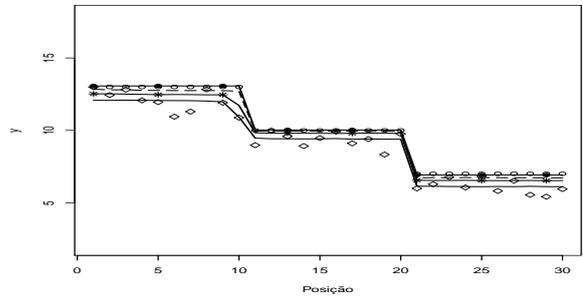
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



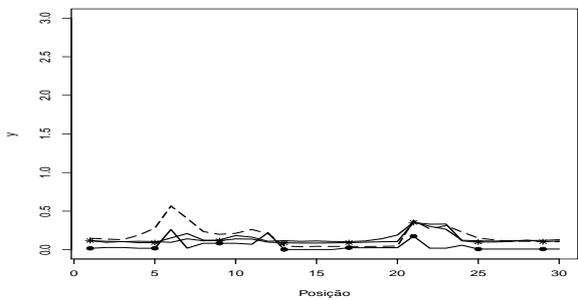
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



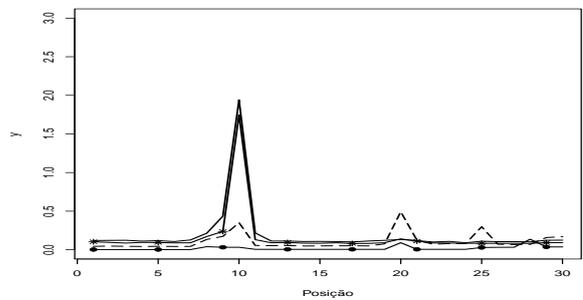
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



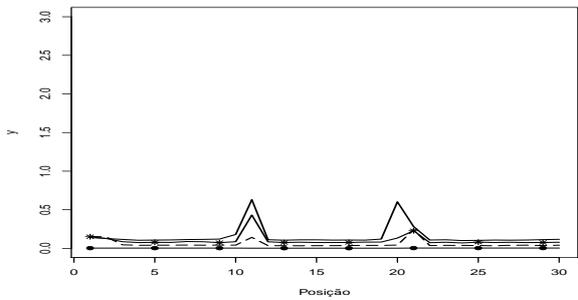
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



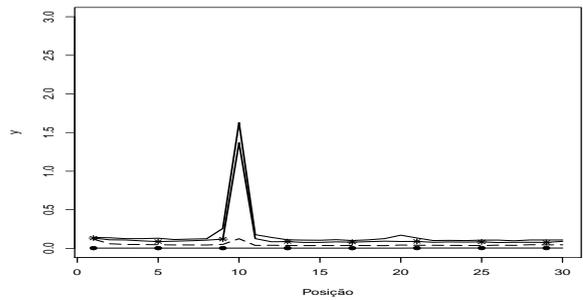
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$

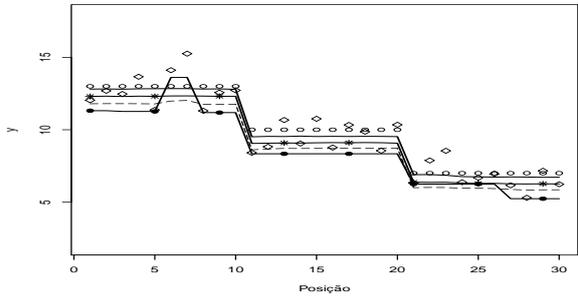


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$

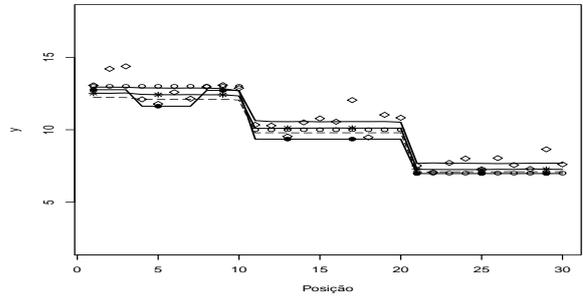


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

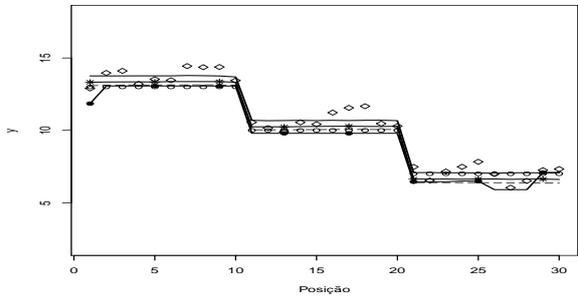
Figura D.9: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 30$



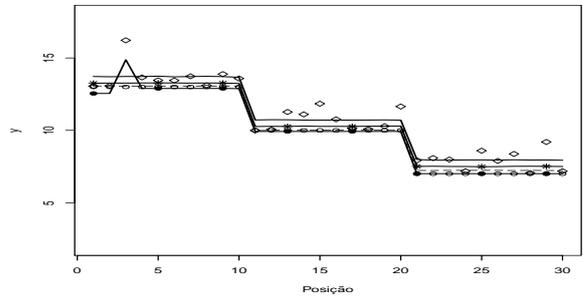
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



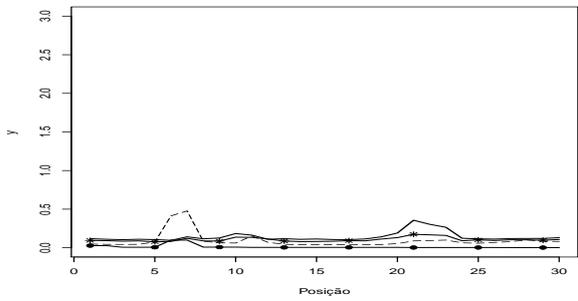
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



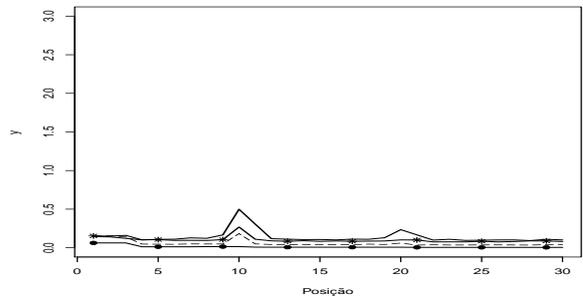
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$



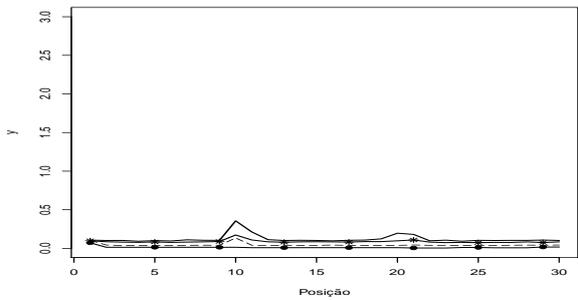
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$



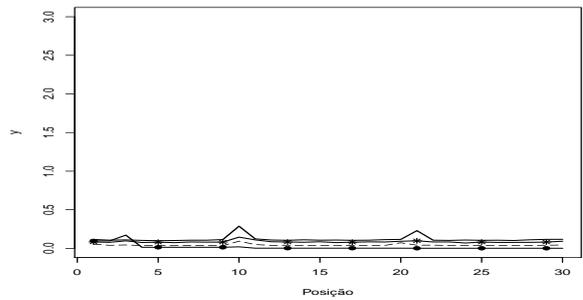
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$

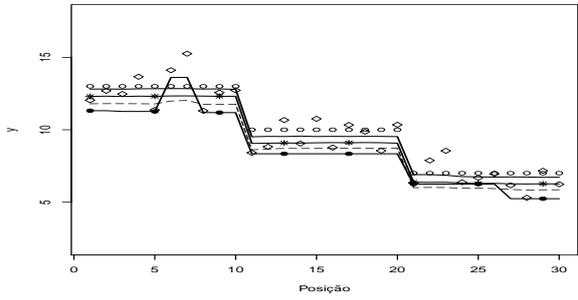


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

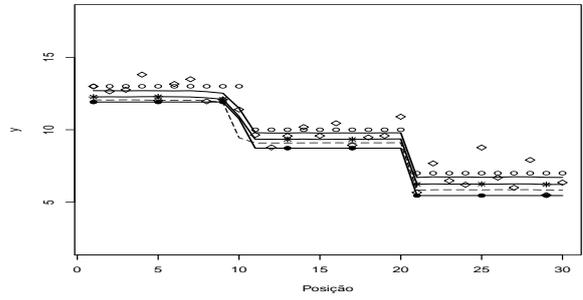


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

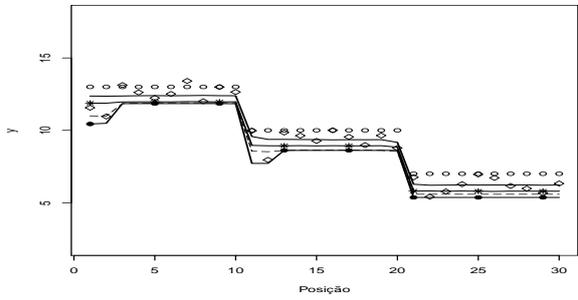
Figura D.10: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 30$ .



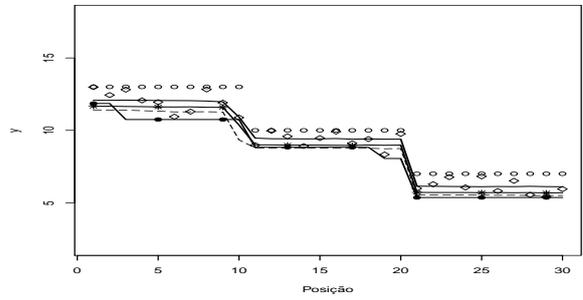
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



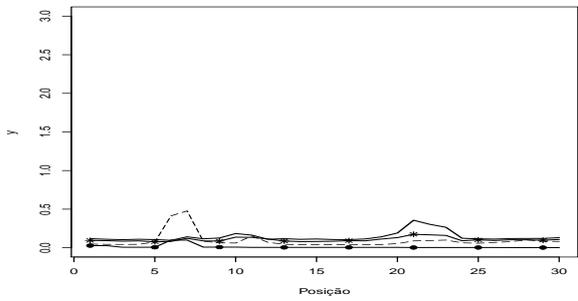
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



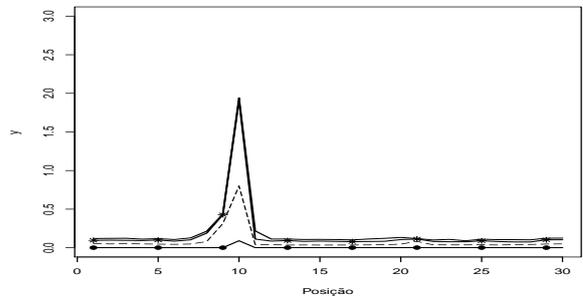
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$



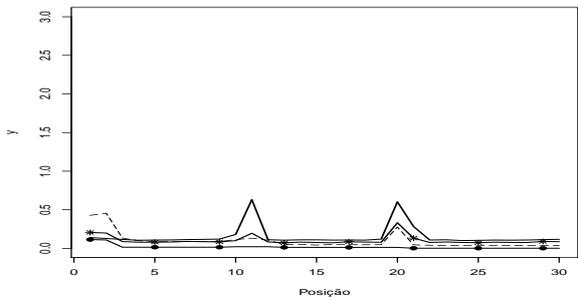
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$



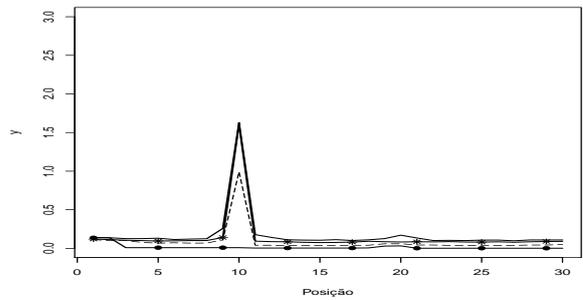
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$

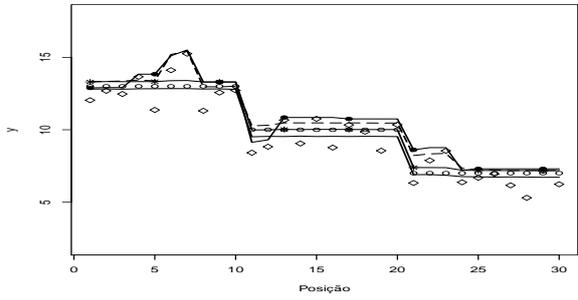


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

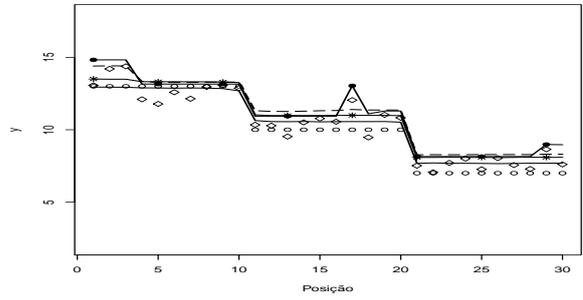


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

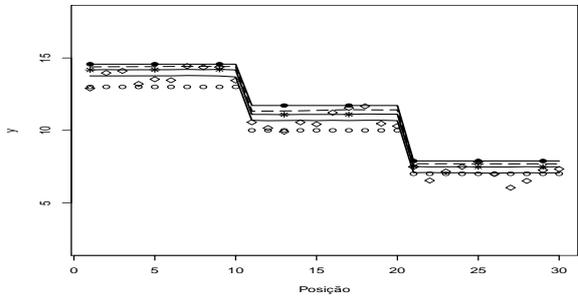
Figura D.11: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 30$ .



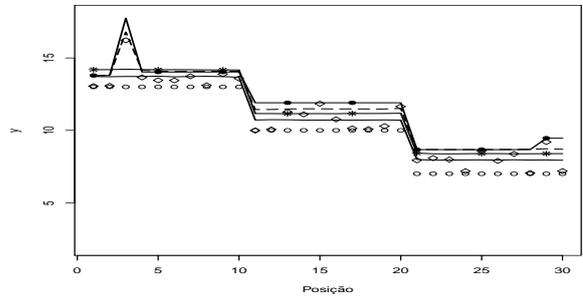
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



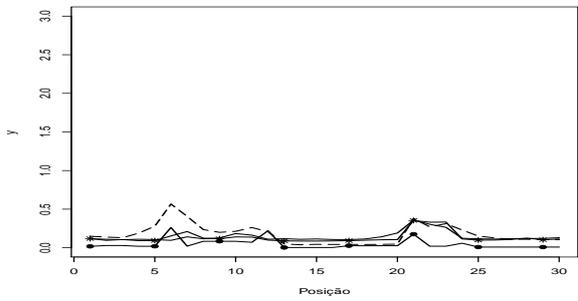
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



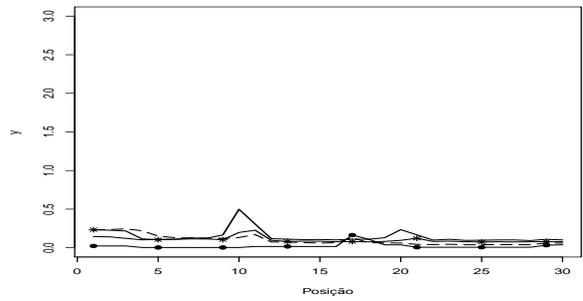
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



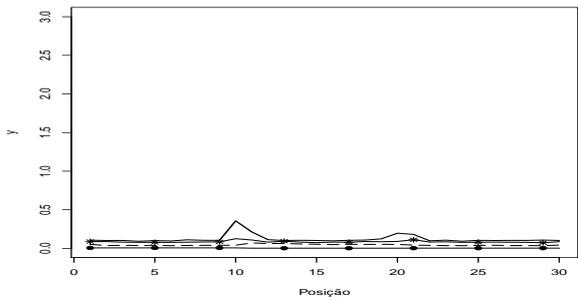
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$



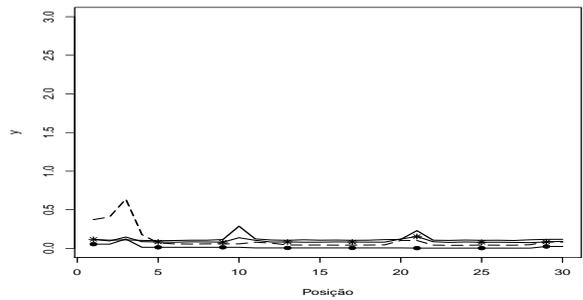
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$

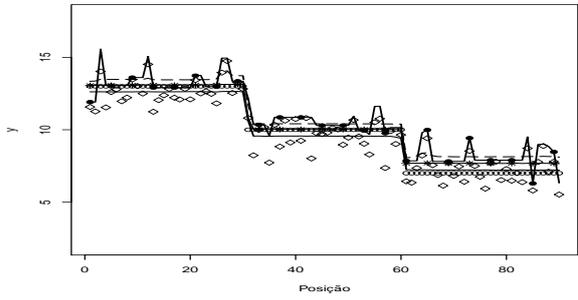


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$

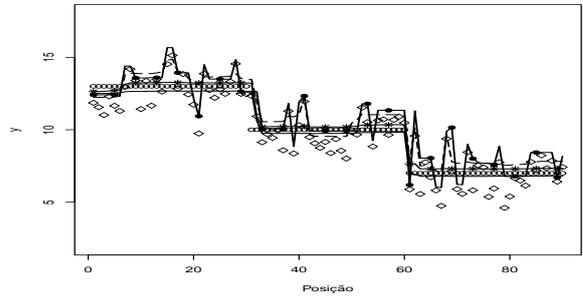


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

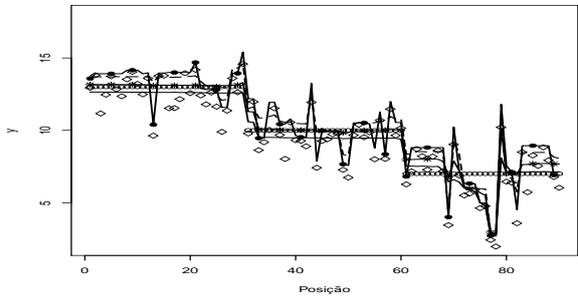
Figura D.12: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 30$ .



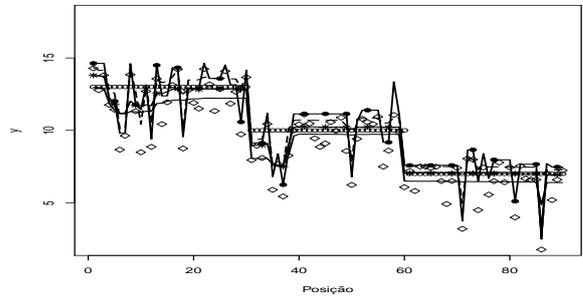
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



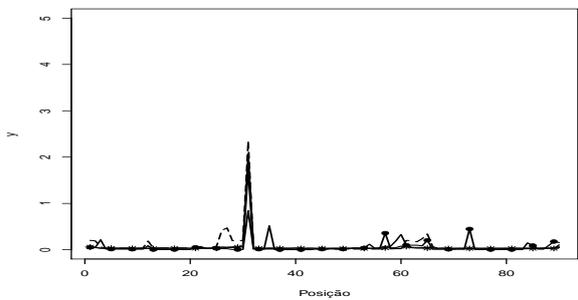
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



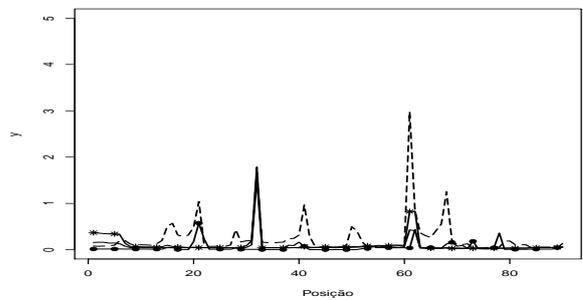
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



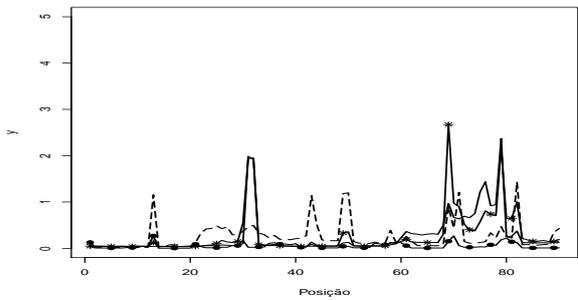
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



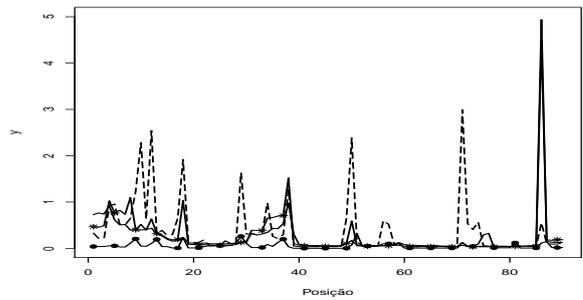
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$

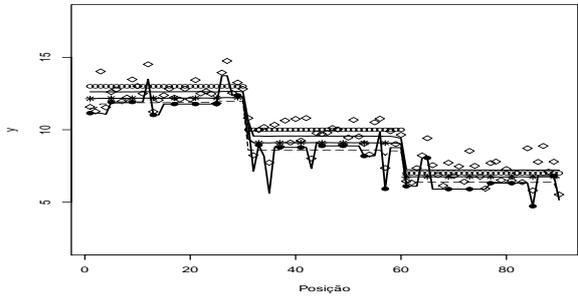


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$

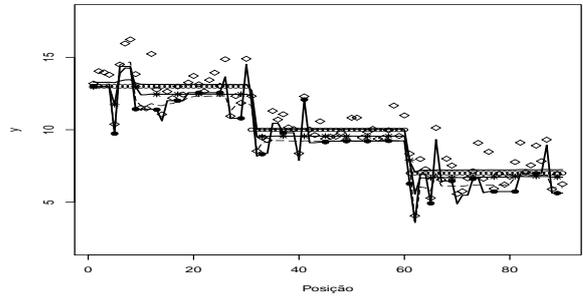


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$

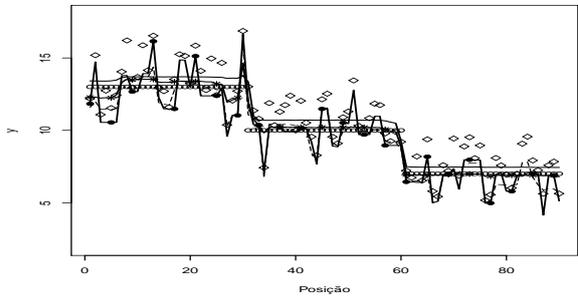
Figura D.13: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 90$ .



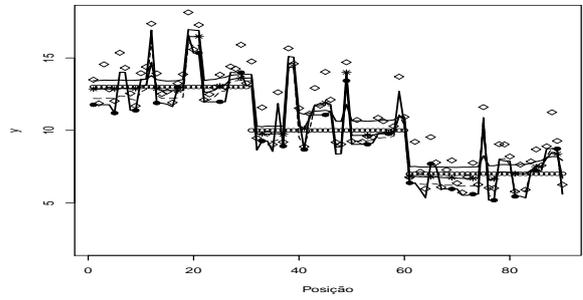
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



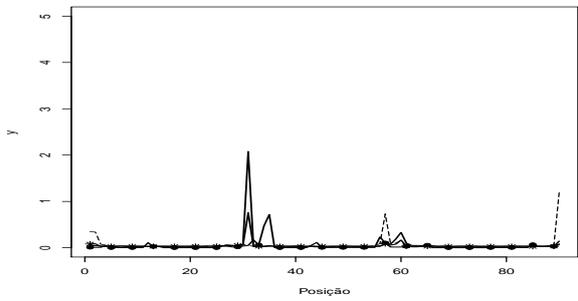
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



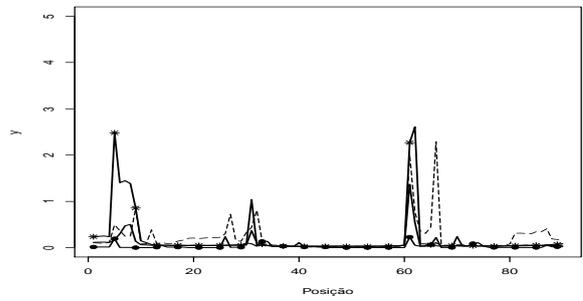
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$



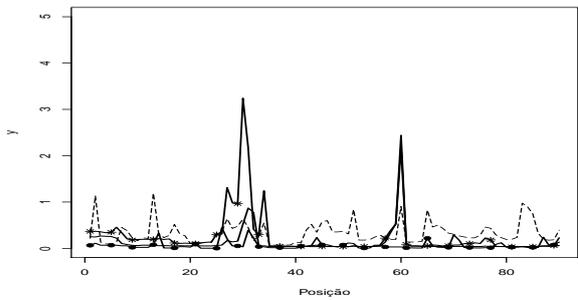
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$



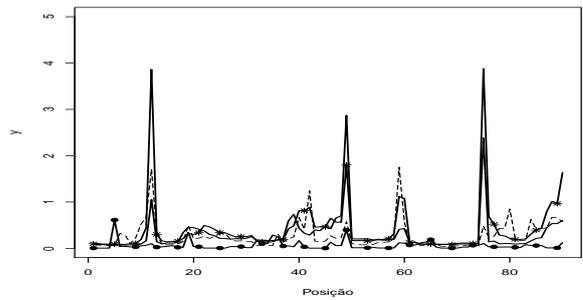
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$

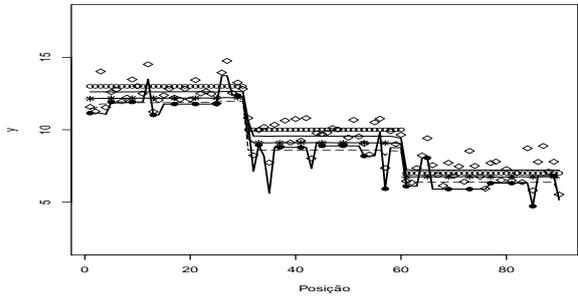


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

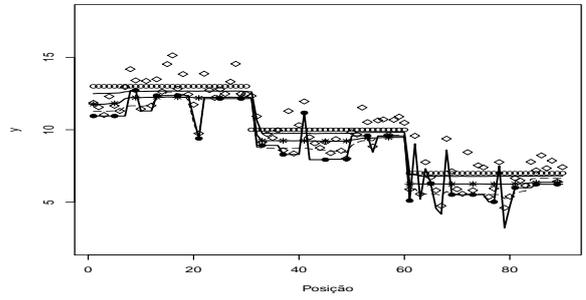


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

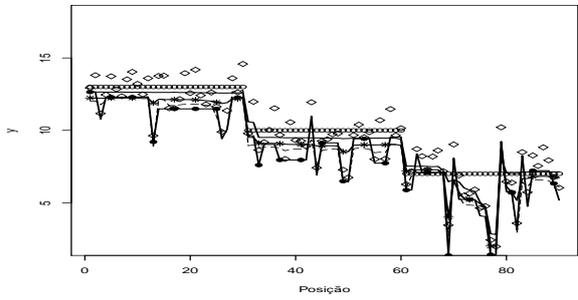
Figura D.14: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 90$ .



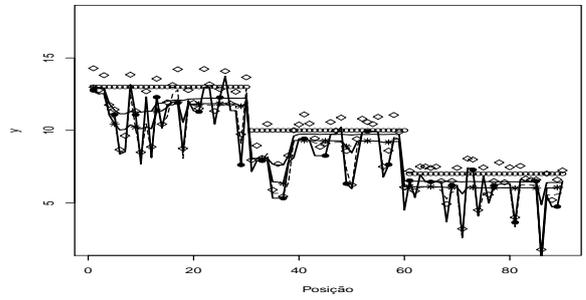
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



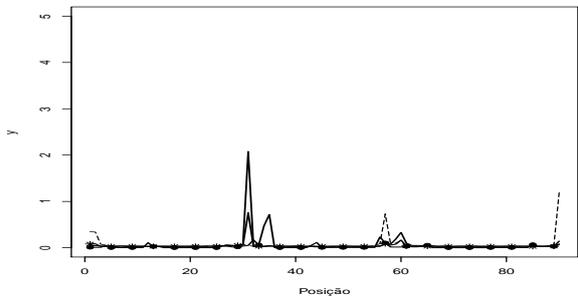
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$



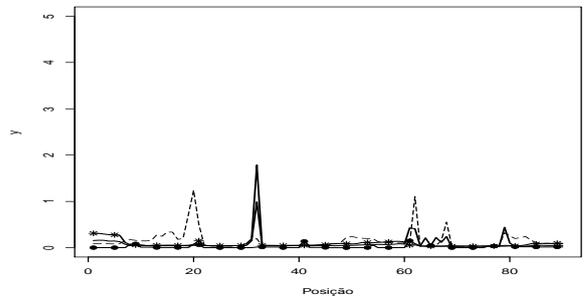
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$



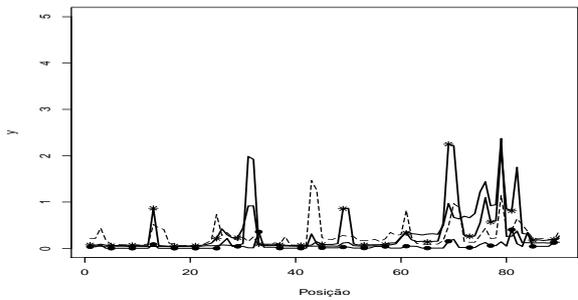
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$



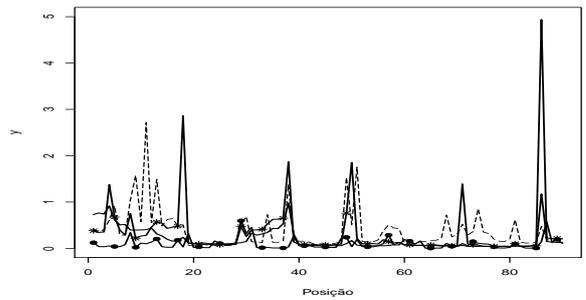
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$

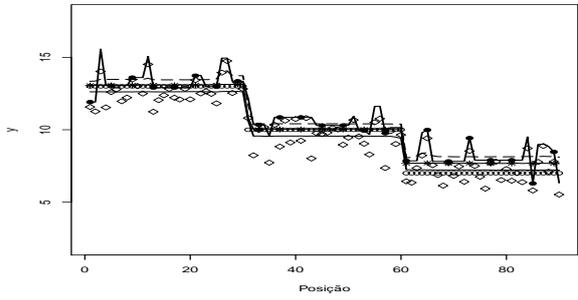


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

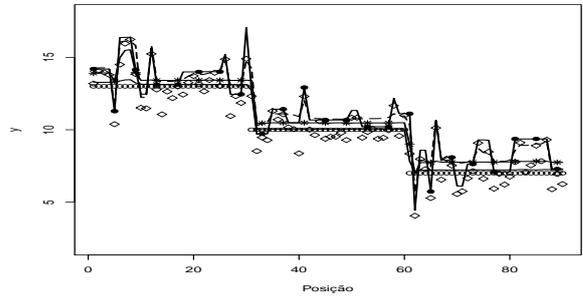


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

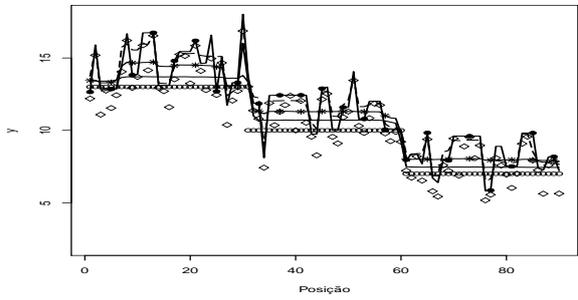
Figura D.15: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 90$ .



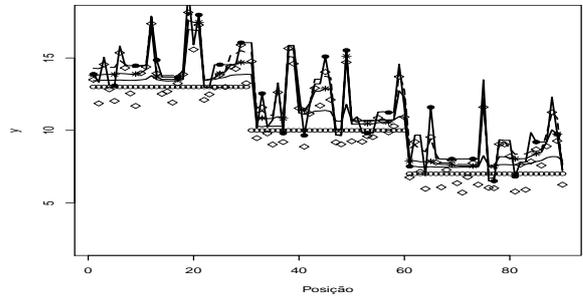
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



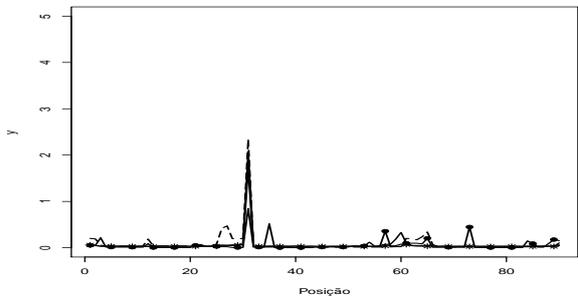
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



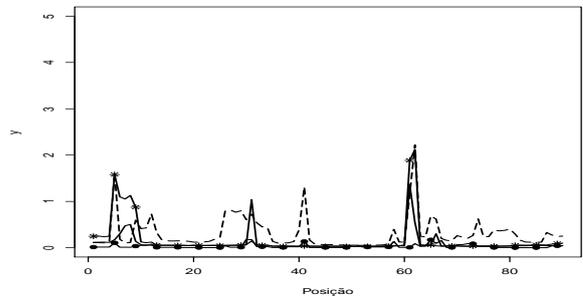
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



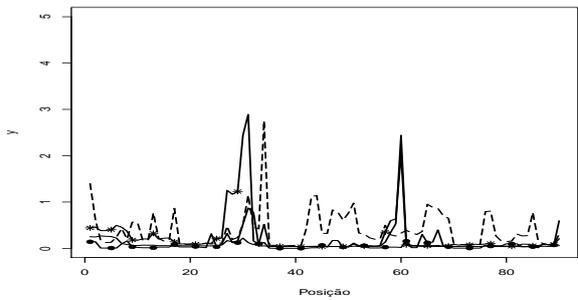
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$



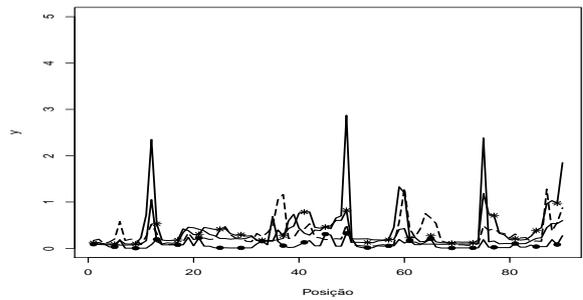
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



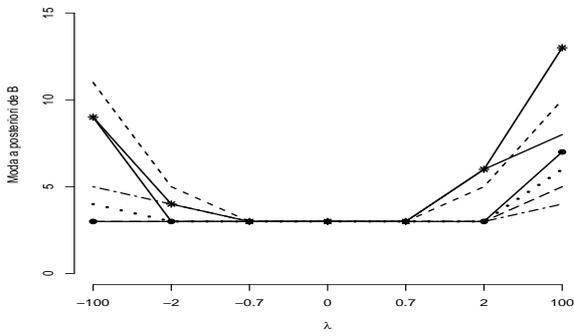
(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



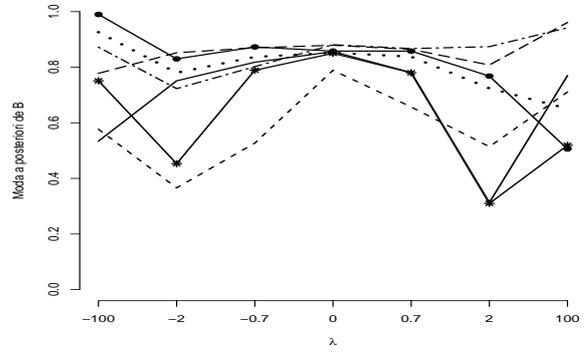
(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

Figura D.16: Estimativas produto e variância *a posteriori* para cada  $\mu_i$ . Média verdadeira ( $\circ$ ), Estimativas para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 90$ .

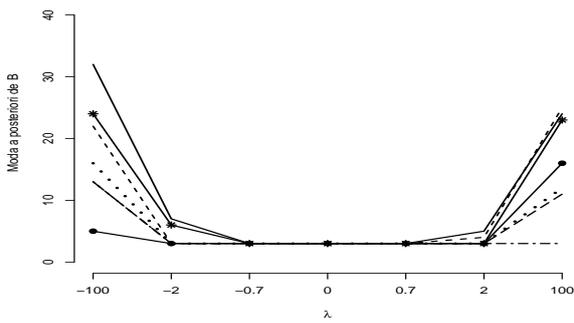
### D.1.2 A influência de $\lambda$ e $n$ nas inferências sobre $\rho$ e $B$



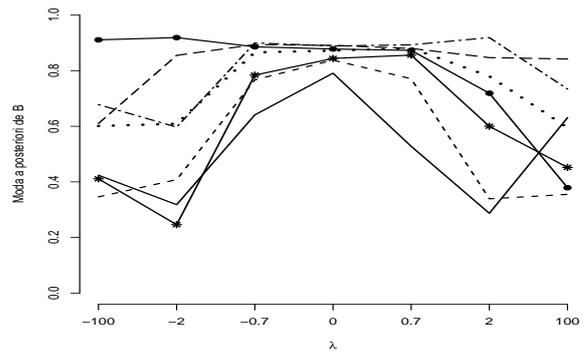
(a) Número Blocos =  $B' - n = 30$



(b)  $P(B') - n = 30$

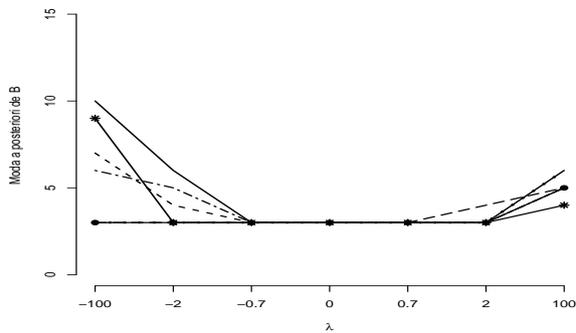


(c) Número Blocos =  $B' - n = 90$

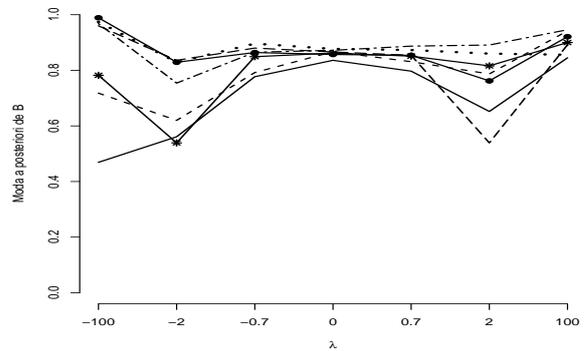


(d)  $P(B') - n = 90$

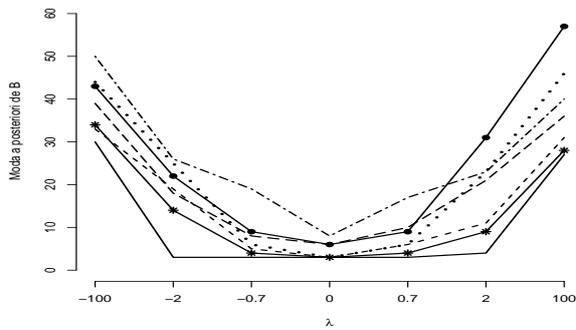
Figura D.17: Moda a *posteriori* de  $B$  e sua probabilidade para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com \*),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia),  $\lambda =$  (linha tracejada curta),  $\lambda =$  (linha pontilhada),  $\lambda =$  (linha tracejada e pontilhada), cenário escada decrescente.



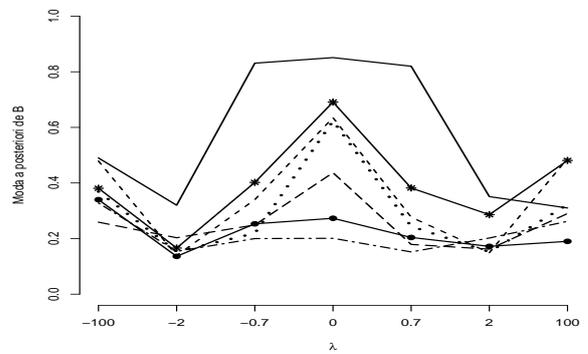
(a) Número Blocos =  $B' - n = 30$



(b)  $P(B') - n = 30$

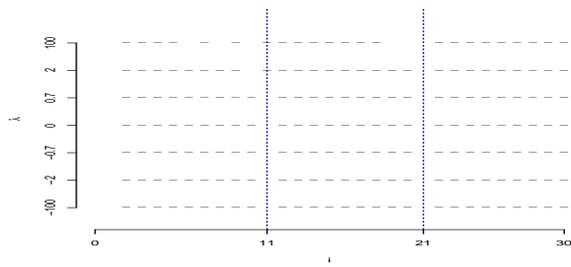


(c) Número Blocos =  $B' - n = 90$

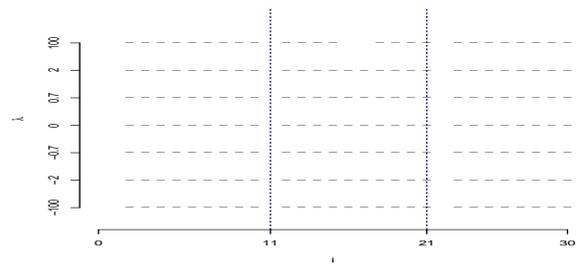


(d)  $P(B') - n = 90$

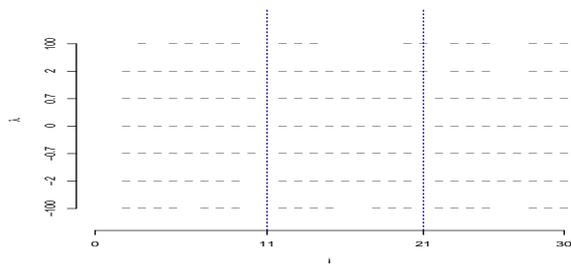
Figura D.18: Moda a *posteriori* de  $B$  e sua probabilidade para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com \*),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia),  $\lambda =$  (linha tracejada curta),  $\lambda =$  (linha pontilhada),  $\lambda =$  (linha tracejada e pontilhada), cenário escada decrescente.



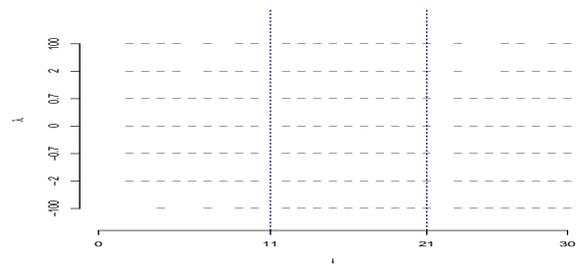
(a) Partição mais provável - Dados  $\lambda = -100$



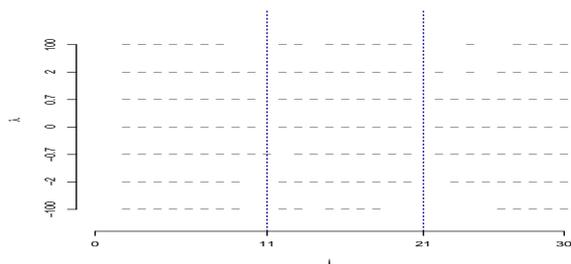
(b) Partição mais provável - Dados  $\lambda = -2$



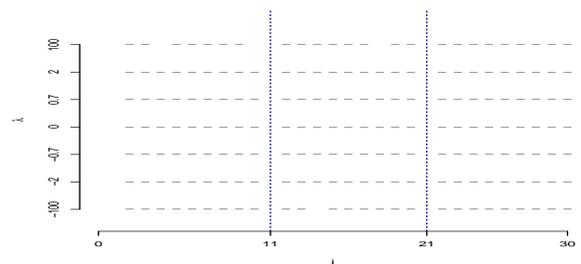
(c) Partição mais provável - Dados  $\lambda = -0.7$



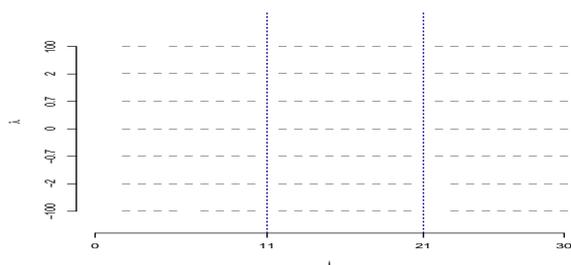
(d) Partição mais provável - Dados  $\lambda = 0$



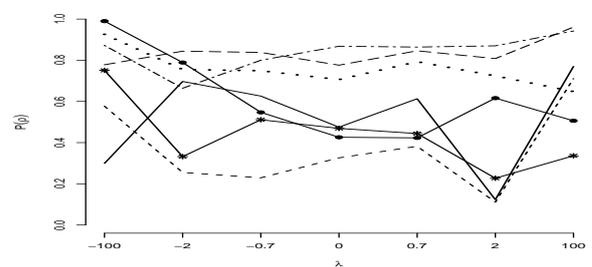
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

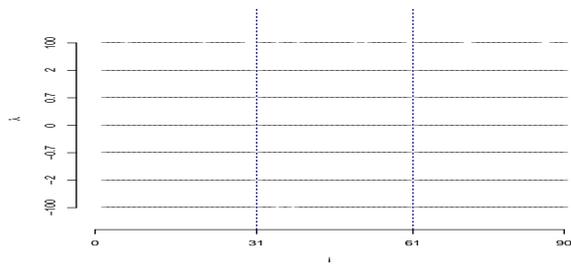


(g) Partição mais provável - Dados  $\lambda = 100$

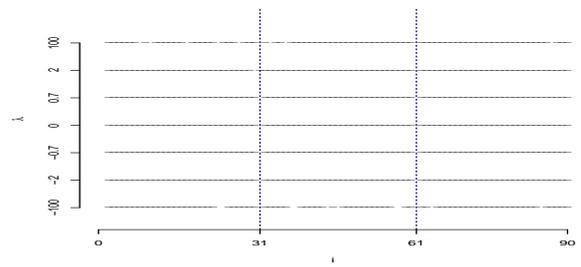


(h)  $P(\rho = \rho')$

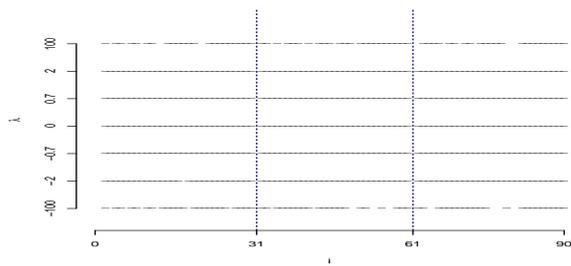
Figura D.19: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com \*),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia),  $\lambda =$  (linha tracejada curta),  $\lambda =$  (linha pontilhada),  $\lambda =$  (linha tracejada e pontilhada), cenário degrau negativo,  $n = 30$ .



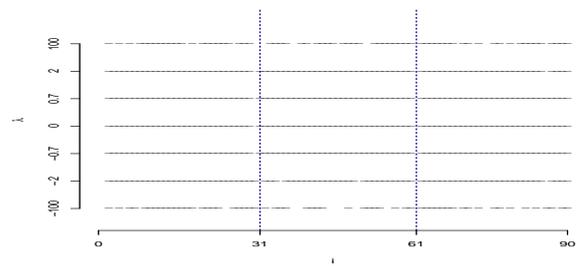
(a) Partição mais provável - Dados  $\lambda = -100$



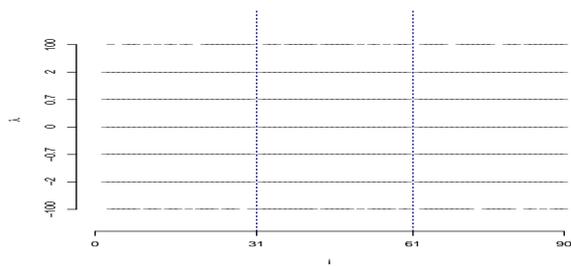
(b) Partição mais provável - Dados  $\lambda = -2$



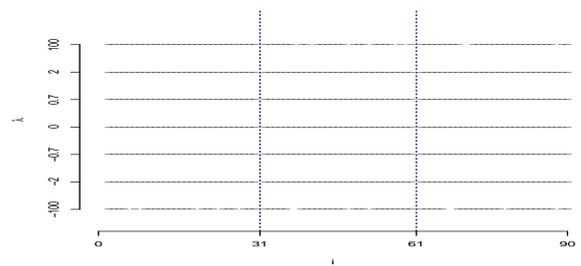
(c) Partição mais provável - Dados  $\lambda = -0.7$



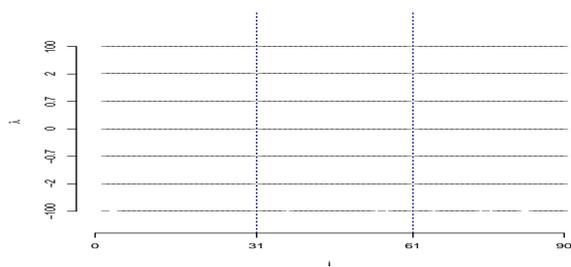
(d) Partição mais provável - Dados  $\lambda = 0$



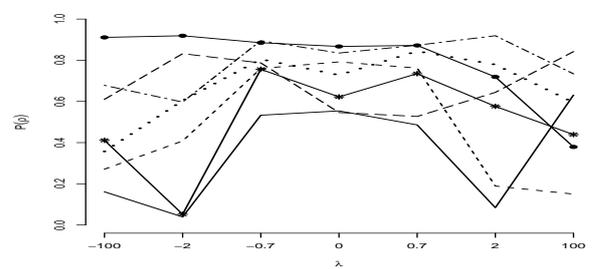
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

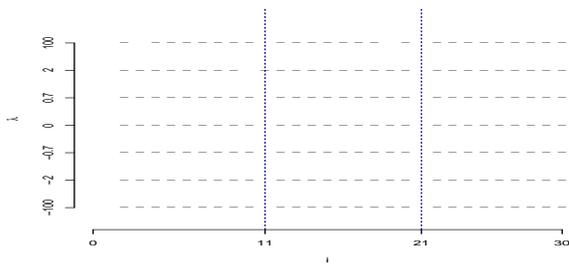


(g) Partição mais provável - Dados  $\lambda = 100$

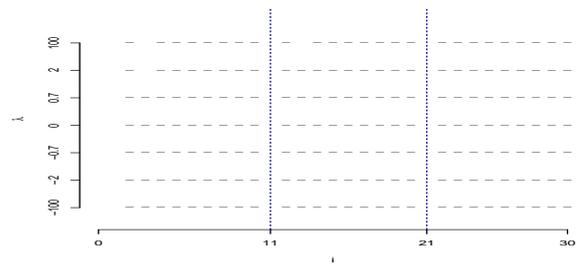


(h)  $P(\rho = \rho')$

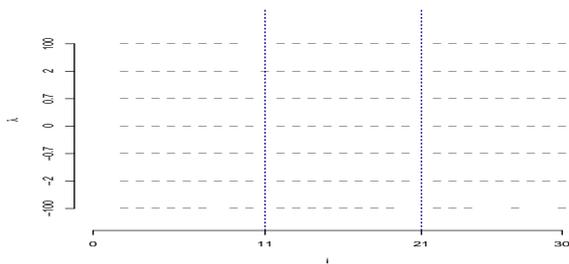
Figura D.20: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com \*),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário degrau negativo,  $n = 90$ .



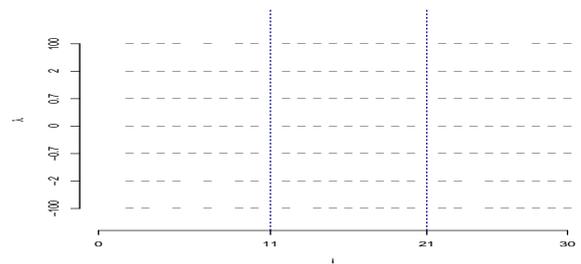
(a) Partição mais provável - Dados  $\lambda = -100$



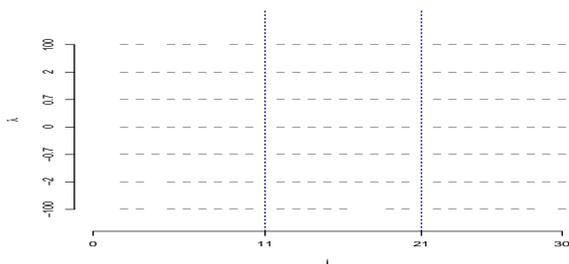
(b) Partição mais provável - Dados  $\lambda = -2$



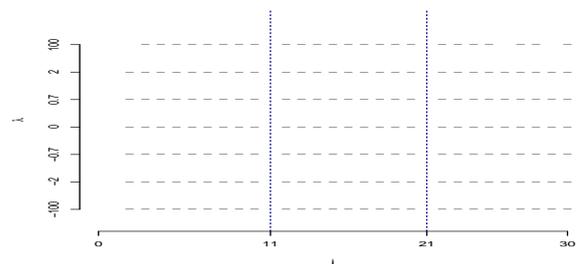
(c) Partição mais provável - Dados  $\lambda = -0.7$



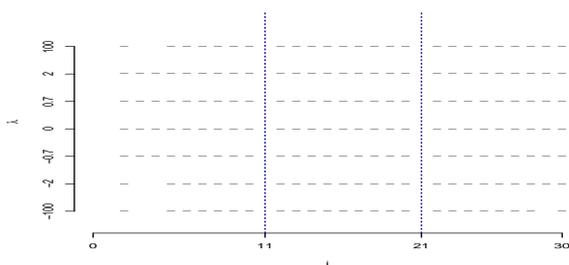
(d) Partição mais provável - Dados  $\lambda = 0$



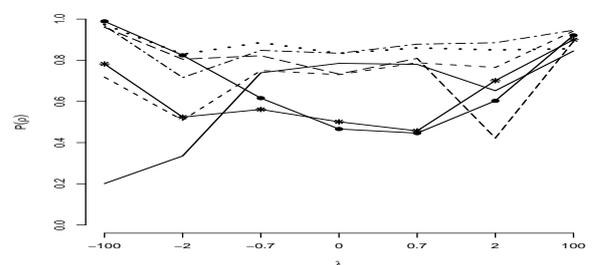
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$

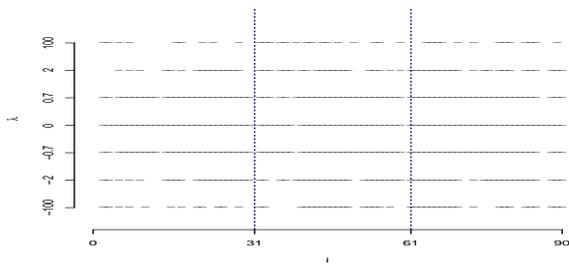


(g) Partição mais provável - Dados  $\lambda = 100$

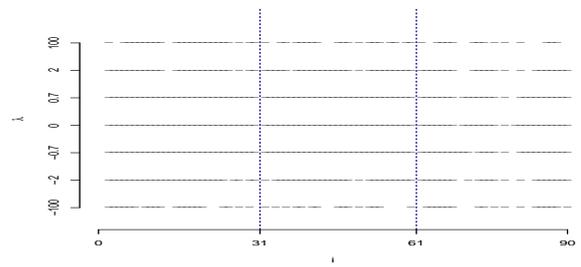


(h)  $P(\rho = \rho')$

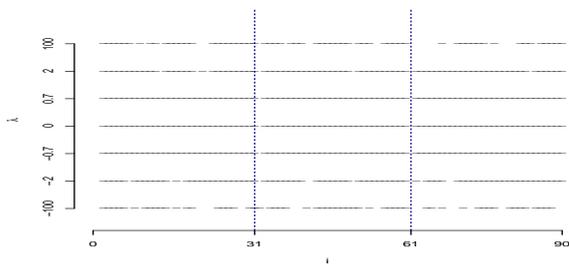
Figura D.21: Partição mais provável *a posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com \*),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com  $\bullet$ ) e  $\lambda = 0$  (linha cheia),  $\lambda =$  (linha tracejada curta),  $\lambda =$  (linha pontilhada),  $\lambda =$  (linha tracejada e pontilhada), cenário escada decrescente,  $n = 30$ .



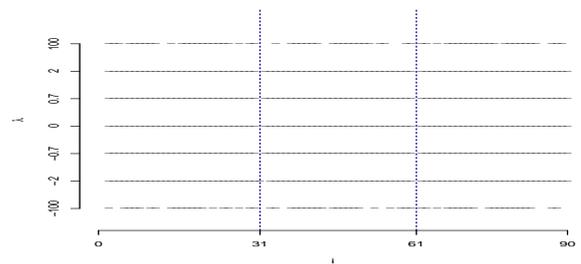
(a) Partição mais provável - Dados  $\lambda = -100$



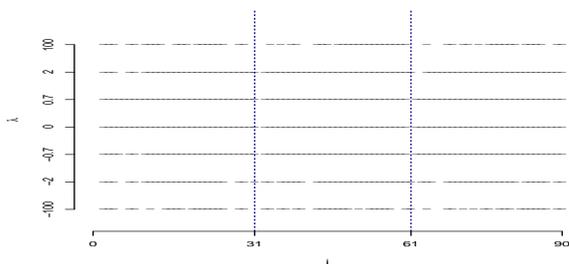
(b) Partição mais provável - Dados  $\lambda = -2$



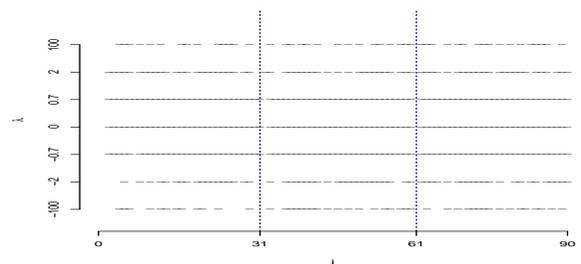
(c) Partição mais provável - Dados  $\lambda = -0.7$



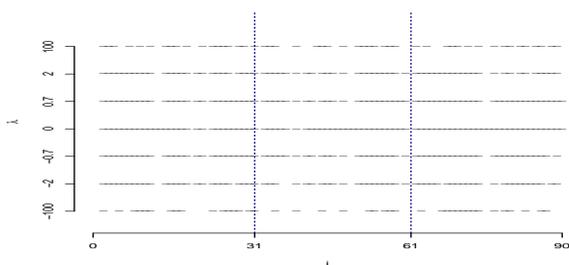
(d) Partição mais provável - Dados  $\lambda = 0$



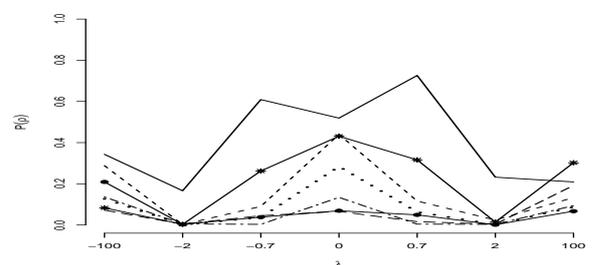
(e) Partição mais provável - Dados  $\lambda = 0.7$



(f) Partição mais provável - Dados  $\lambda = 2$



(g) Partição mais provável - Dados  $\lambda = 100$



(h)  $P(\rho = \rho')$

Figura D.22: Partição mais provável a *posteriori* e sua probabilidade(h) para cada  $\lambda$ . Estimativas para  $\lambda = -100$  (linha com \*),  $\lambda = -2$  (linha tracejada longa),  $\lambda = -0.7$  (linha com ●) e  $\lambda = 0$  (linha cheia),  $\lambda = 0.7$  (linha tracejada curta),  $\lambda = 2$  (linha pontilhada),  $\lambda = 100$  (linha tracejada e pontilhada), cenário escada decrescente,  $n = 90$ .

### D.1.3 A influência de $\lambda$ e $n$ nas inferências sobre $P(A_l)$

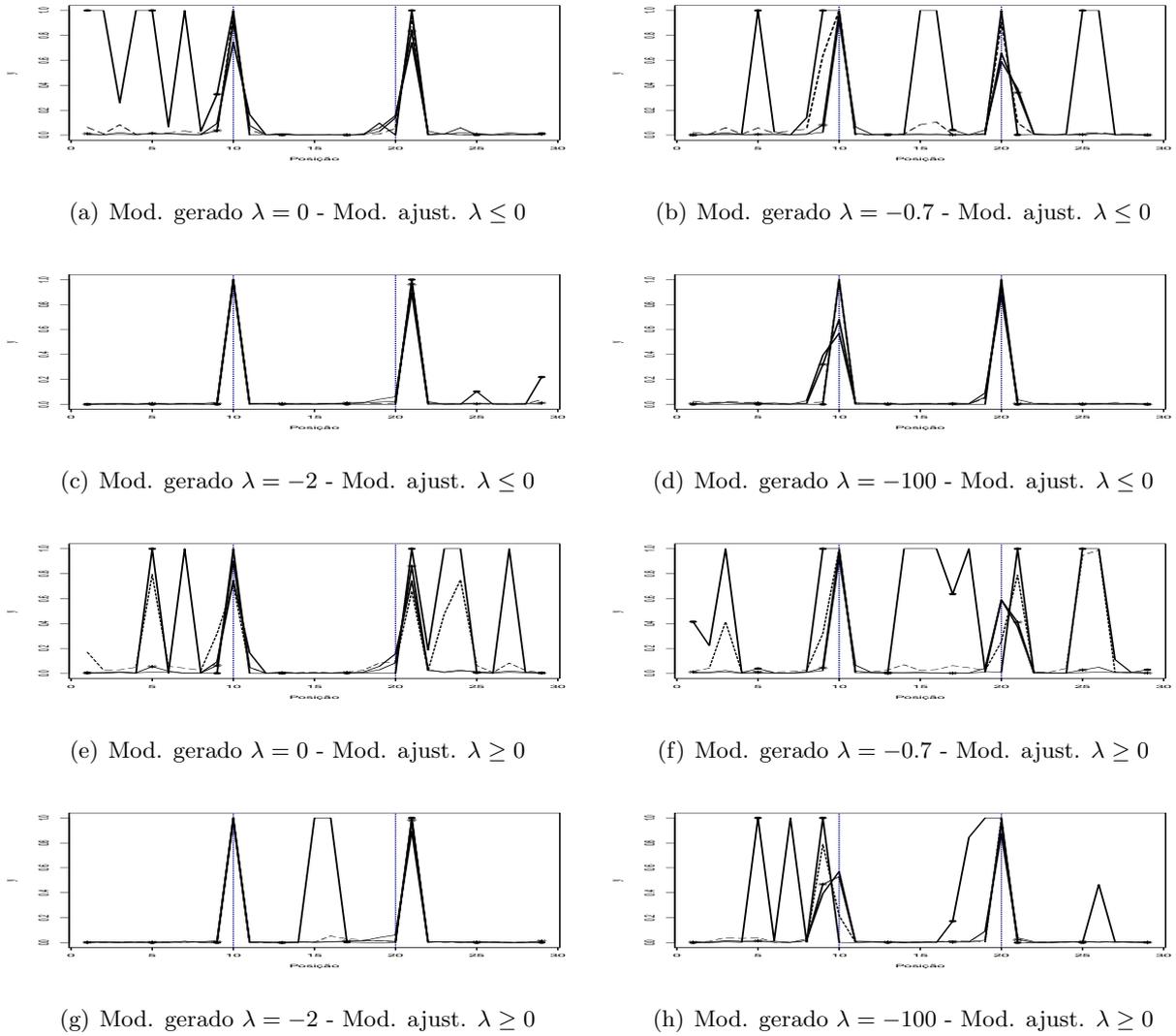
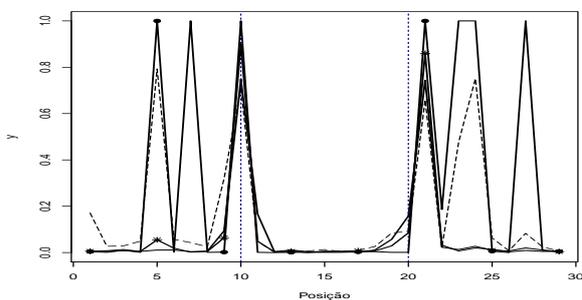
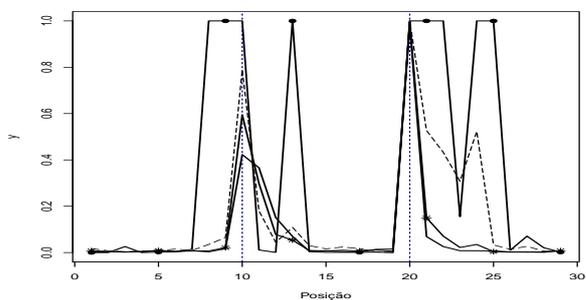


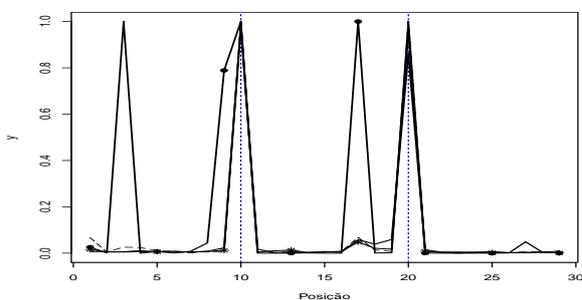
Figura D.23: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 30$ .



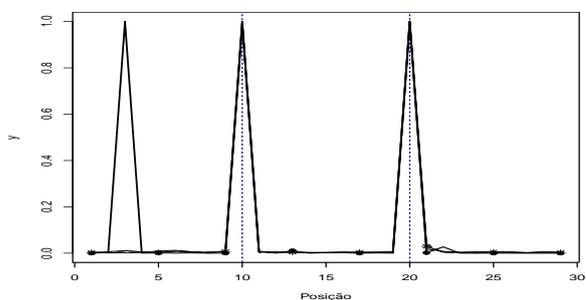
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



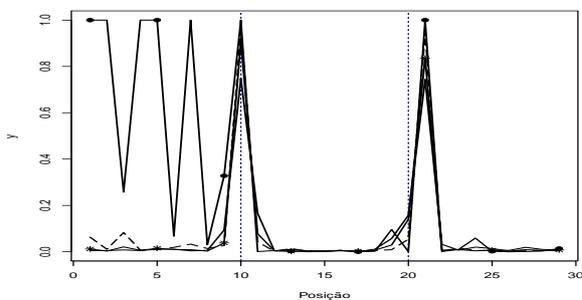
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



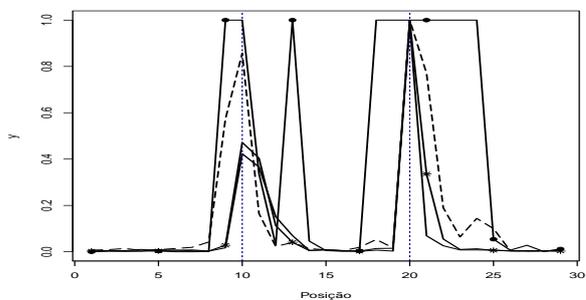
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$



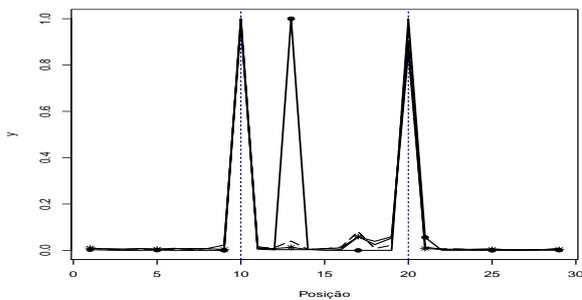
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$



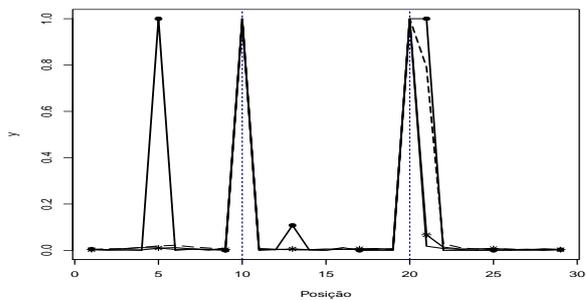
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$

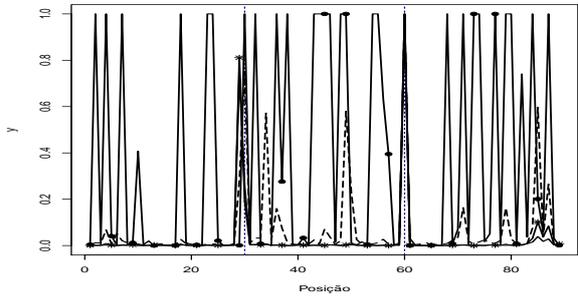


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$

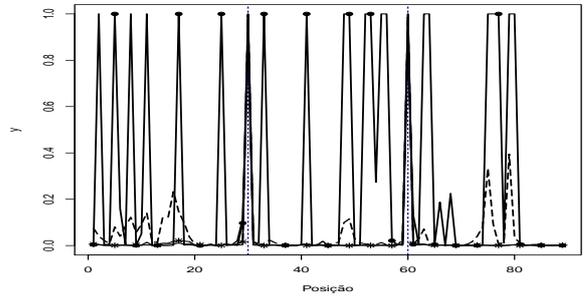


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

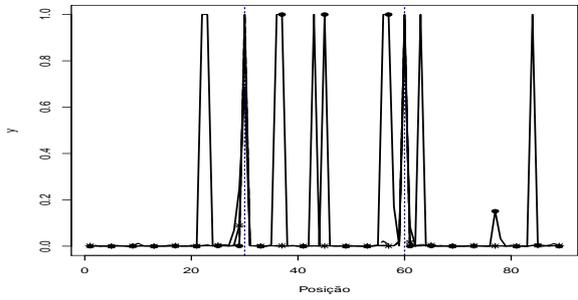
Figura D.24: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 30$ .



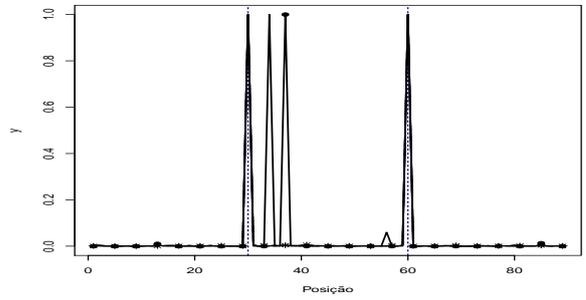
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



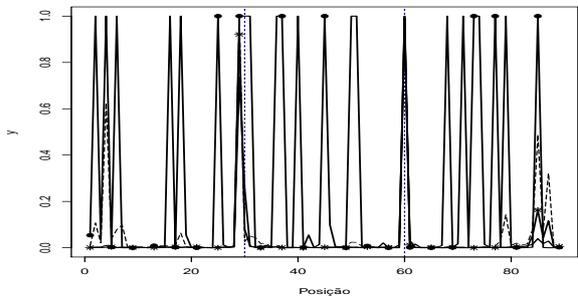
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



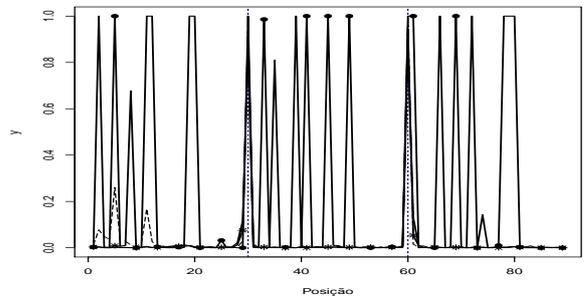
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



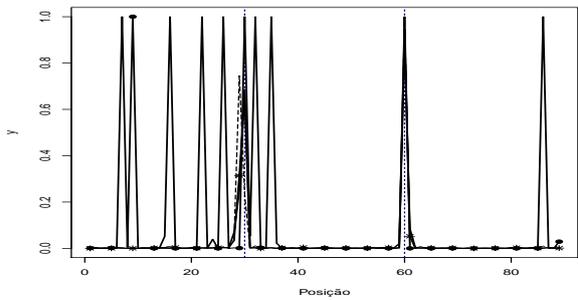
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



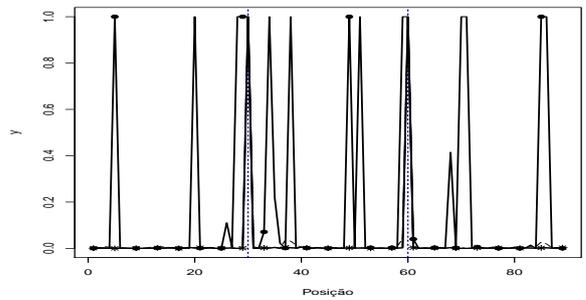
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$

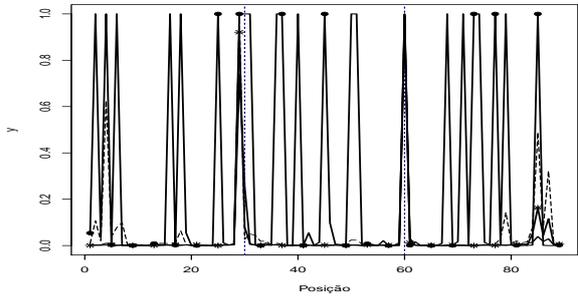


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

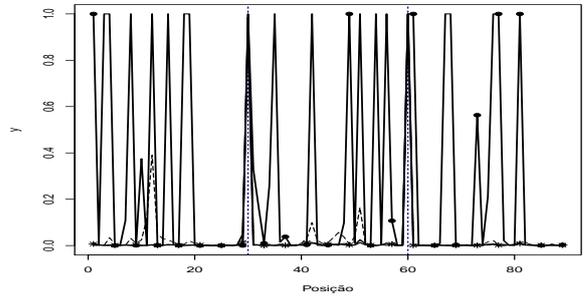


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

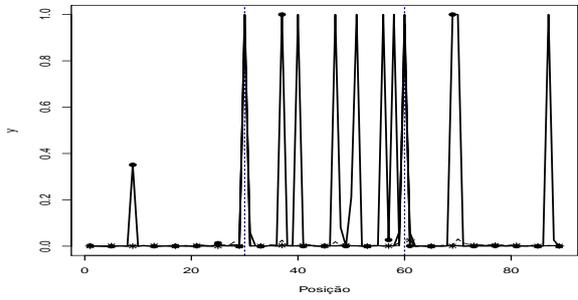
Figura D.25: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 90$ .



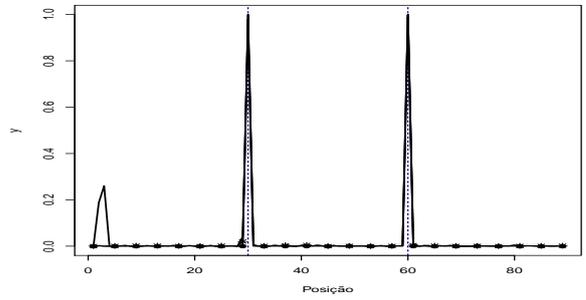
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



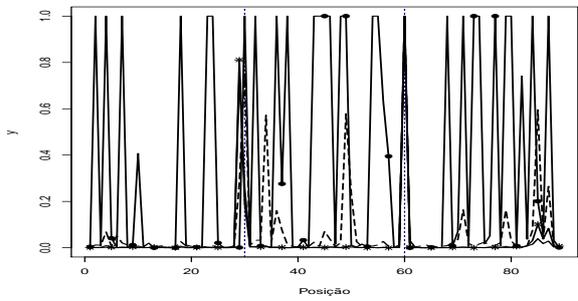
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



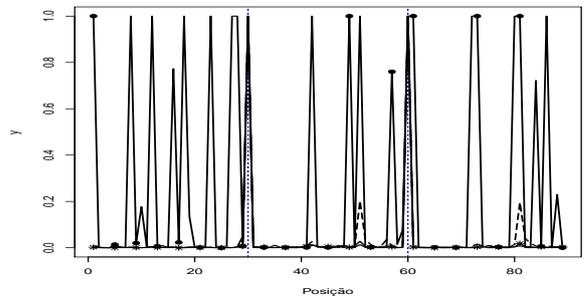
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$



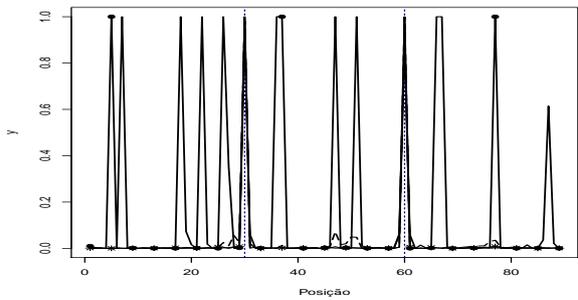
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$



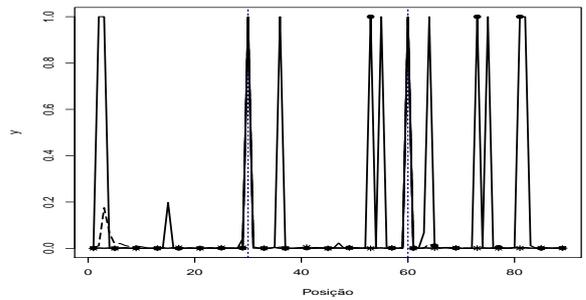
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$

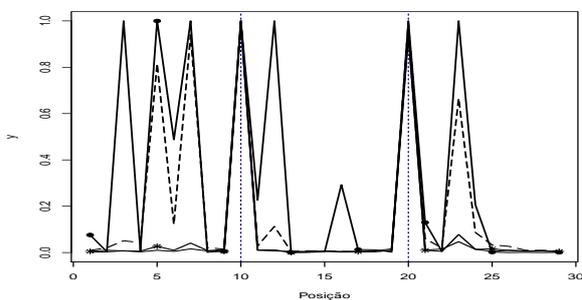


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$

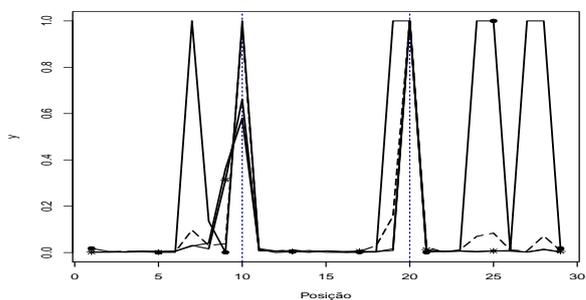


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

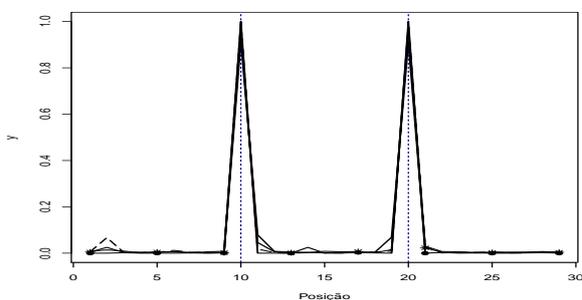
Figura D.26: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário degrau negativo,  $n = 90$ .



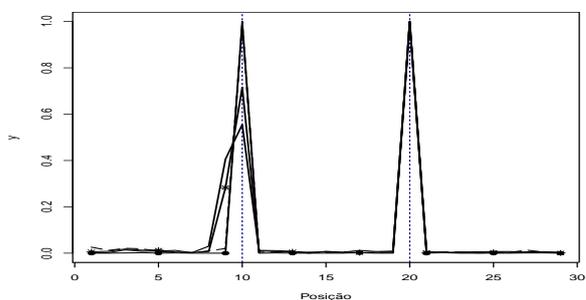
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



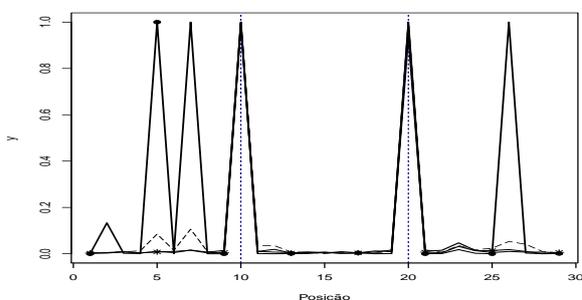
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



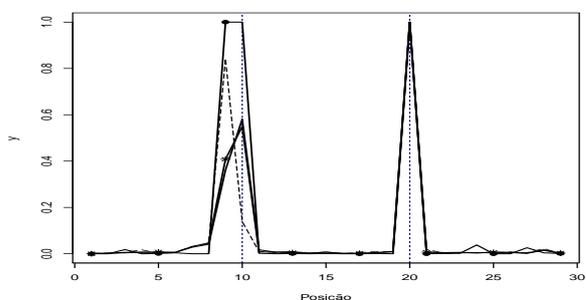
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



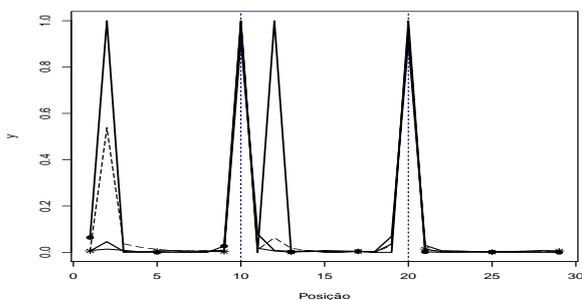
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



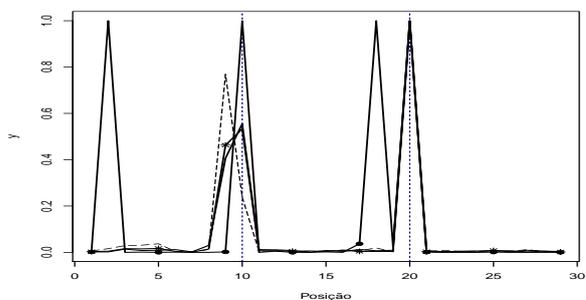
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$

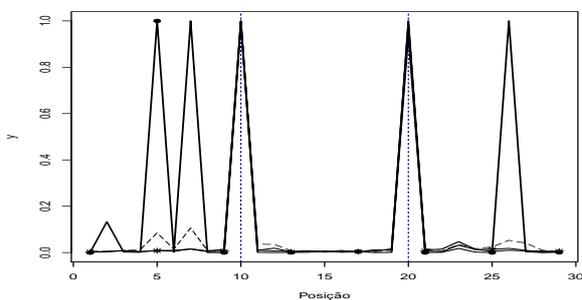


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

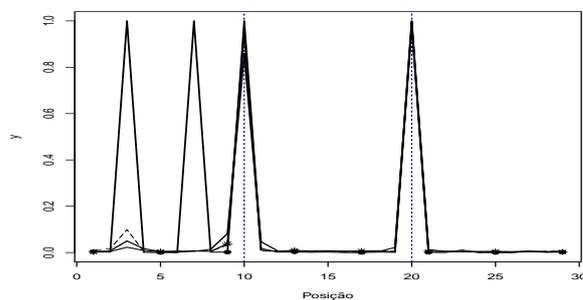


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

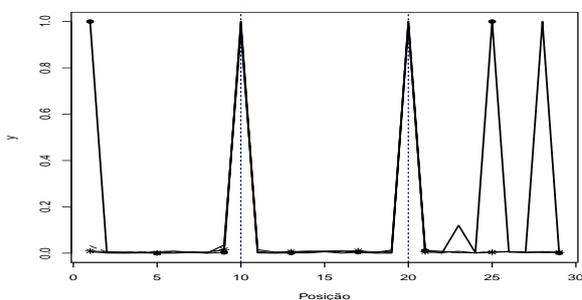
Figura D.27: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 30$ .



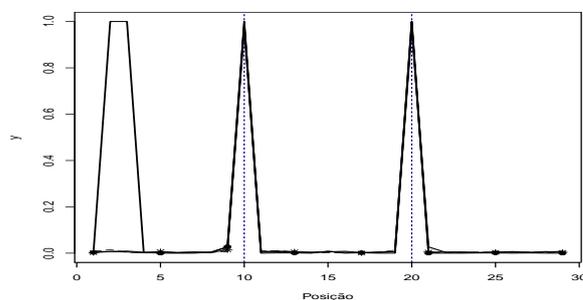
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



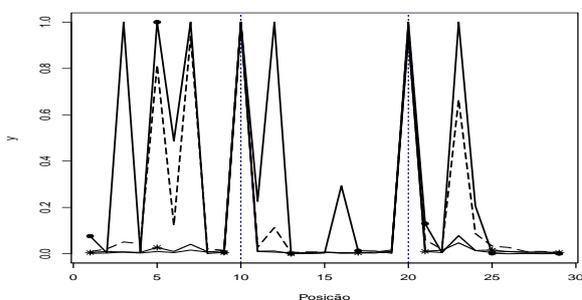
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



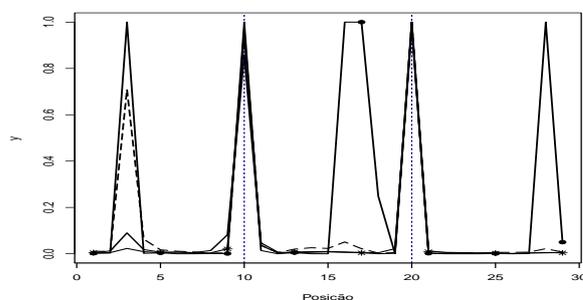
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$



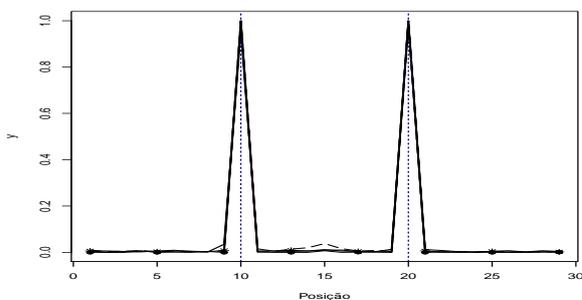
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$



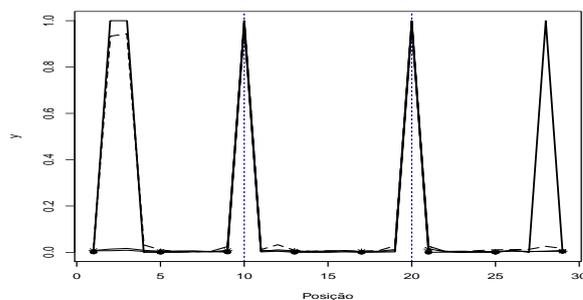
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$

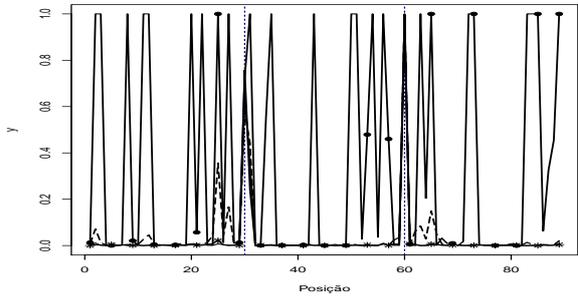


(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$

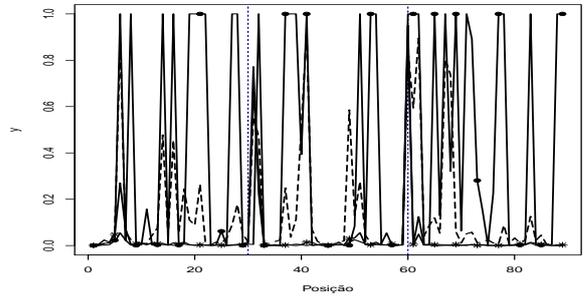


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

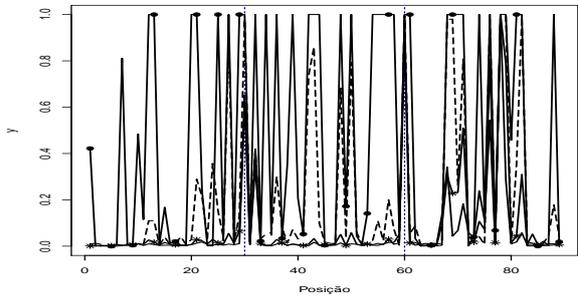
Figura D.28: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 30$ .



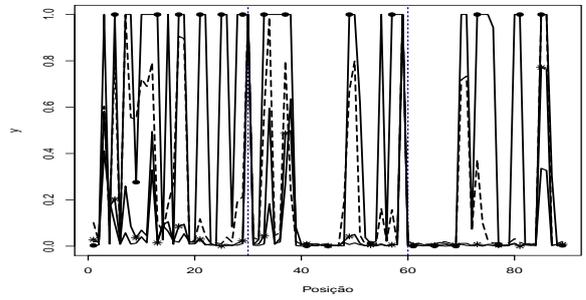
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



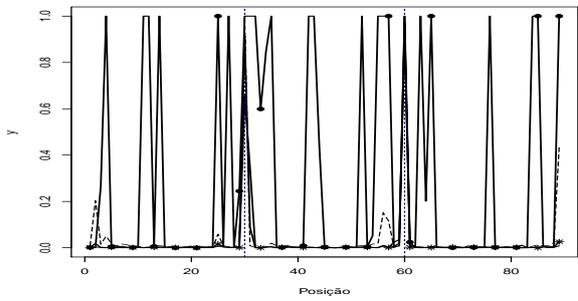
(b) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \leq 0$



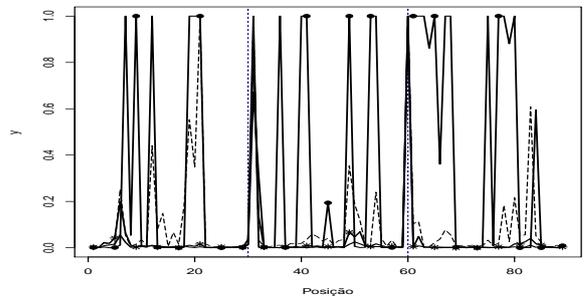
(c) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \leq 0$



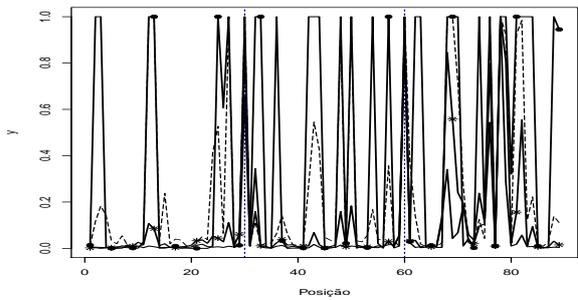
(d) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \leq 0$



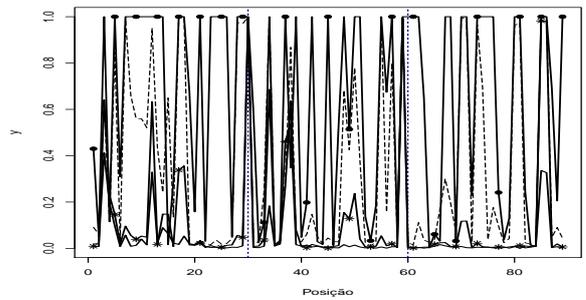
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



(f) Mod. gerado  $\lambda = -0.7$  - Mod. ajust.  $\lambda \geq 0$

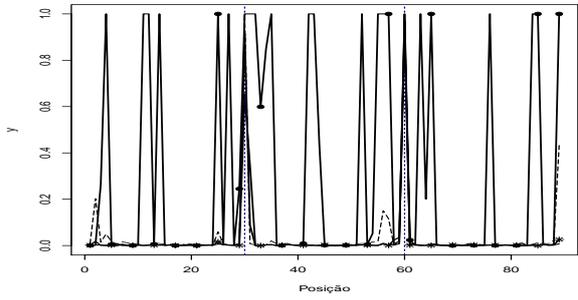


(g) Mod. gerado  $\lambda = -2$  - Mod. ajust.  $\lambda \geq 0$

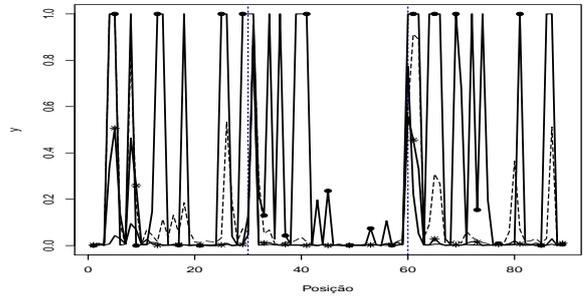


(h) Mod. gerado  $\lambda = -100$  - Mod. ajust.  $\lambda \geq 0$

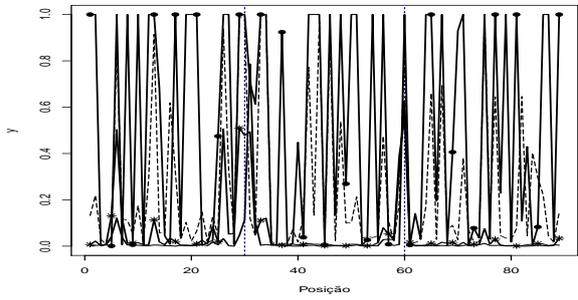
Figura D.29: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 90$ .



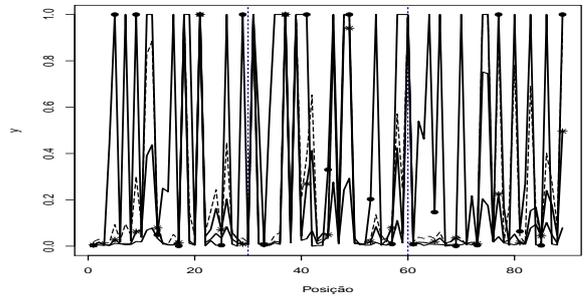
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



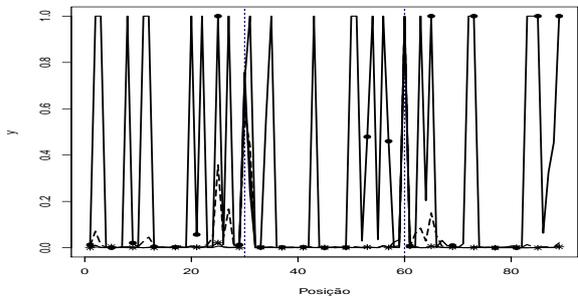
(b) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



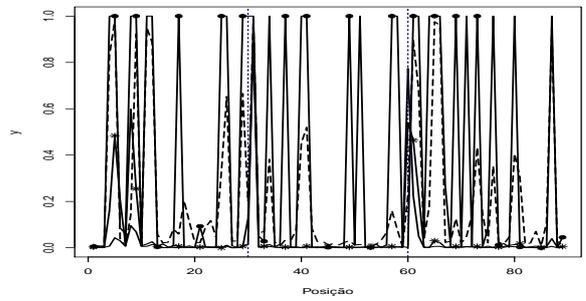
(c) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$



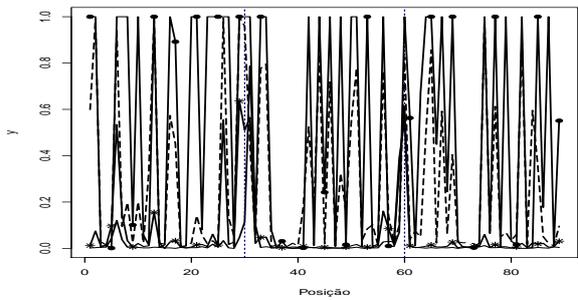
(d) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$



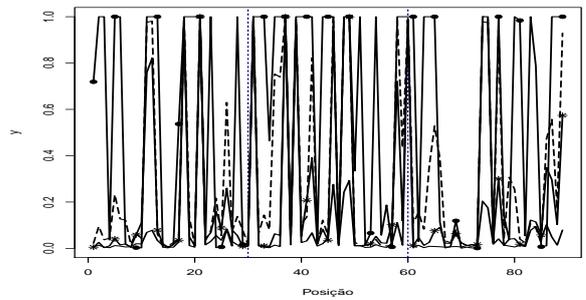
(e) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



(g) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$

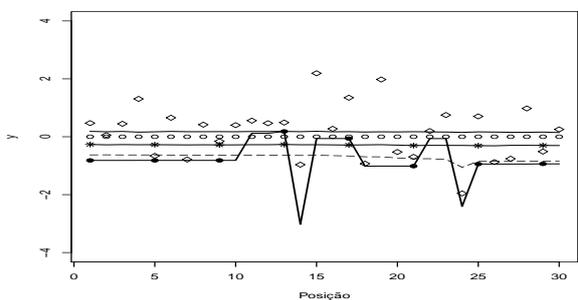


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

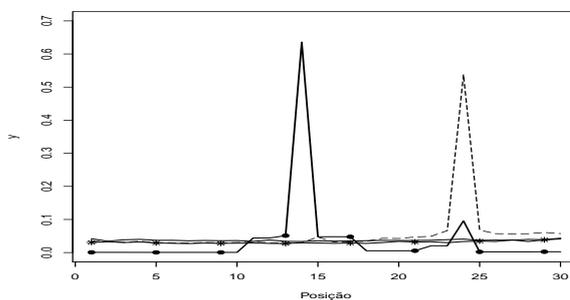
Figura D.30: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com \*),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com •) e  $\lambda = 0$  (linha cheia), cenário escada decrescente,  $n = 90$ .

#### **D.1.4 Gráficos do estudo de sensibilidade informal - Modelo II: robustez na média e no instante da mudança**

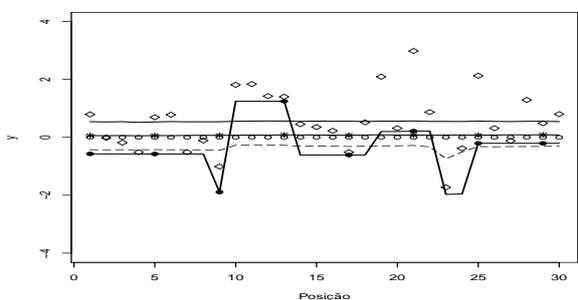
Os gráficos dessa seção referem-se a todos os cenários avaliados no estudo de sensibilidade nos dados, no entanto, foi observado um padrão no comportamento da estimativas, ou seja, se os dados são simulados com um parâmetro de assimetria negativo e o modelo é ajustado com assimetria positiva, então as estimativas tendem a ser subestimadas. Se o contrário ocorre, os dados são simulados com um parâmetro de assimetria positivo e o modelo é ajustado com assimetria negativa, então as estimativas serão superestimadas.



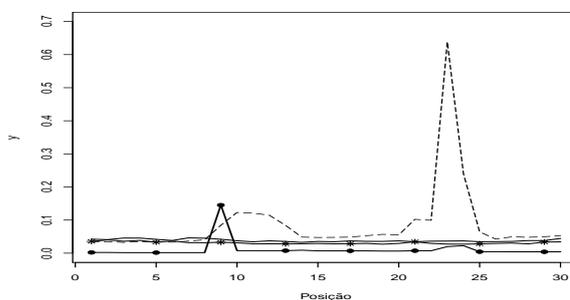
(a) Mod. gerado  $\lambda = 0$



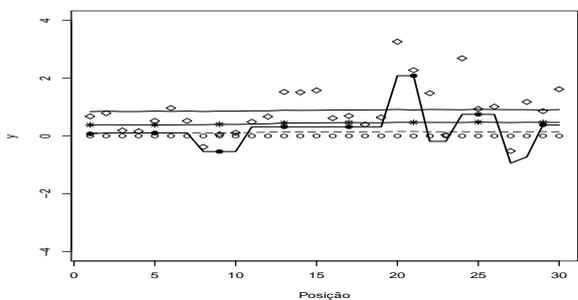
(b) Mod. gerado  $\lambda = 0$



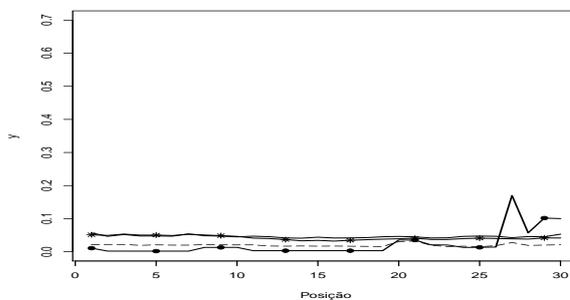
(c) Mod. gerado  $\lambda = 0.7$



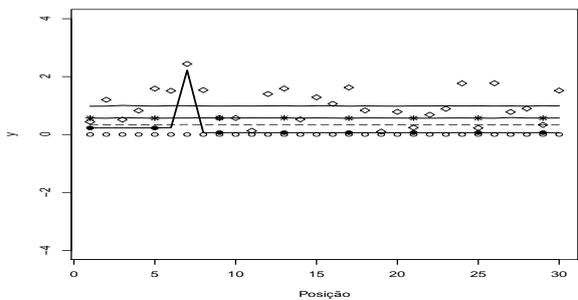
(d) Mod. gerado  $\lambda = 0.7$



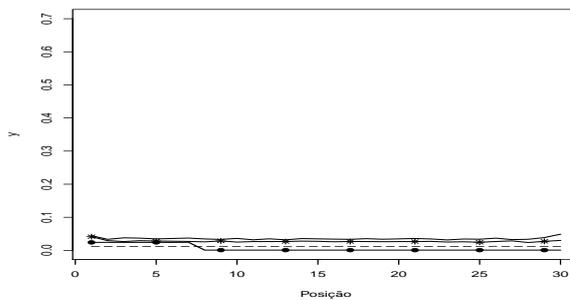
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

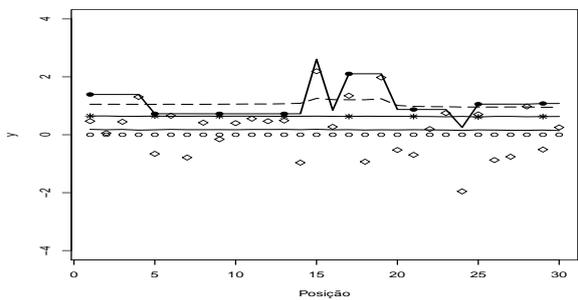


(g) Mod. gerado  $\lambda = 100$

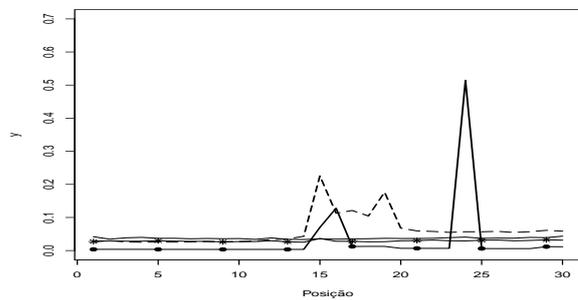


(h) Mod. gerado  $\lambda = 100$

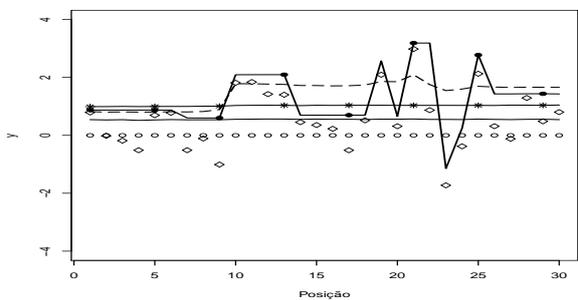
Figura D.31: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 30$ .



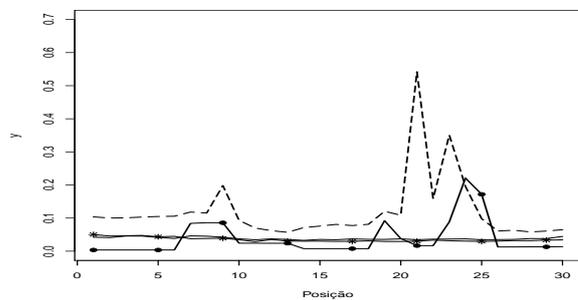
(a) Mod. gerado  $\lambda = 0$



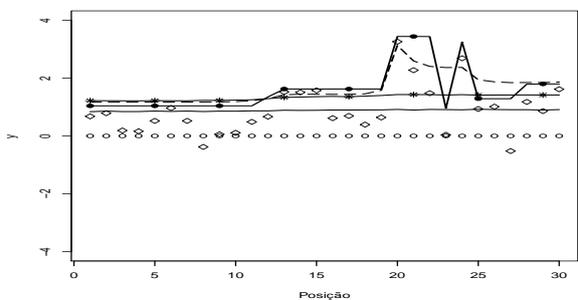
(b) Mod. gerado  $\lambda = 0$



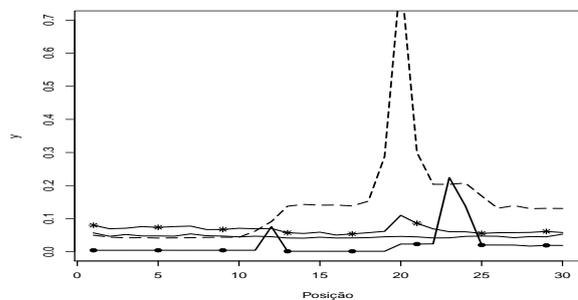
(c) Mod. gerado  $\lambda = 0.7$



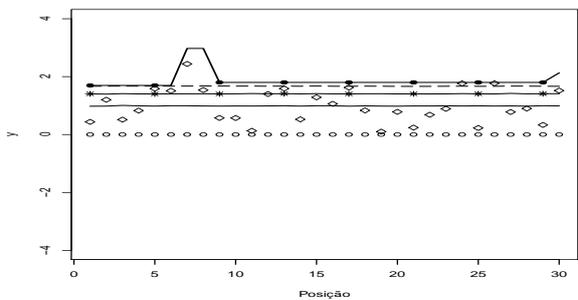
(d) Mod. gerado  $\lambda = 0.7$



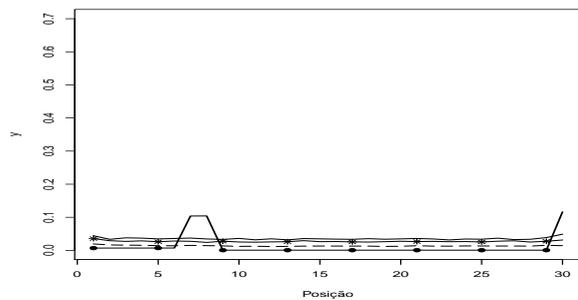
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

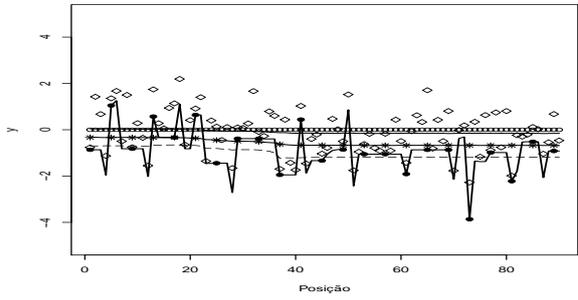


(g) Mod. gerado  $\lambda = 100$

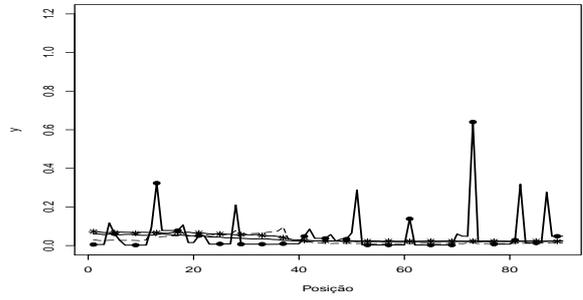


(h) Mod. gerado  $\lambda = 100$

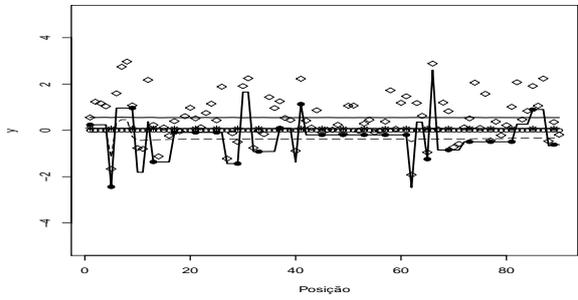
Figura D.32: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 30$ .



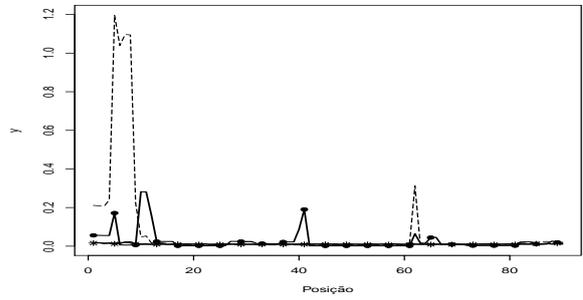
(a) Mod. gerado  $\lambda = 0$



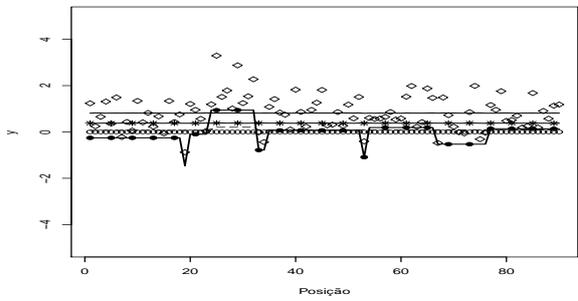
(b) Mod. gerado  $\lambda = 0$



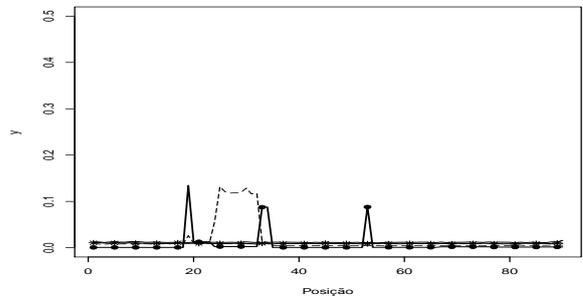
(c) Mod. gerado  $\lambda = 0.7$



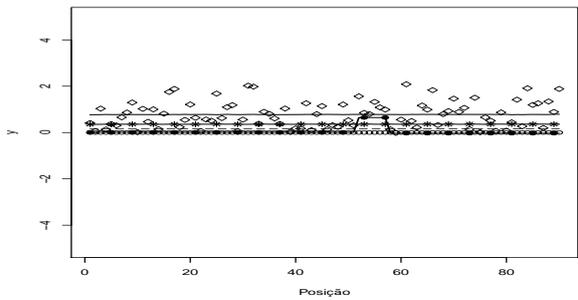
(d) Mod. gerado  $\lambda = 0.7$



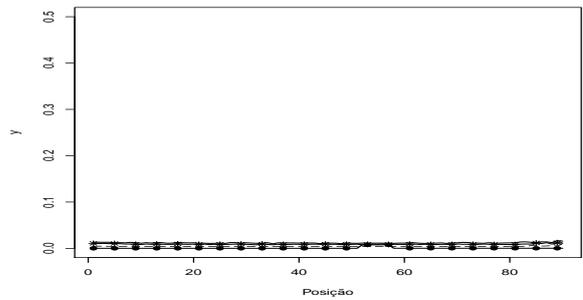
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

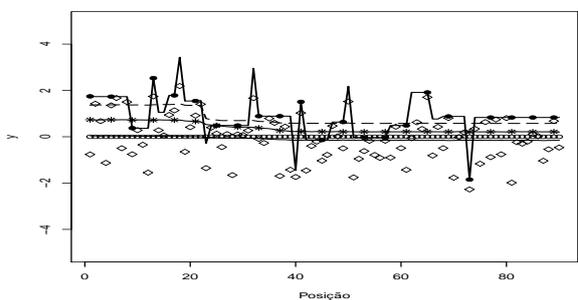


(g) Mod. gerado  $\lambda = 100$

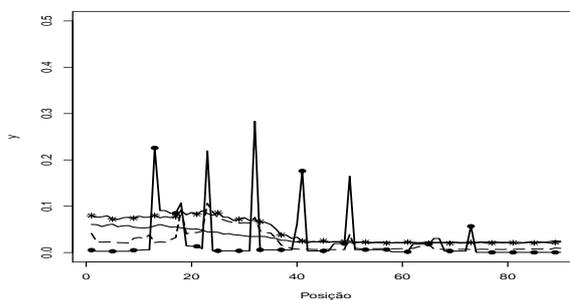


(h) Mod. gerado  $\lambda = 100$

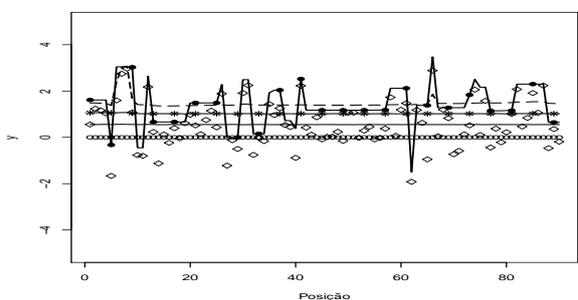
Figura D.33: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 90$ .



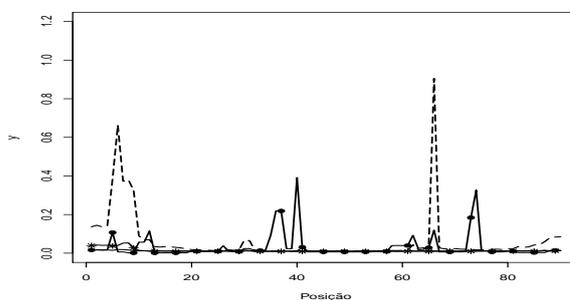
(a) Mod. gerado  $\lambda = 0$



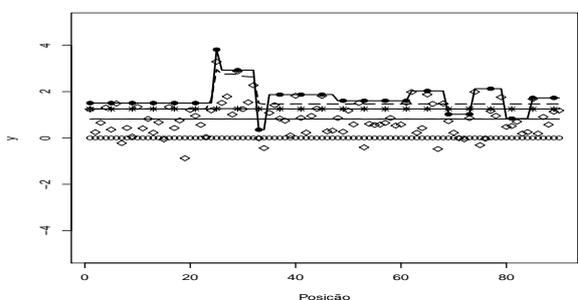
(b) Mod. gerado  $\lambda = 0$



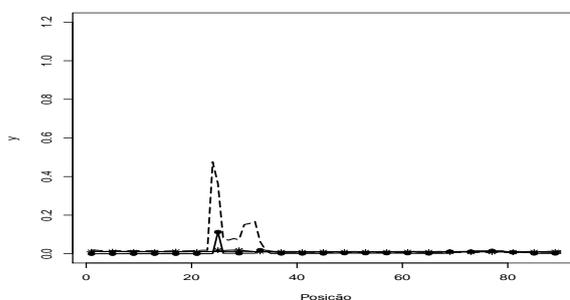
(c) Mod. gerado  $\lambda = 0.7$



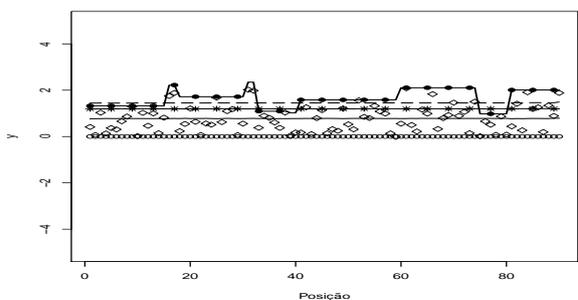
(d) Mod. gerado  $\lambda = 0.7$



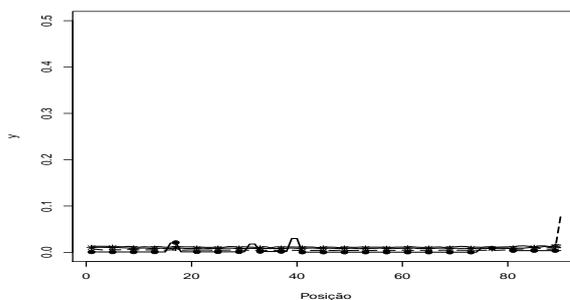
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

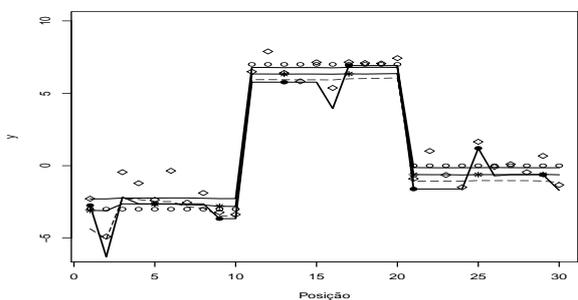


(g) Mod. gerado  $\lambda = 100$

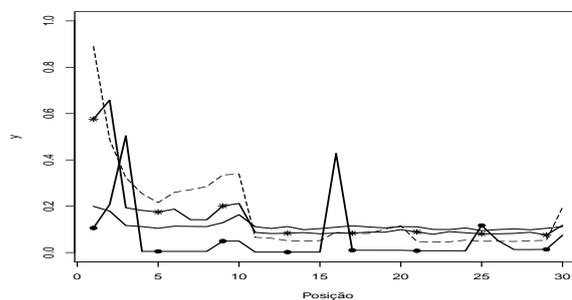


(h) Mod. gerado  $\lambda = 100$

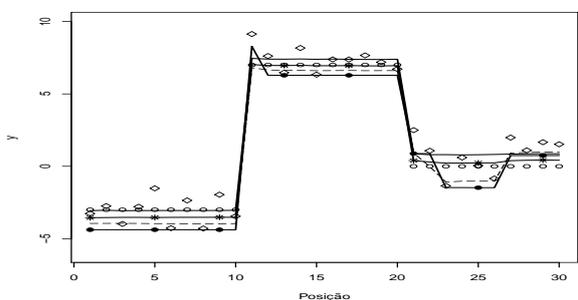
Figura D.34: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 90$ .



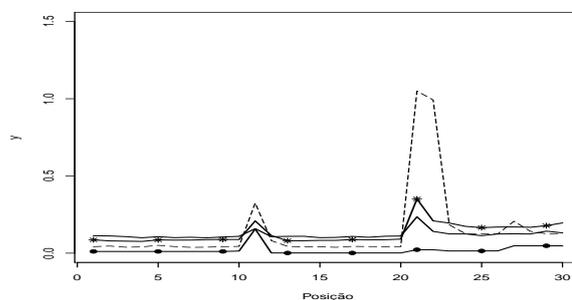
(a) Mod. gerado  $\lambda = 0$



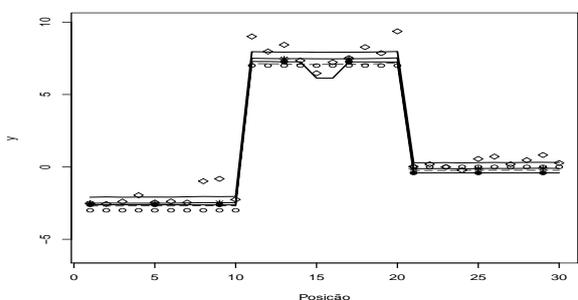
(b) Mod. gerado  $\lambda = 0$



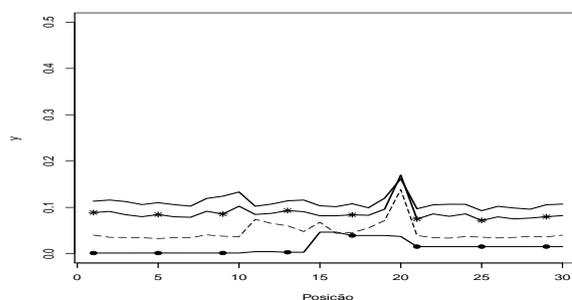
(c) Mod. gerado  $\lambda = 0.7$



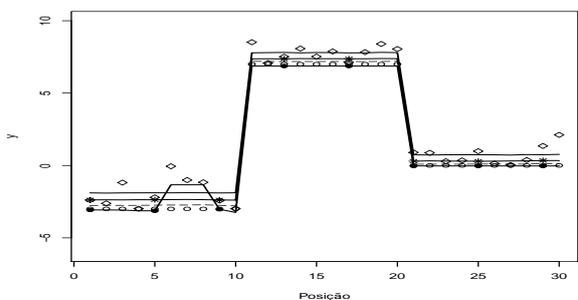
(d) Mod. gerado  $\lambda = 0.7$



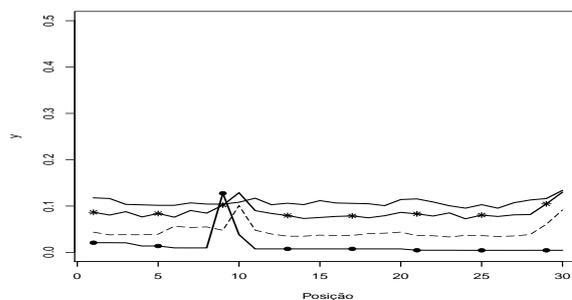
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

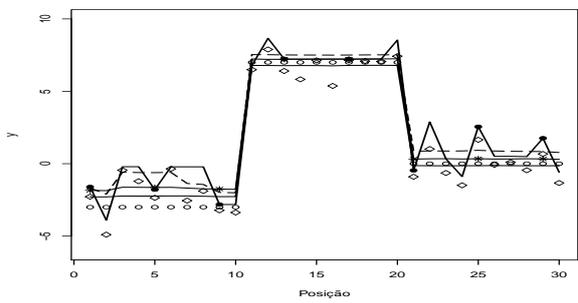


(g) Mod. gerado  $\lambda = 100$

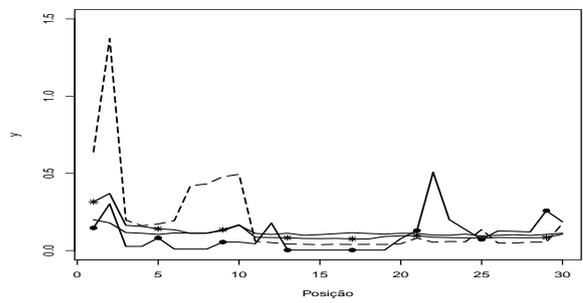


(h) Mod. gerado  $\lambda = 100$

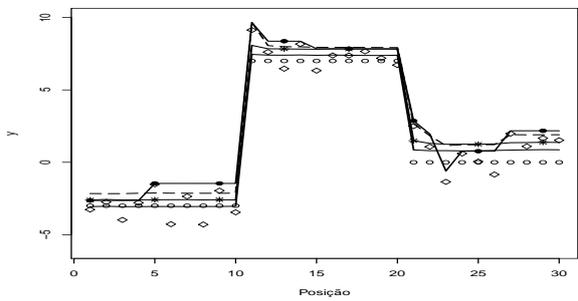
Figura D.35: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 30$ .



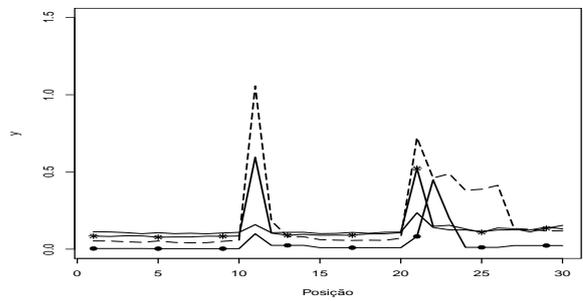
(a) Mod. gerado  $\lambda = 0$



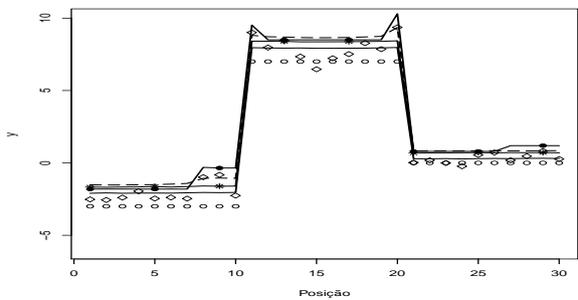
(b) Mod. gerado  $\lambda = 0$



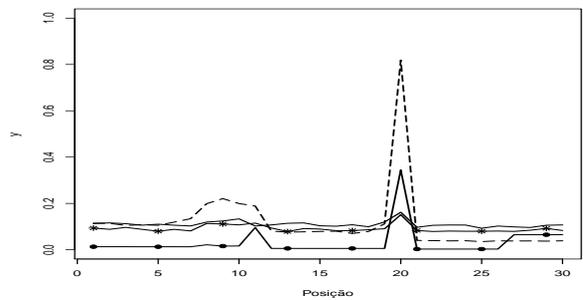
(c) Mod. gerado  $\lambda = 0.7$



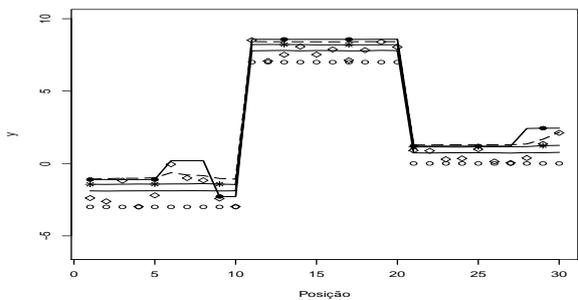
(d) Mod. gerado  $\lambda = 0.7$



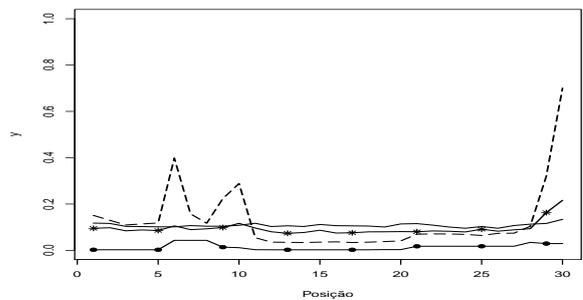
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

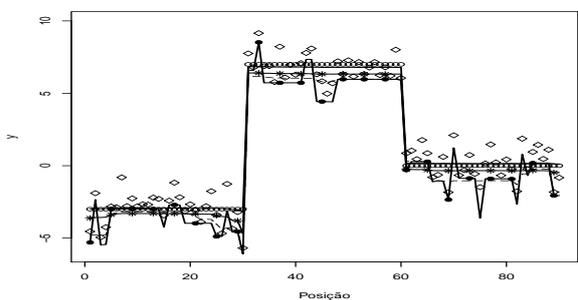


(g) Mod. gerado  $\lambda = 100$

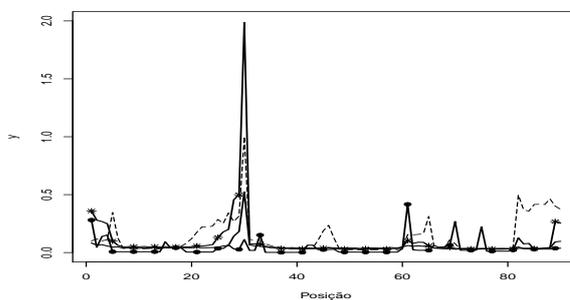


(h) Mod. gerado  $\lambda = 100$

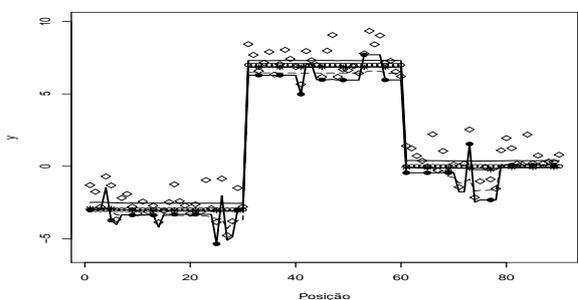
Figura D.36: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 30$ .



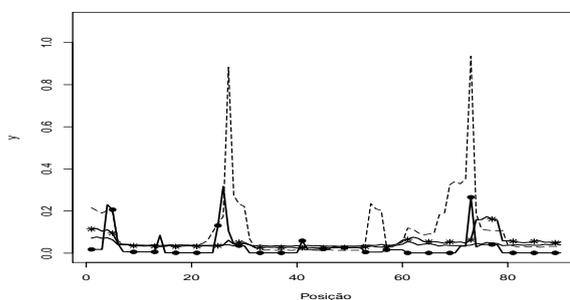
(a) Mod. gerado  $\lambda = 0$



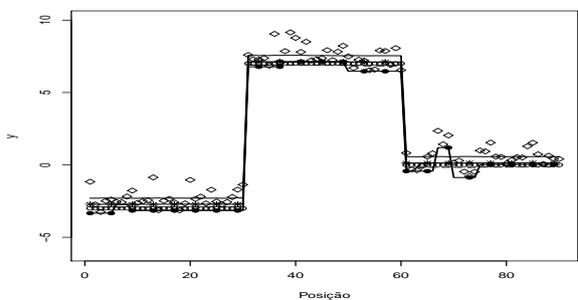
(b) Mod. gerado  $\lambda = 0$



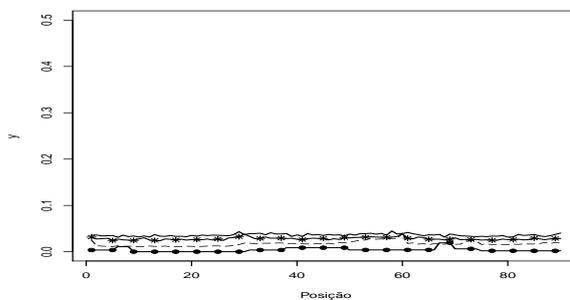
(c) Mod. gerado  $\lambda = 0.7$



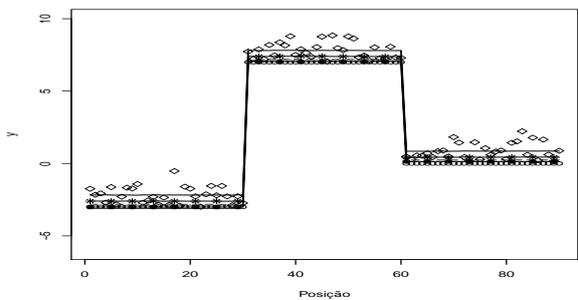
(d) Mod. gerado  $\lambda = 0.7$



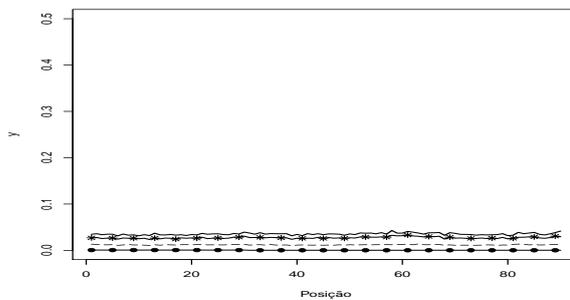
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

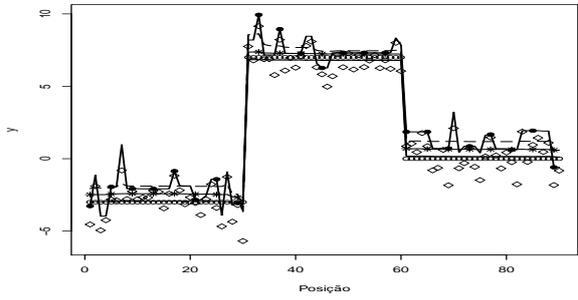


(g) Mod. gerado  $\lambda = 100$

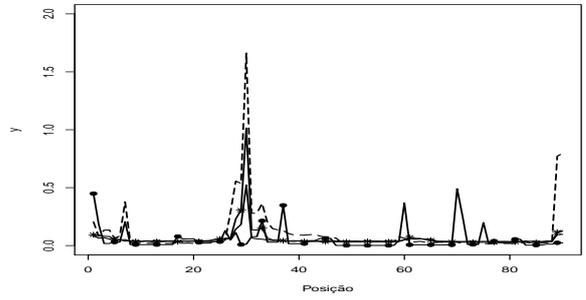


(h) Mod. gerado  $\lambda = 100$

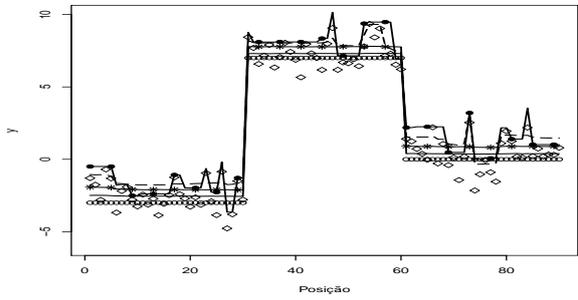
Figura D.37: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 90$ .



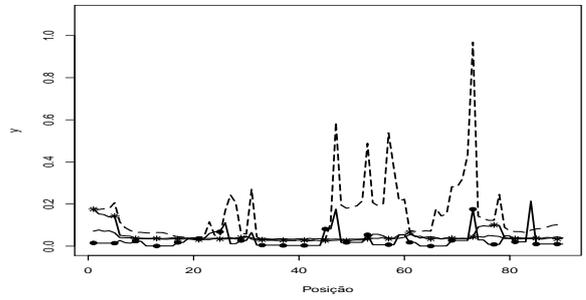
(a) Mod. gerado  $\lambda = 0$



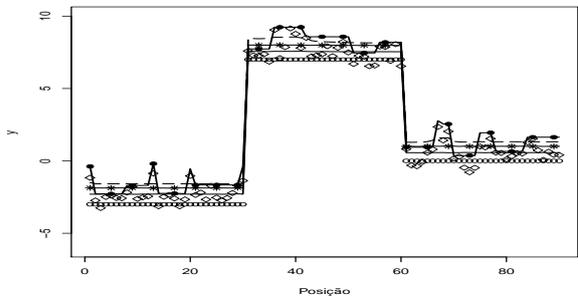
(b) Mod. gerado  $\lambda = 0$



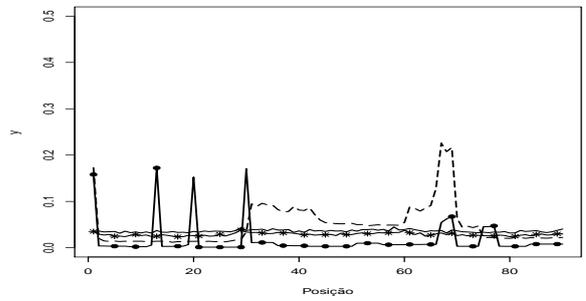
(c) Mod. gerado  $\lambda = 0.7$



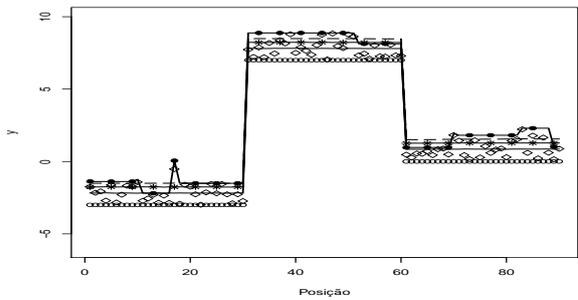
(d) Mod. gerado  $\lambda = 0.7$



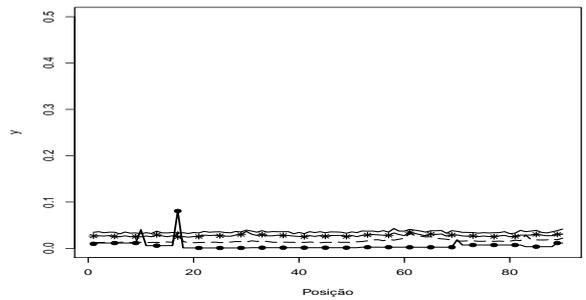
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

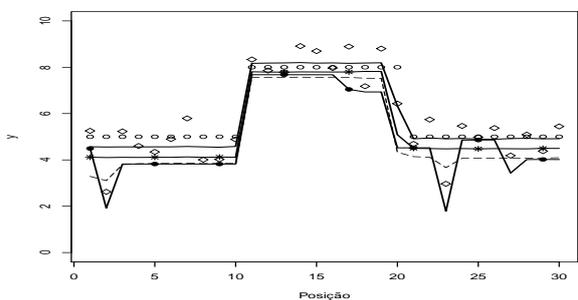


(g) Mod. gerado  $\lambda = 100$

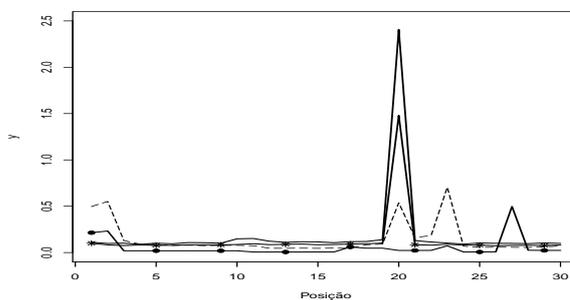


(h) Mod. gerado  $\lambda = 100$

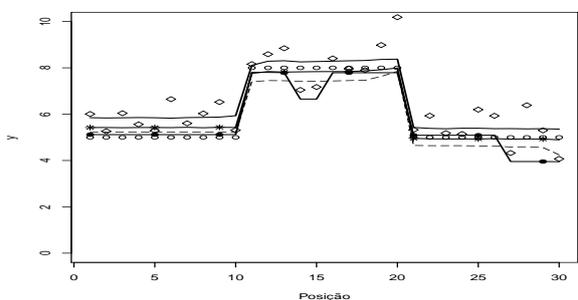
Figura D.38: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 90$ .



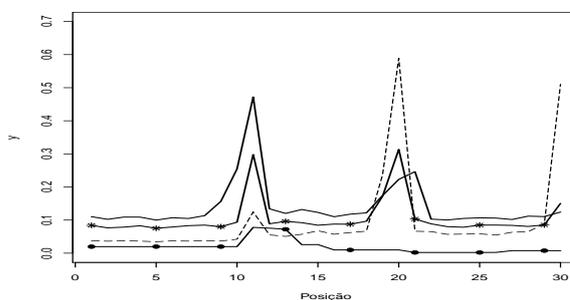
(a) Mod. gerado  $\lambda = 0$



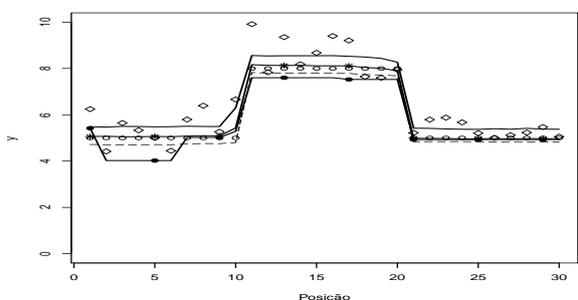
(b) Mod. gerado  $\lambda = 0$



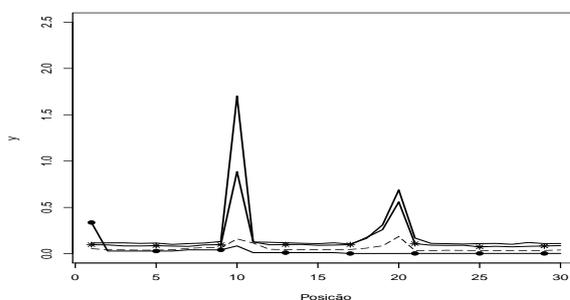
(c) Mod. gerado  $\lambda = 0.7$



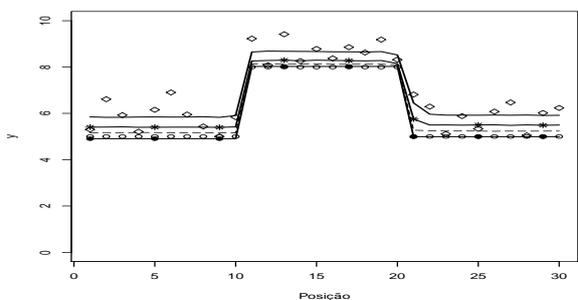
(d) Mod. gerado  $\lambda = 0.7$



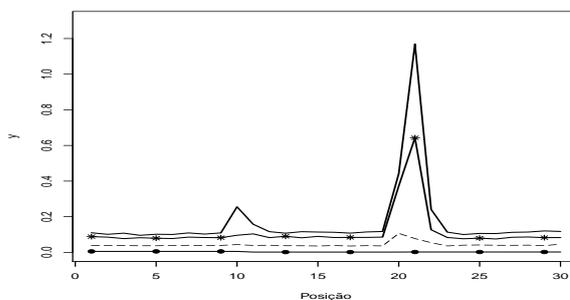
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

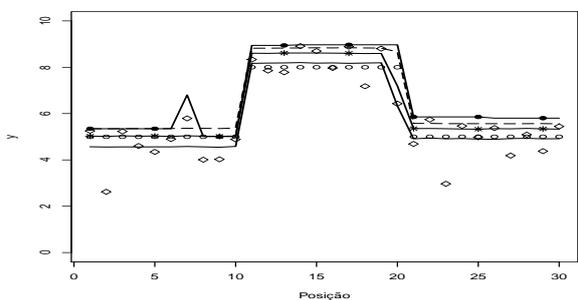


(g) Mod. gerado  $\lambda = 100$

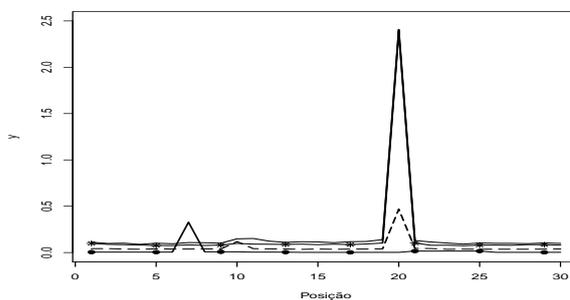


(h) Mod. gerado  $\lambda = 100$

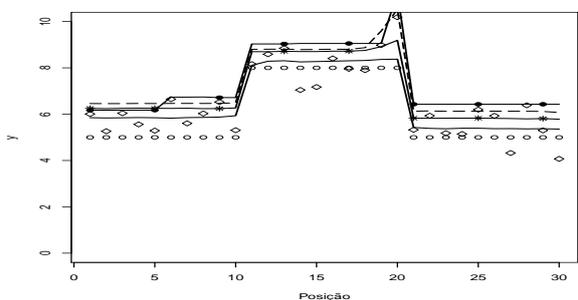
Figura D.39: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 30$ .



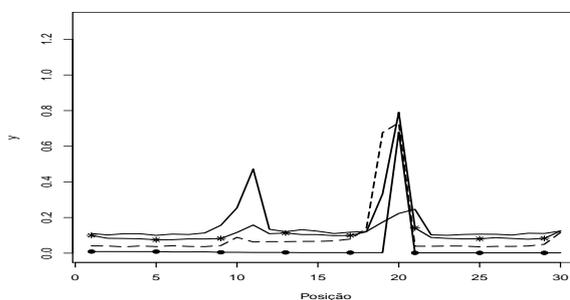
(a) Mod. gerado  $\lambda = 0$



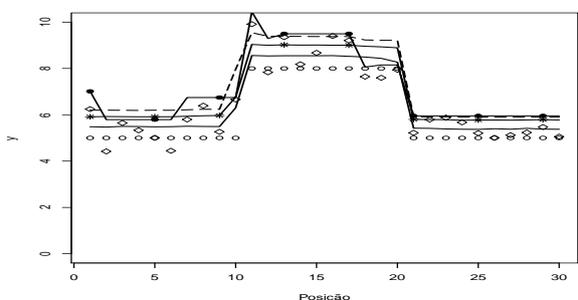
(b) Mod. gerado  $\lambda = 0$



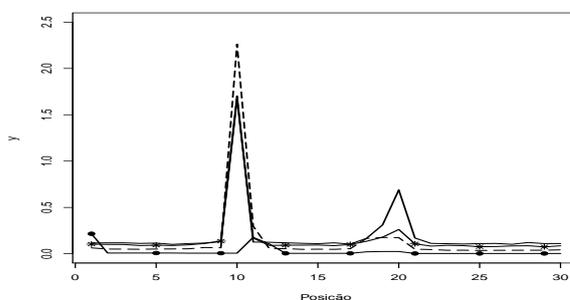
(c) Mod. gerado  $\lambda = 0.7$



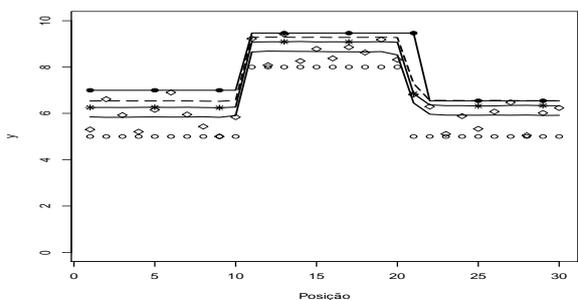
(d) Mod. gerado  $\lambda = 0.7$



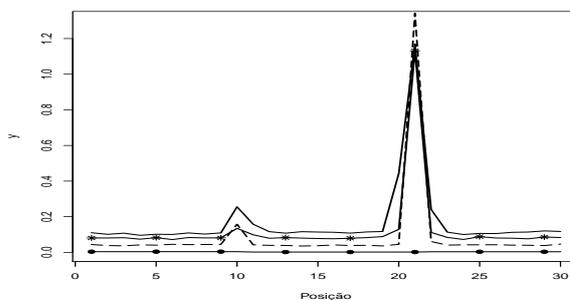
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

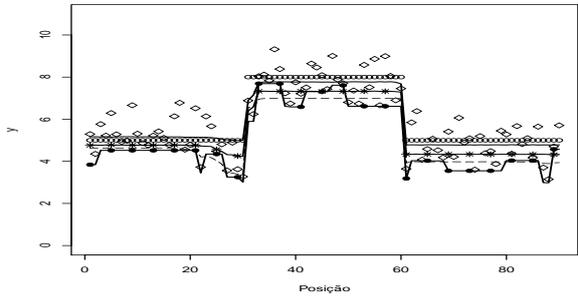


(g) Mod. gerado  $\lambda = 100$

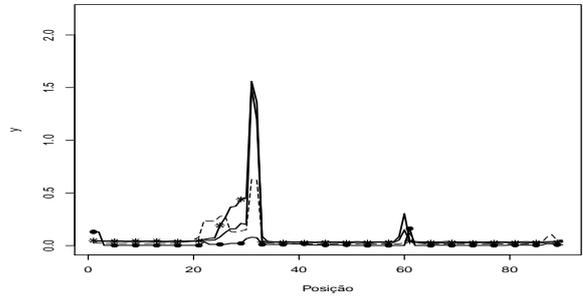


(h) Mod. gerado  $\lambda = 100$

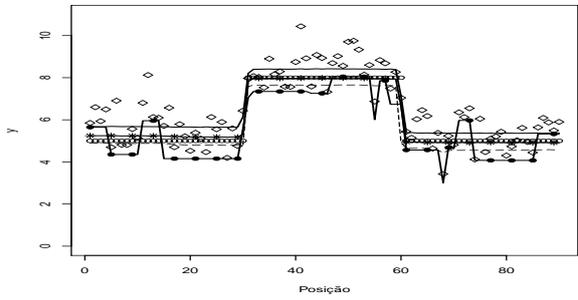
Figura D.40: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 30$ .



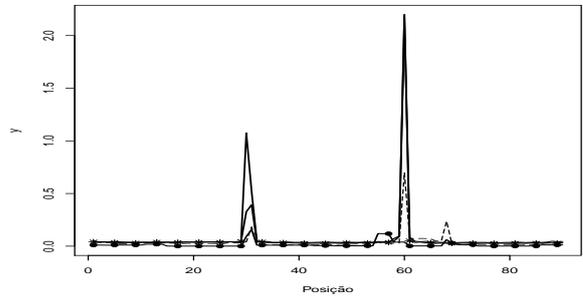
(a) Mod. gerado  $\lambda = 0$



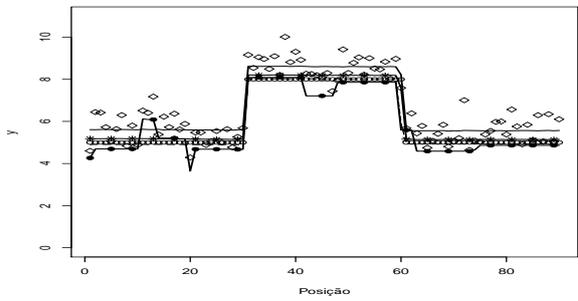
(b) Mod. gerado  $\lambda = 0$



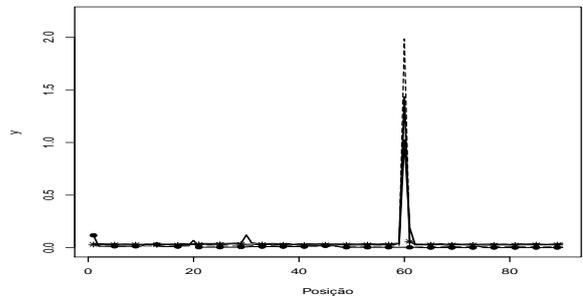
(c) Mod. gerado  $\lambda = 0.7$



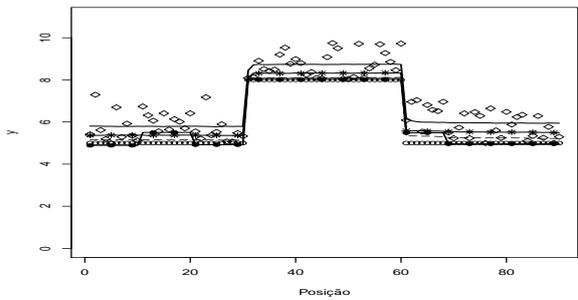
(d) Mod. gerado  $\lambda = 0.7$



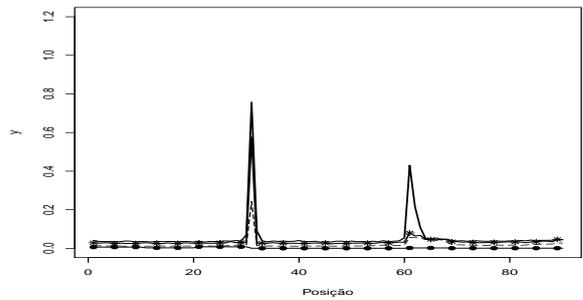
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

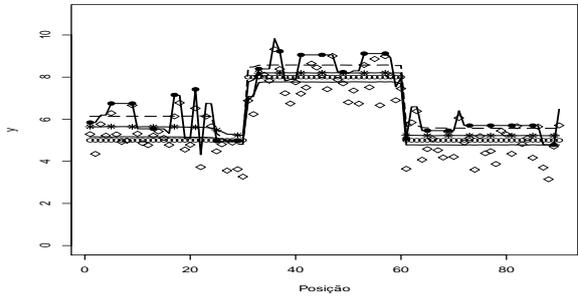


(g) Mod. gerado  $\lambda = 100$

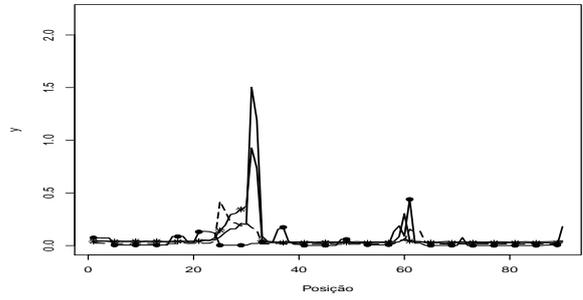


(h) Mod. gerado  $\lambda = 100$

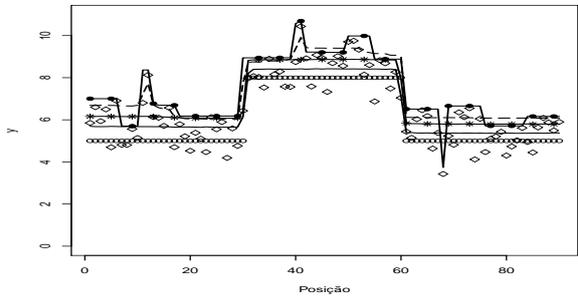
Figura D.41: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 90$ .



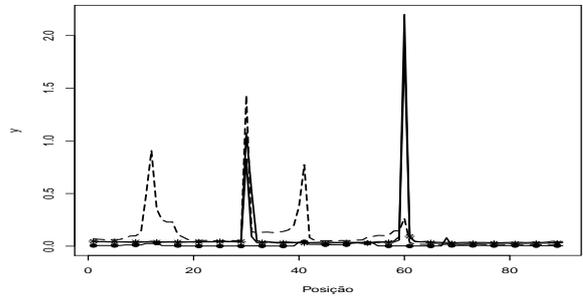
(a) Mod. gerado  $\lambda = 0$



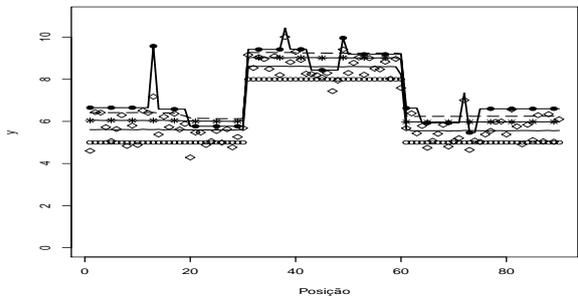
(b) Mod. gerado  $\lambda = 0$



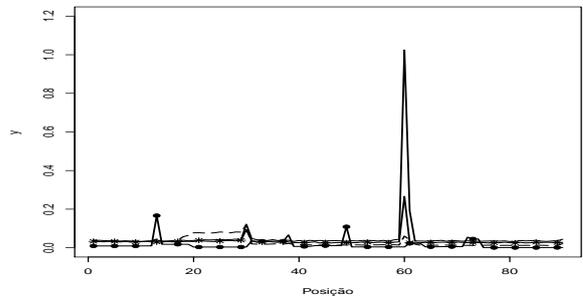
(c) Mod. gerado  $\lambda = 0.7$



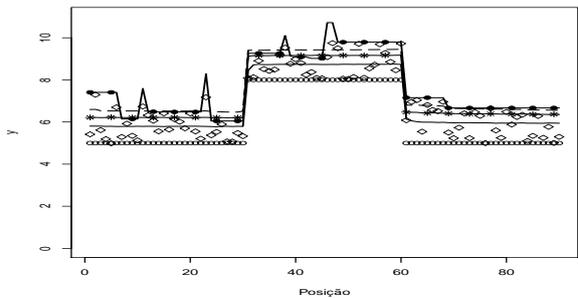
(d) Mod. gerado  $\lambda = 0.7$



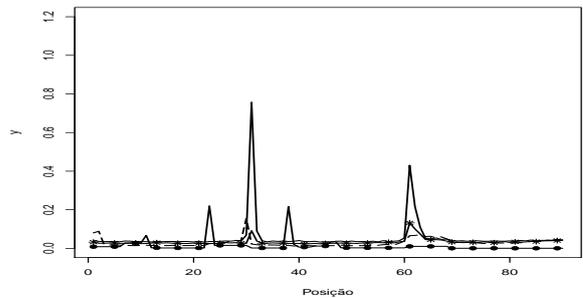
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

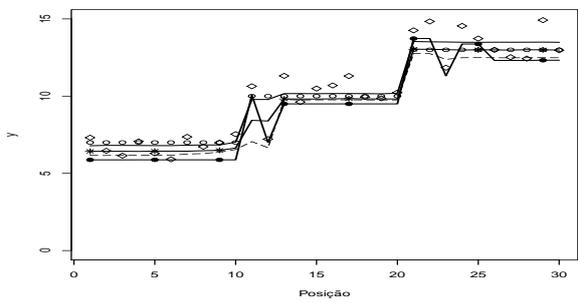


(g) Mod. gerado  $\lambda = 100$

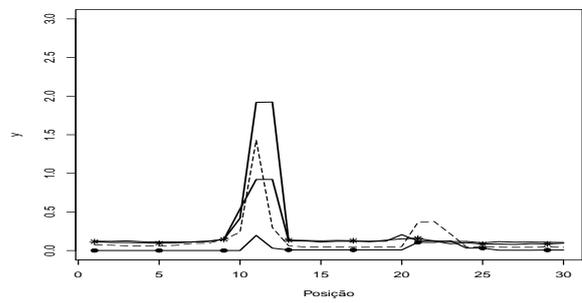


(h) Mod. gerado  $\lambda = 100$

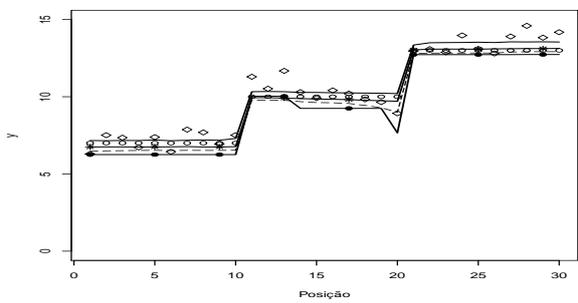
Figura D.42: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 90$ .



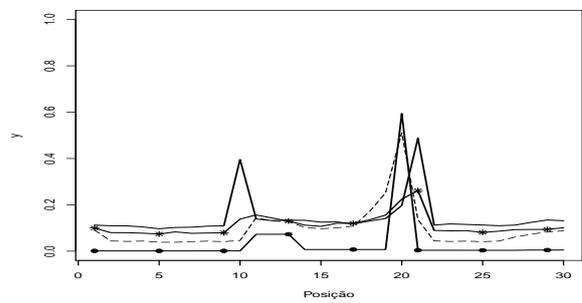
(a) Mod. gerado  $\lambda = 0$



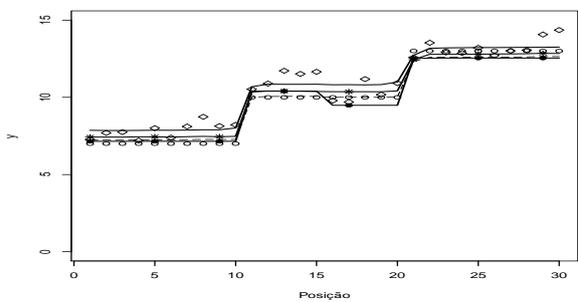
(b) Mod. gerado  $\lambda = 0$



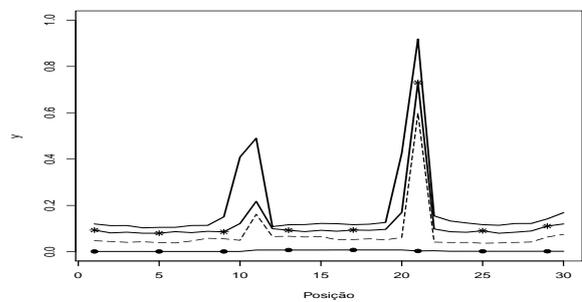
(c) Mod. gerado  $\lambda = 0.7$



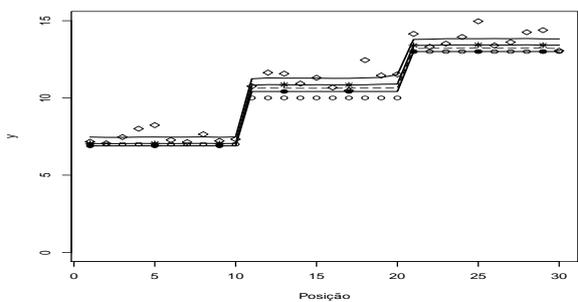
(d) Mod. gerado  $\lambda = 0.7$



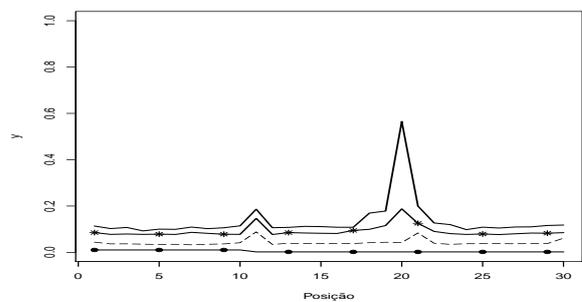
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

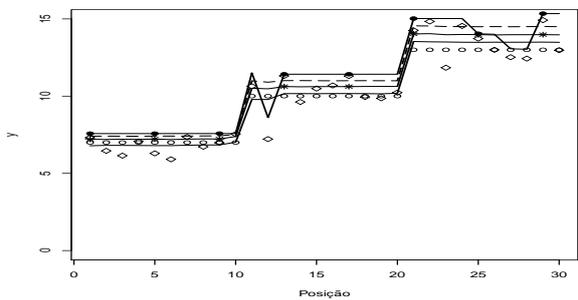


(g) Mod. gerado  $\lambda = 100$

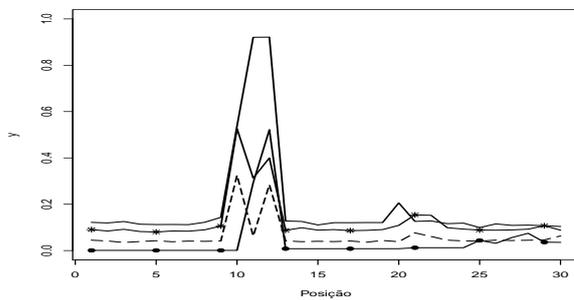


(h) Mod. gerado  $\lambda = 100$

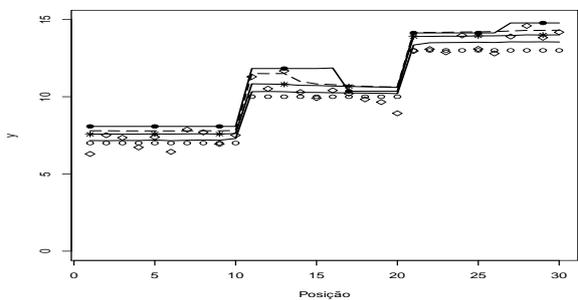
Figura D.43: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 30$ .



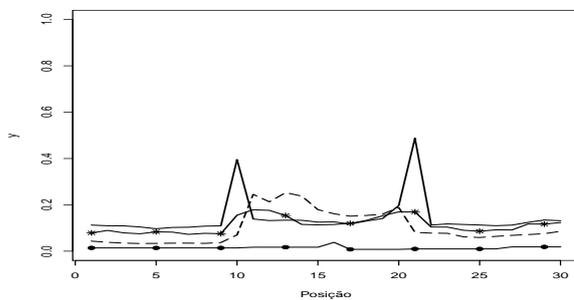
(a) Mod. gerado  $\lambda = 0$



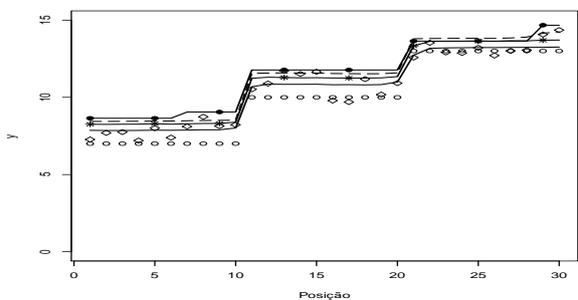
(b) Mod. gerado  $\lambda = 0$



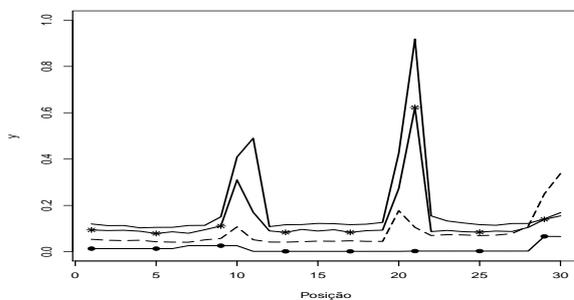
(c) Mod. gerado  $\lambda = 0.7$



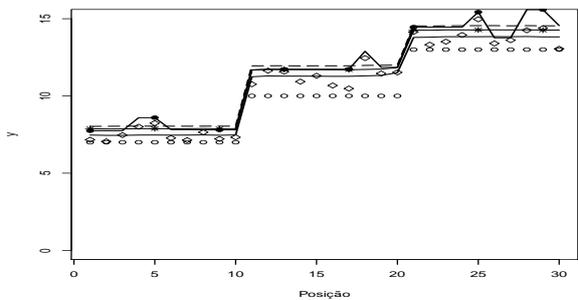
(d) Mod. gerado  $\lambda = 0.7$



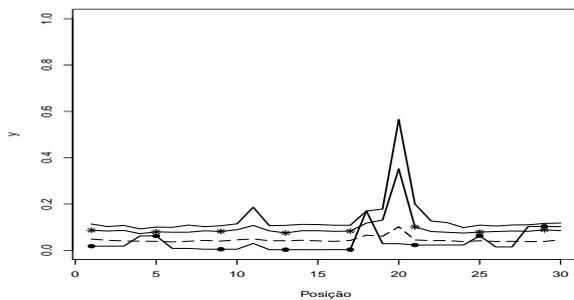
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

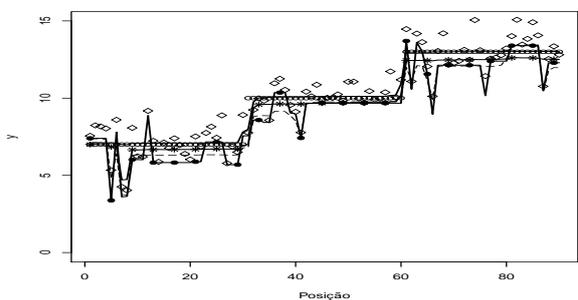


(g) Mod. gerado  $\lambda = 100$

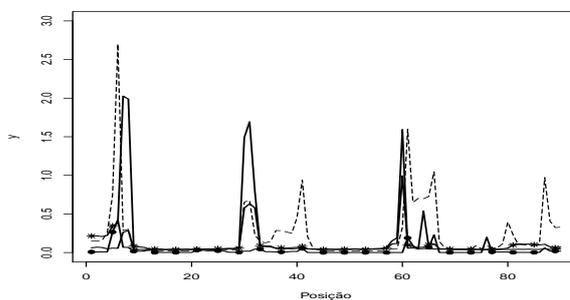


(h) Mod. gerado  $\lambda = 100$

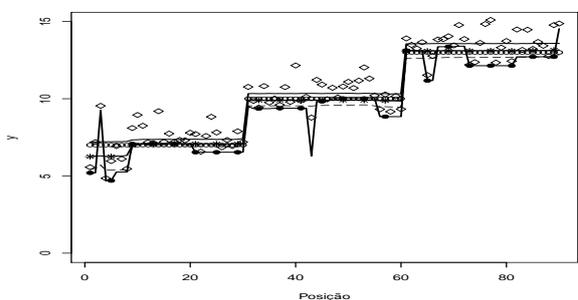
Figura D.44: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 30$ .



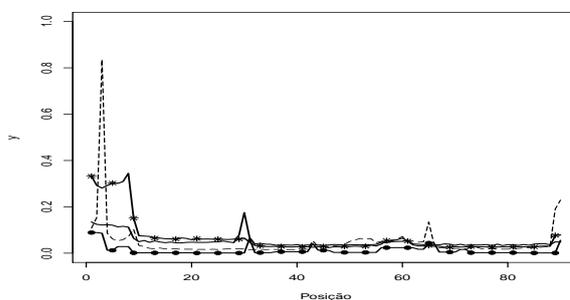
(a) Mod. gerado  $\lambda = 0$



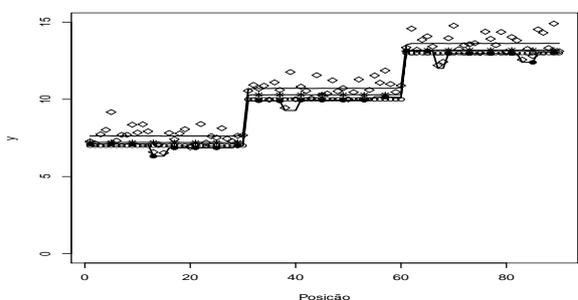
(b) Mod. gerado  $\lambda = 0$



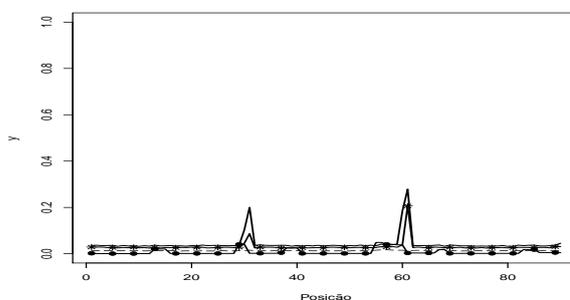
(c) Mod. gerado  $\lambda = 0.7$



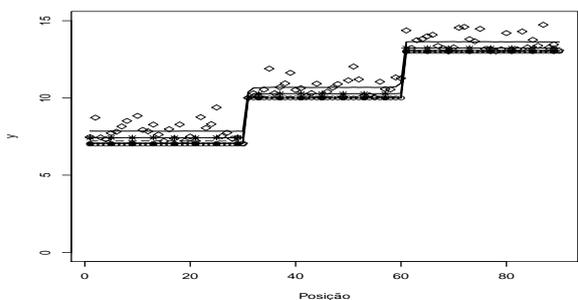
(d) Mod. gerado  $\lambda = 0.7$



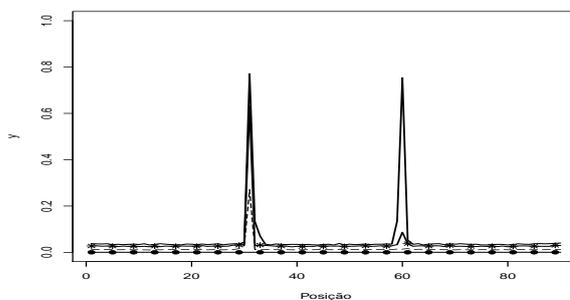
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

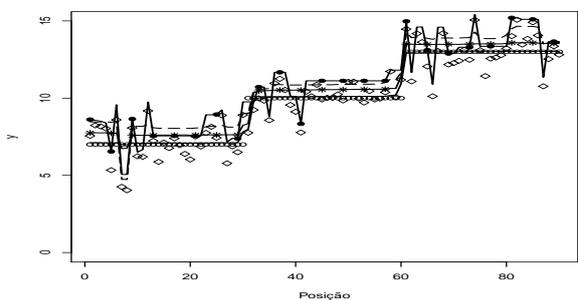


(g) Mod. gerado  $\lambda = 100$

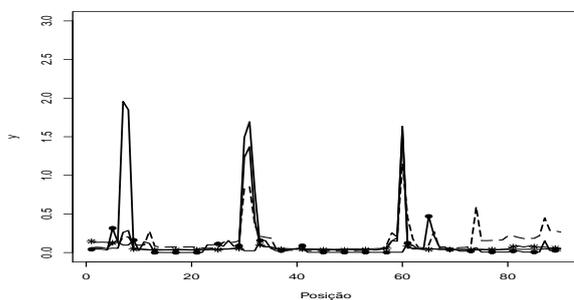


(h) Mod. gerado  $\lambda = 100$

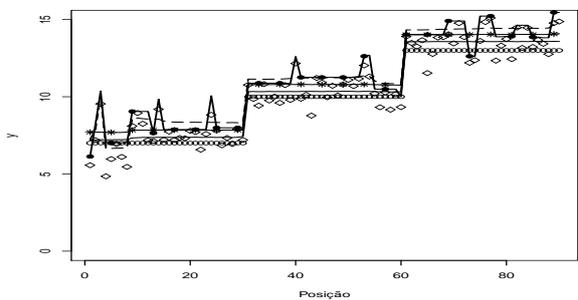
Figura D.45: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 90$ .



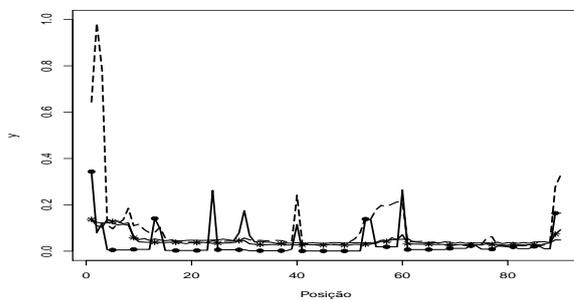
(a) Mod. gerado  $\lambda = 0$



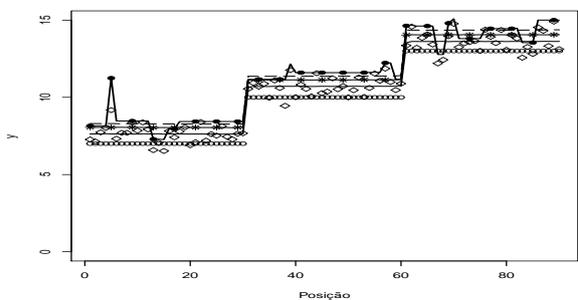
(b) Mod. gerado  $\lambda = 0$



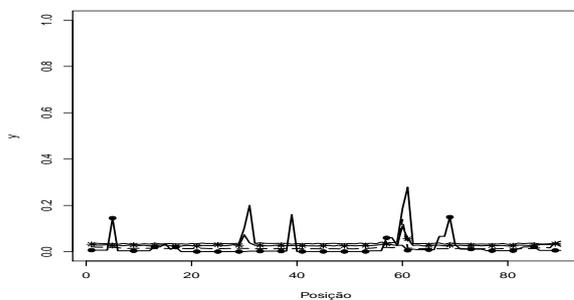
(c) Mod. gerado  $\lambda = 0.7$



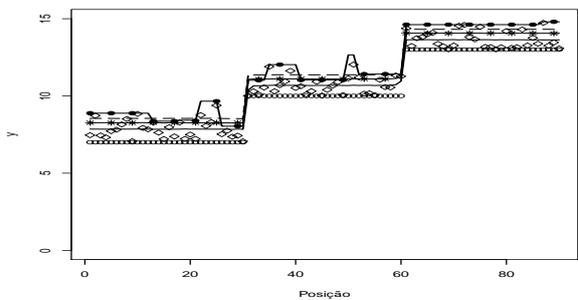
(d) Mod. gerado  $\lambda = 0.7$



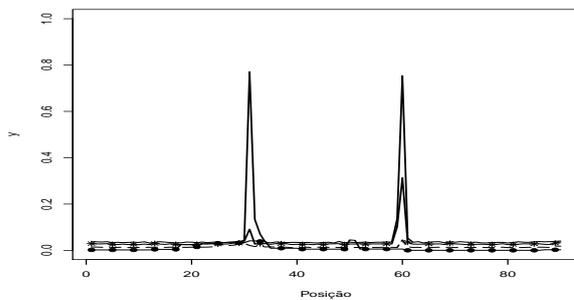
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

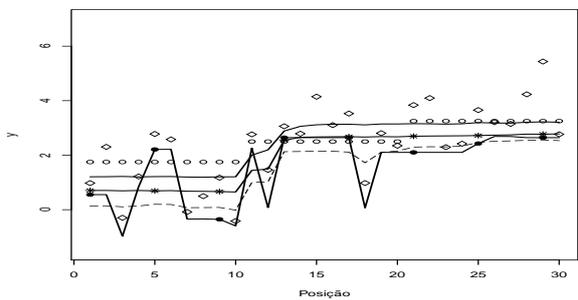


(g) Mod. gerado  $\lambda = 100$

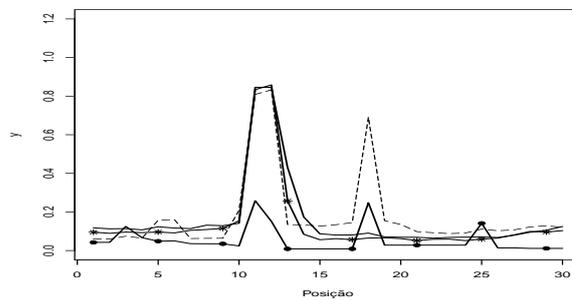


(h) Mod. gerado  $\lambda = 100$

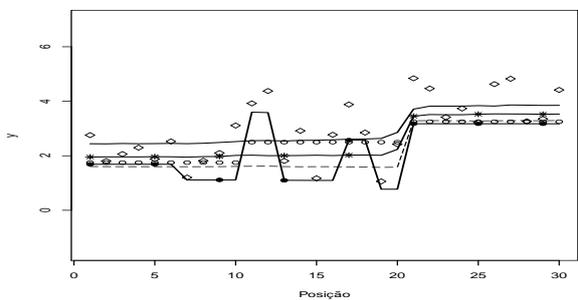
Figura D.46: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 90$ .



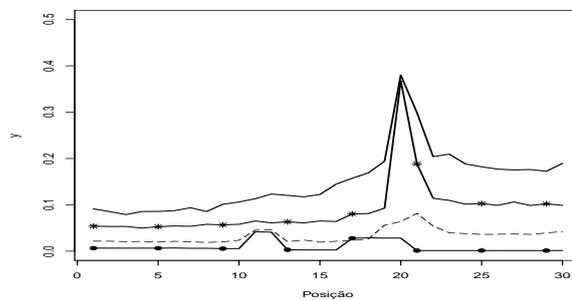
(a) Mod. gerado  $\lambda = 0$



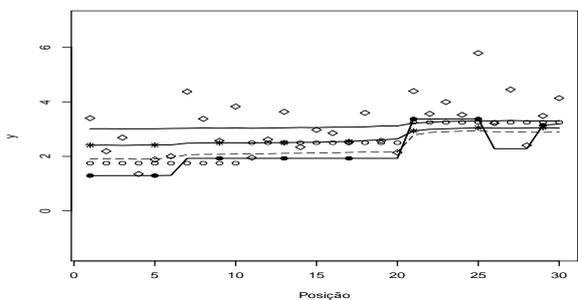
(b) Mod. gerado  $\lambda = 0$



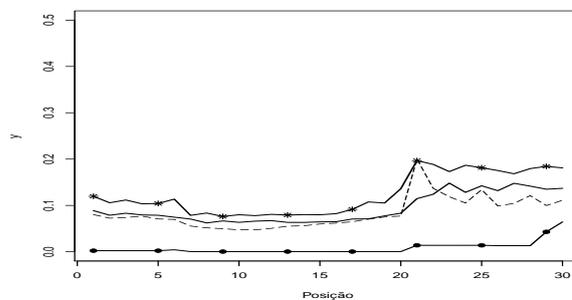
(c) Mod. gerado  $\lambda = 0.7$



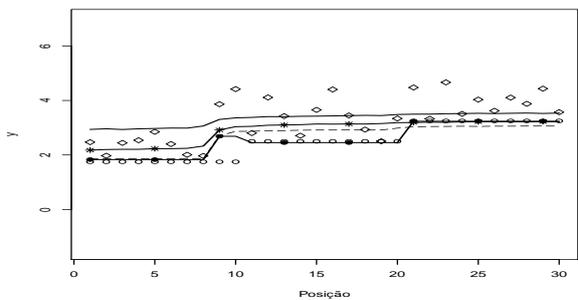
(d) Mod. gerado  $\lambda = 0.7$



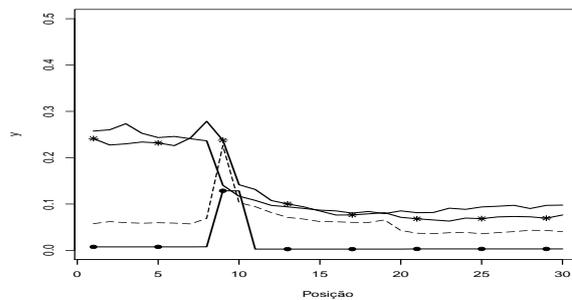
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

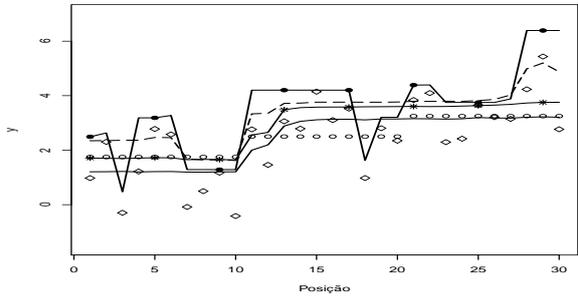


(g) Mod. gerado  $\lambda = 100$

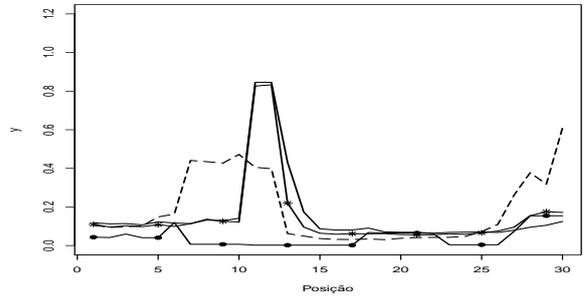


(h) Mod. gerado  $\lambda = 100$

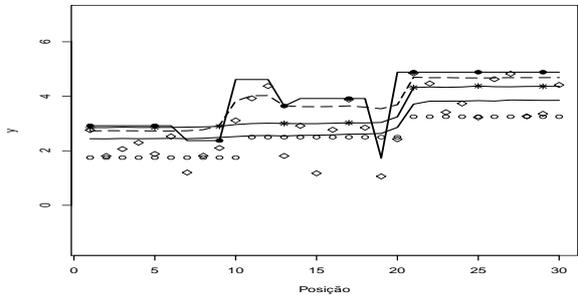
Figura D.47: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 30$ .



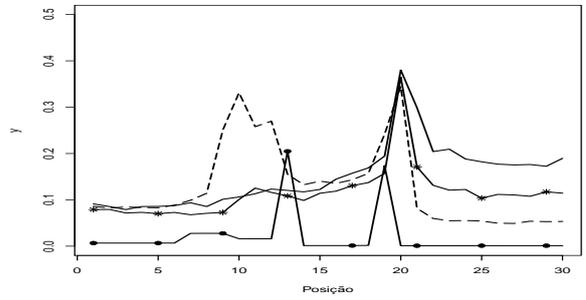
(a) Mod. gerado  $\lambda = 0$



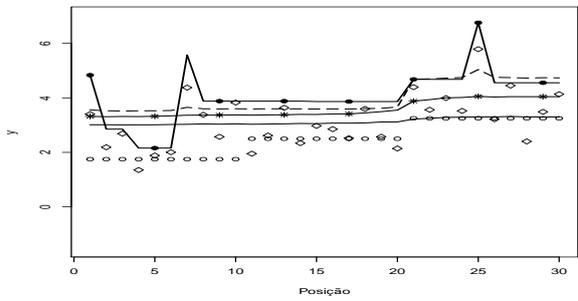
(b) Mod. gerado  $\lambda = 0$



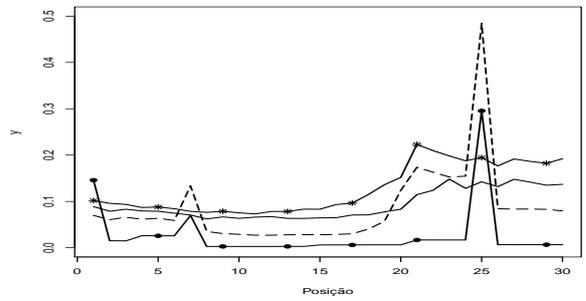
(c) Mod. gerado  $\lambda = 0.7$



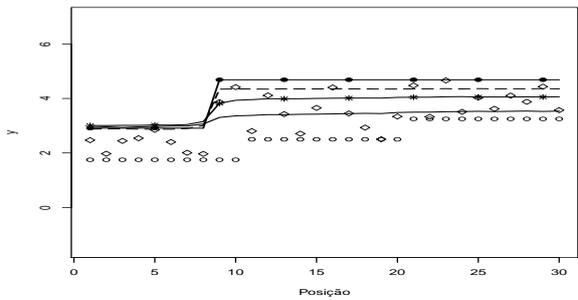
(d) Mod. gerado  $\lambda = 0.7$



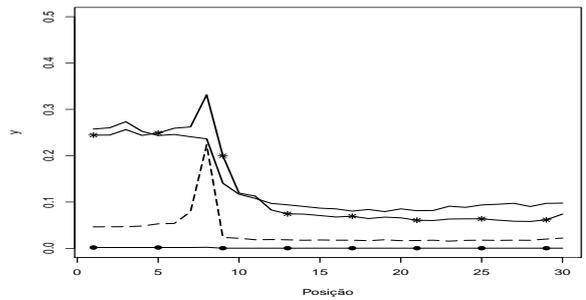
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

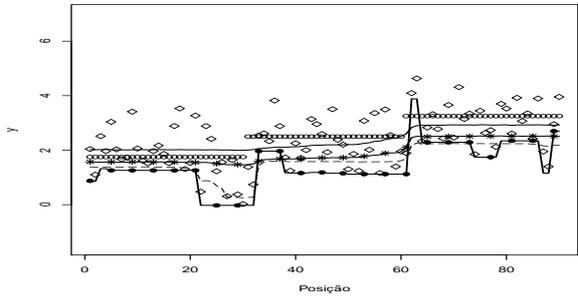


(g) Mod. gerado  $\lambda = 100$

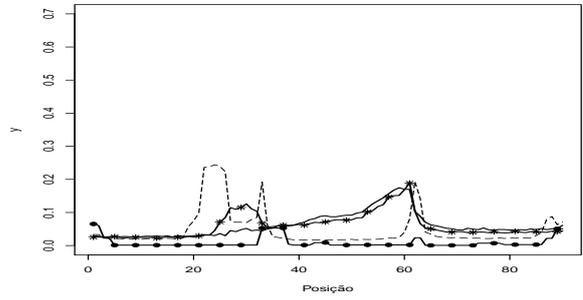


(h) Mod. gerado  $\lambda = 100$

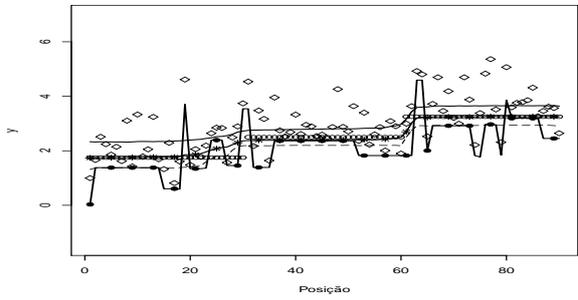
Figura D.48: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (o), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 30$ .



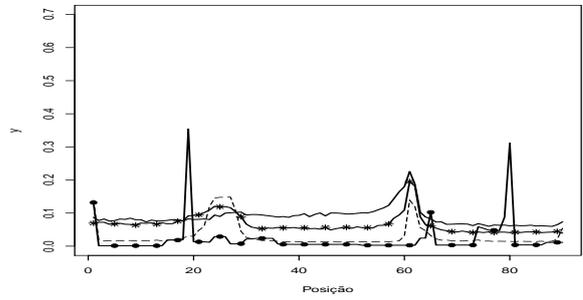
(a) Mod. gerado  $\lambda = 0$



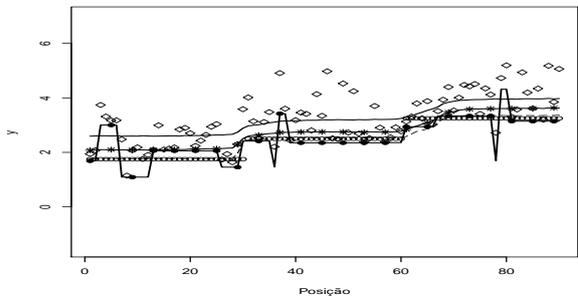
(b) Mod. gerado  $\lambda = 0$



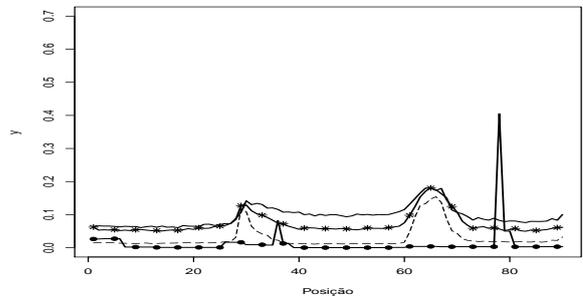
(c) Mod. gerado  $\lambda = 0.7$



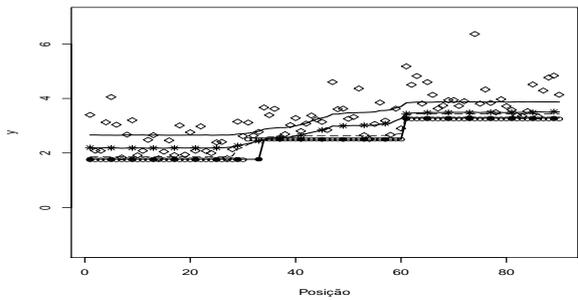
(d) Mod. gerado  $\lambda = 0.7$



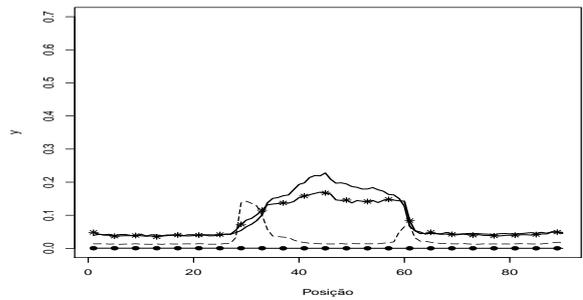
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

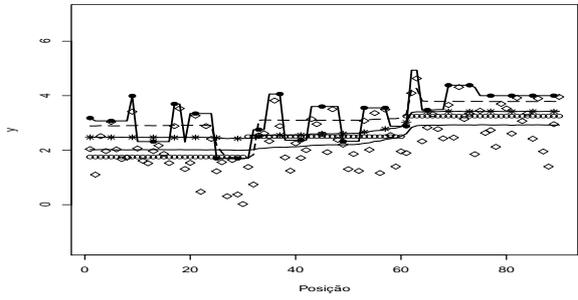


(g) Mod. gerado  $\lambda = 100$

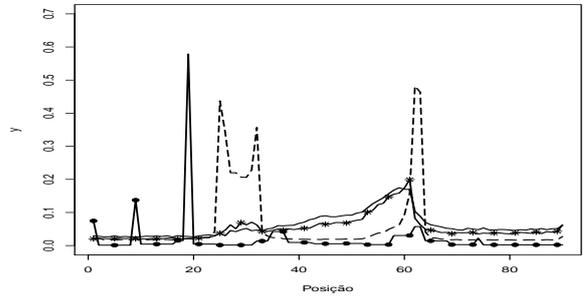


(h) Mod. gerado  $\lambda = 100$

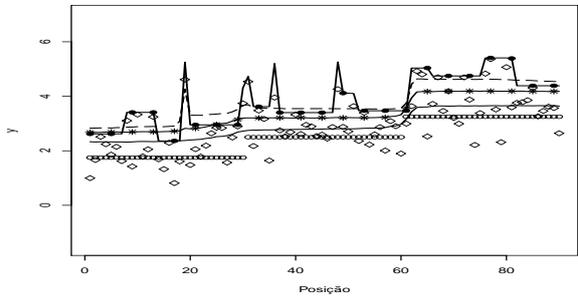
Figura D.49: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 90$ .



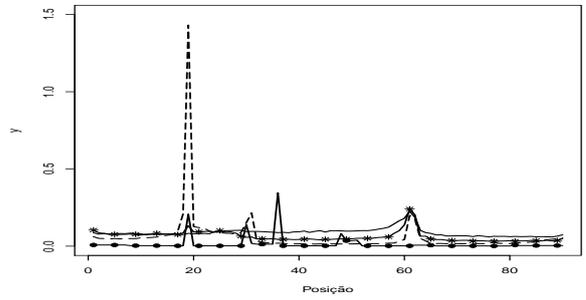
(a) Mod. gerado  $\lambda = 0$



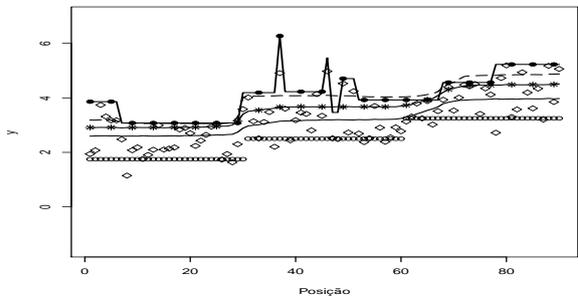
(b) Mod. gerado  $\lambda = 0$



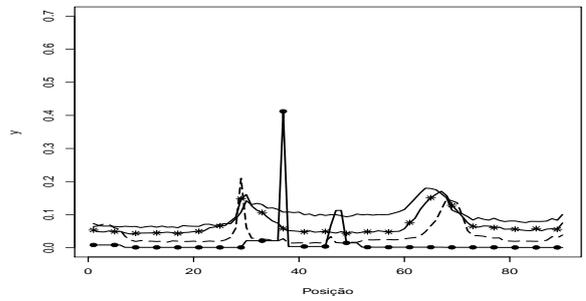
(c) Mod. gerado  $\lambda = 0.7$



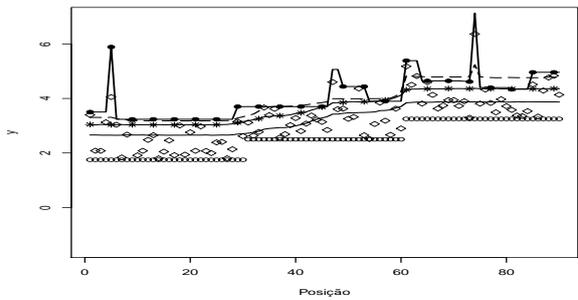
(d) Mod. gerado  $\lambda = 0.7$



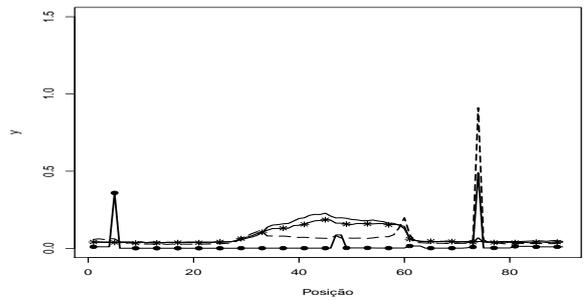
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

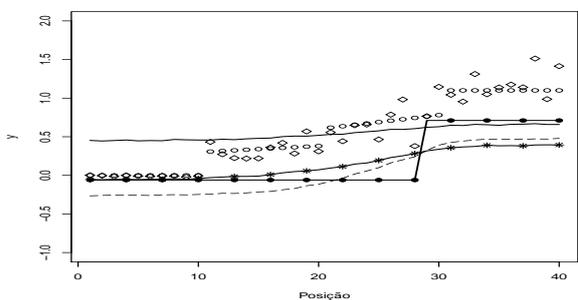


(g) Mod. gerado  $\lambda = 100$

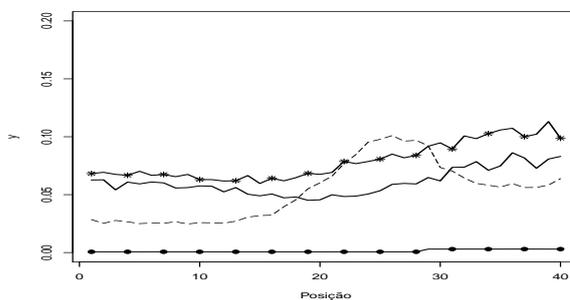


(h) Mod. gerado  $\lambda = 100$

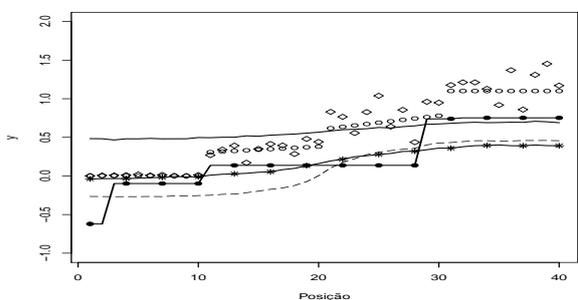
Figura D.50: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 90$ .



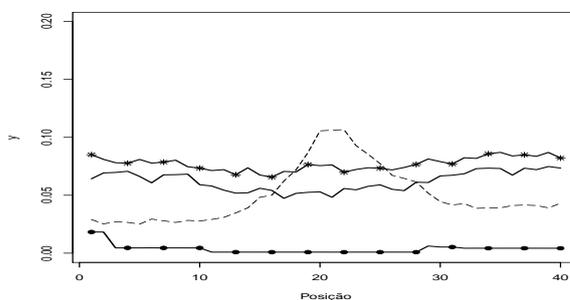
(a) Mod. gerado  $\lambda = 0$



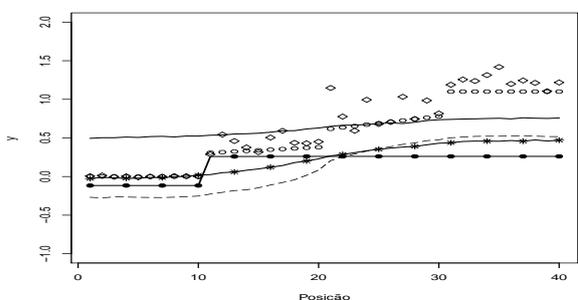
(b) Mod. gerado  $\lambda = 0$



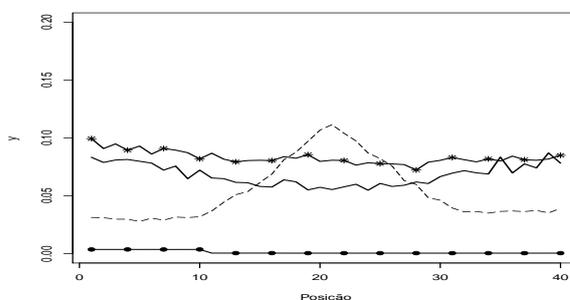
(c) Mod. gerado  $\lambda = 0.7$



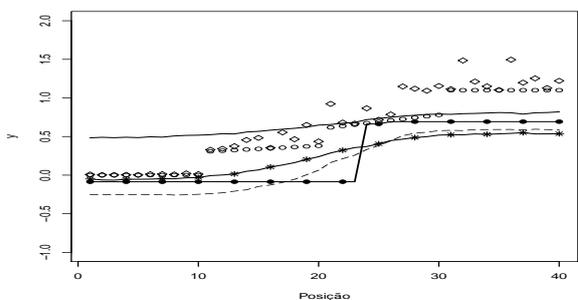
(d) Mod. gerado  $\lambda = 0.7$



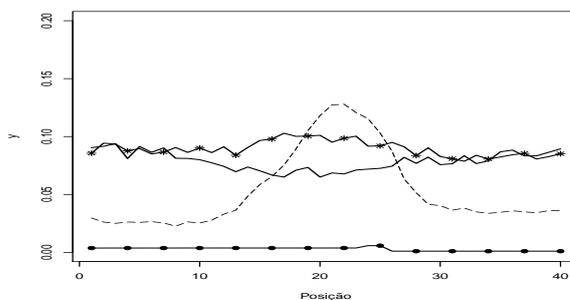
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

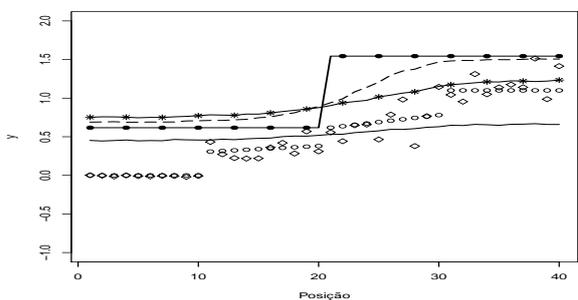


(g) Mod. gerado  $\lambda = 100$

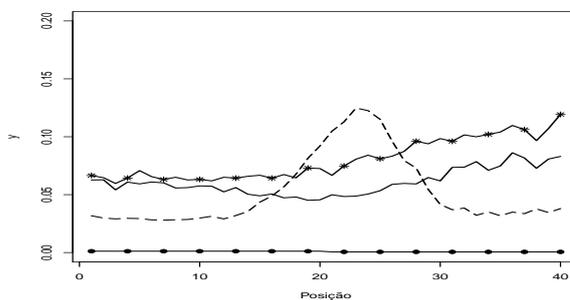


(h) Mod. gerado  $\lambda = 100$

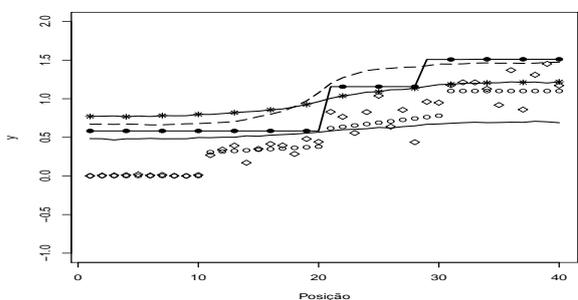
Figura D.51: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 30$ .



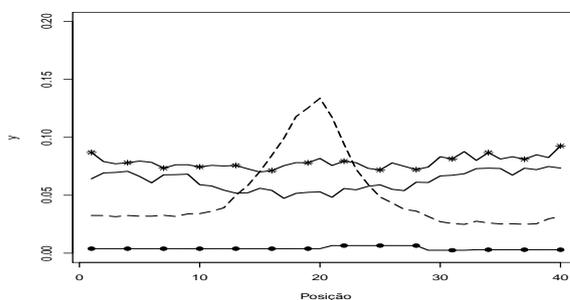
(a) Mod. gerado  $\lambda = 0$



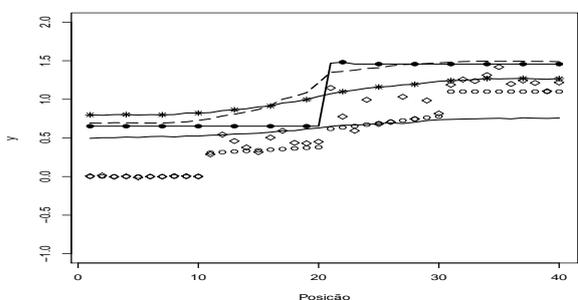
(b) Mod. gerado  $\lambda = 0$



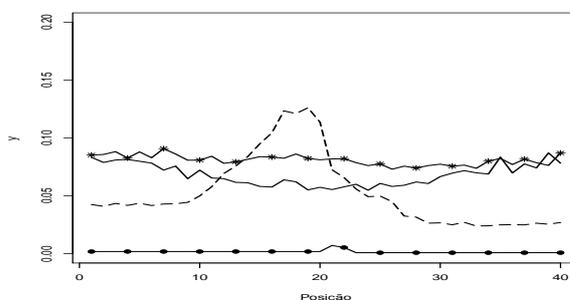
(c) Mod. gerado  $\lambda = 0.7$



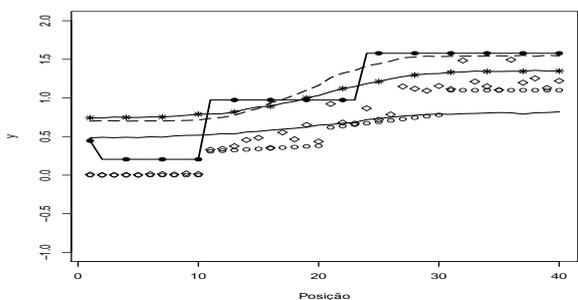
(d) Mod. gerado  $\lambda = 0.7$



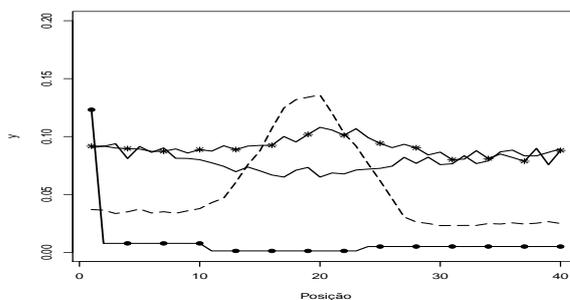
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

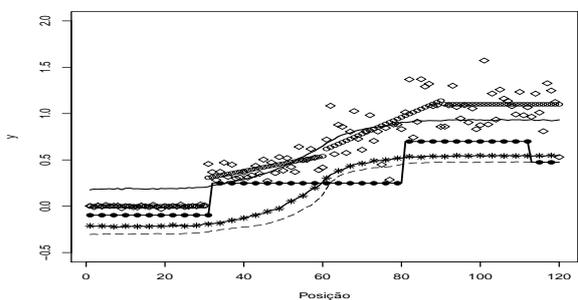


(g) Mod. gerado  $\lambda = 100$

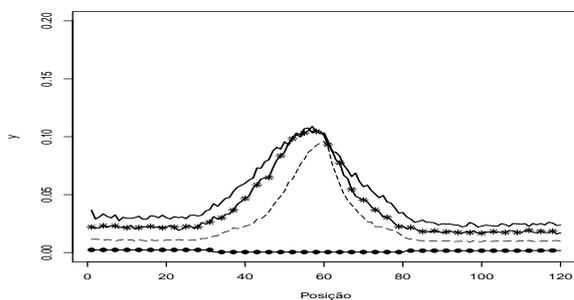


(h) Mod. gerado  $\lambda = 100$

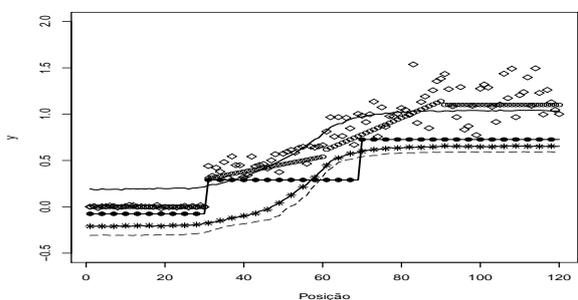
Figura D.52: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 30$ .



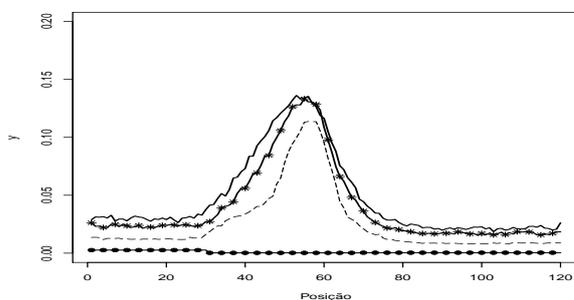
(a) Mod. gerado  $\lambda = 0$



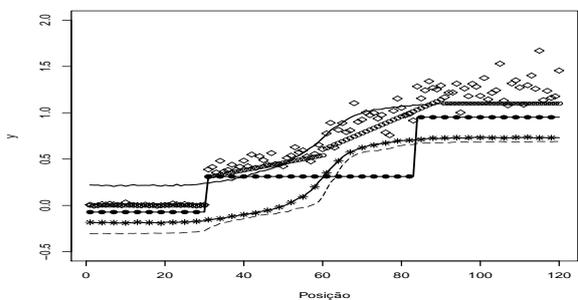
(b) Mod. gerado  $\lambda = 0$



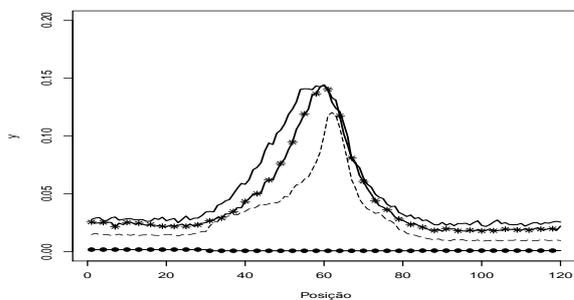
(c) Mod. gerado  $\lambda = 0.7$



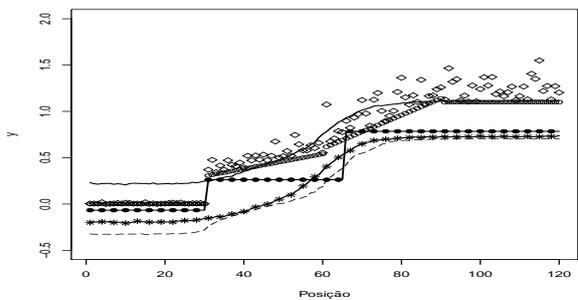
(d) Mod. gerado  $\lambda = 0.7$



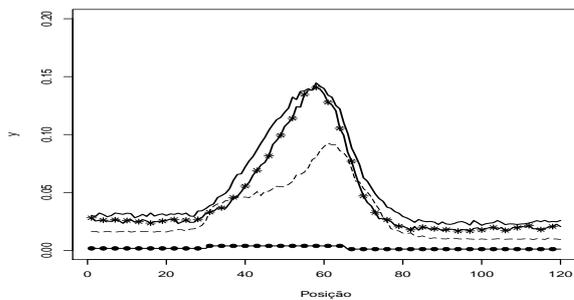
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

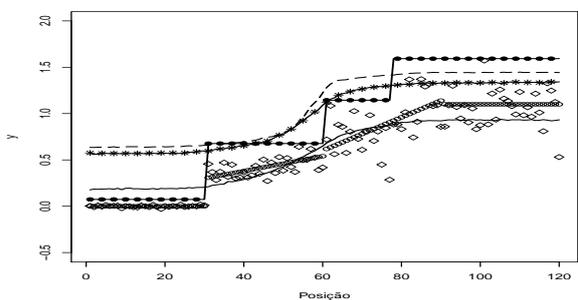


(g) Mod. gerado  $\lambda = 100$

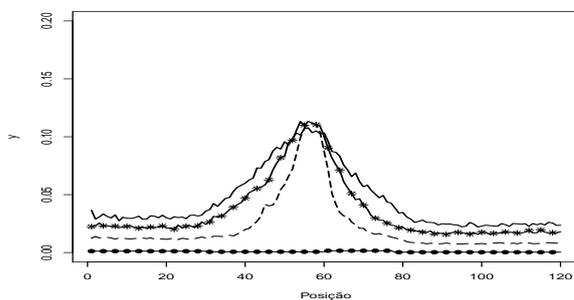


(h) Mod. gerado  $\lambda = 100$

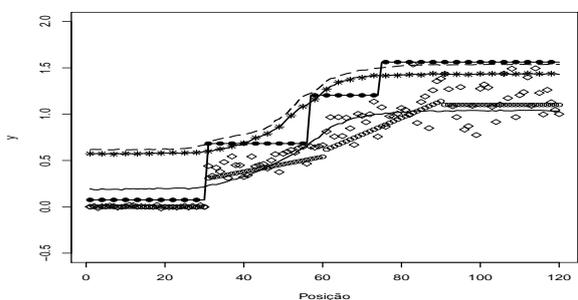
Figura D.53: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = 100$  (linha com  $\bullet$ ),  $\lambda = 2$  (linha tracejada),  $\lambda = 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 90$ .



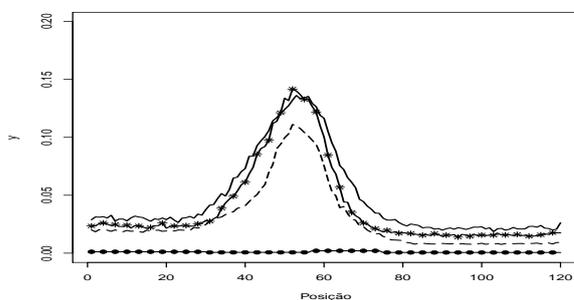
(a) Mod. gerado  $\lambda = 0$



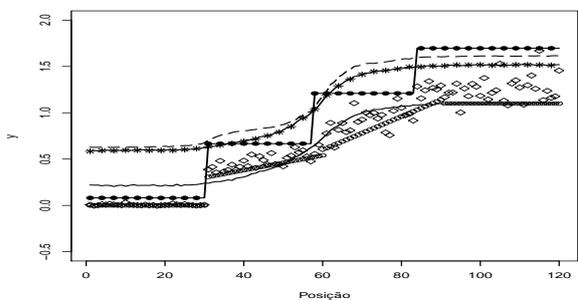
(b) Mod. gerado  $\lambda = 0$



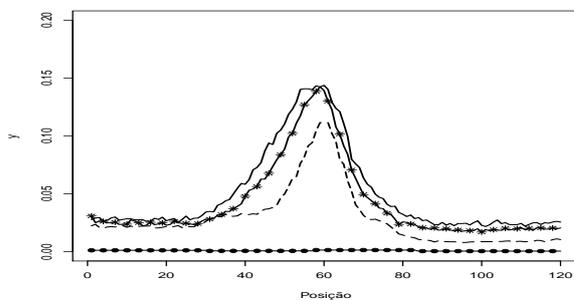
(c) Mod. gerado  $\lambda = 0.7$



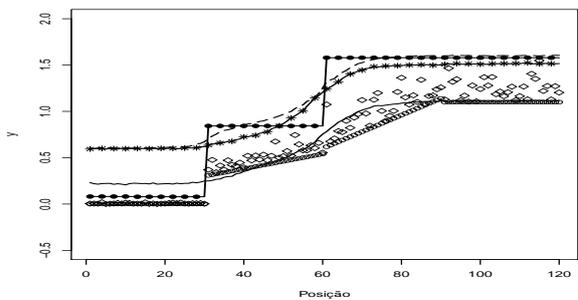
(d) Mod. gerado  $\lambda = 0.7$



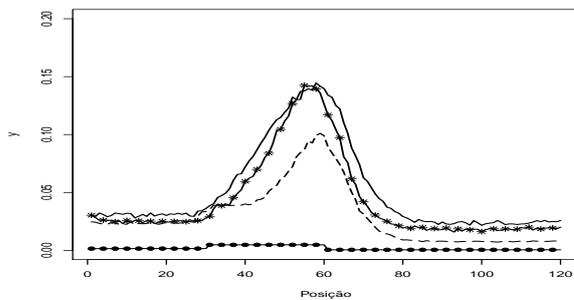
(e) Mod. gerado  $\lambda = 2$



(f) Mod. gerado  $\lambda = 2$

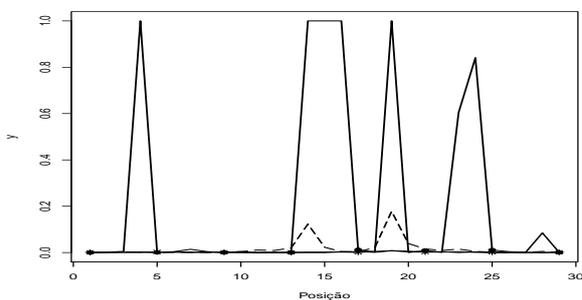


(g) Mod. gerado  $\lambda = 100$

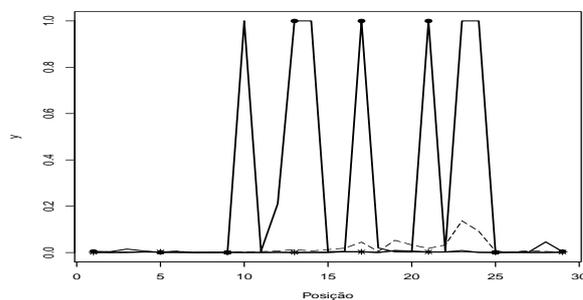


(h) Mod. gerado  $\lambda = 100$

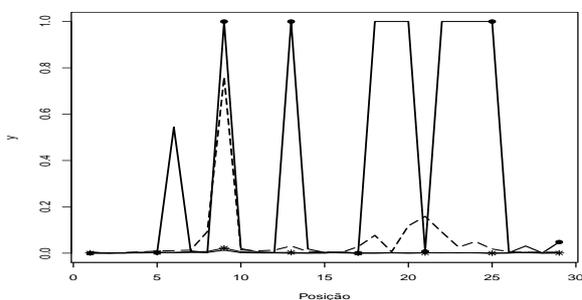
Figura D.54: Estimativas produto (a), (c), (e) e (g) e variância (b), (d), (f) e (h) *a posteriori* para cada  $\mu_i$ . Média verdadeira (c), estimativas para  $\lambda = -100$  (linha com  $\bullet$ ),  $\lambda = -2$  (linha tracejada),  $\lambda = -0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 90$ .



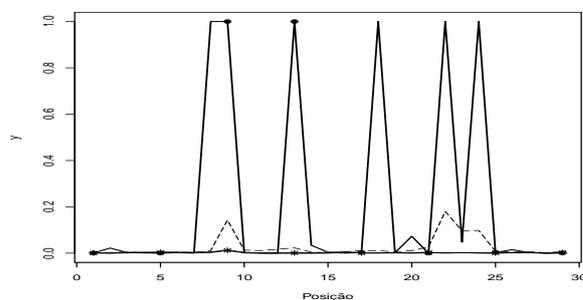
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



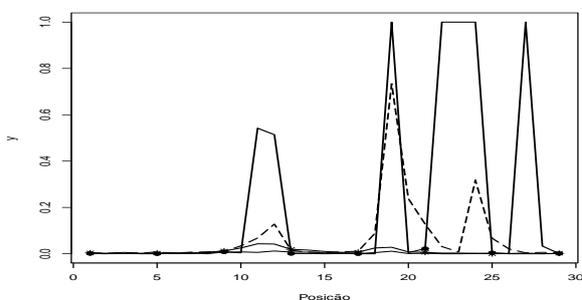
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



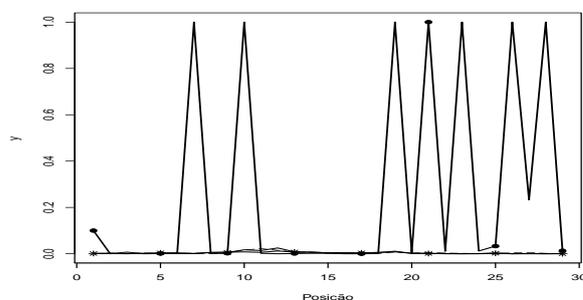
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



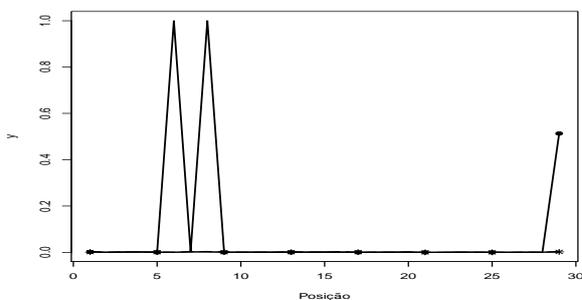
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



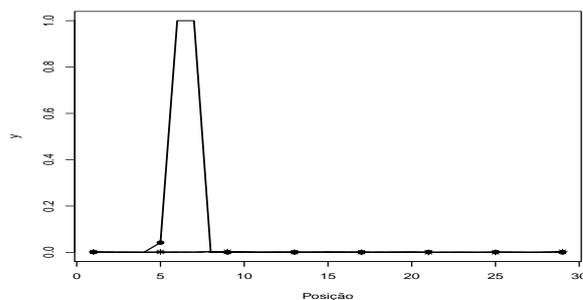
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

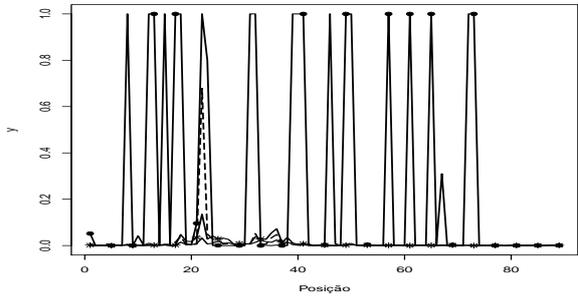


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

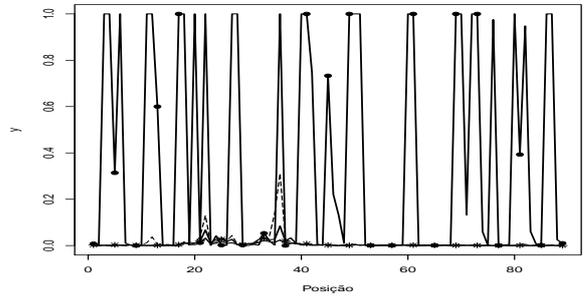


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

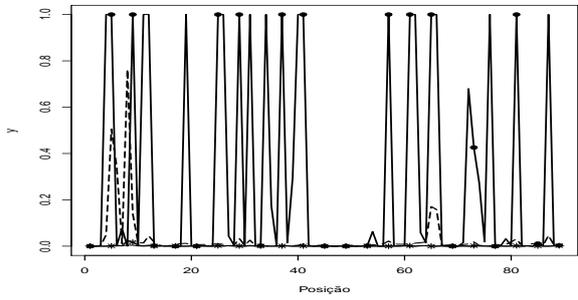
Figura D.55: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 30$ .



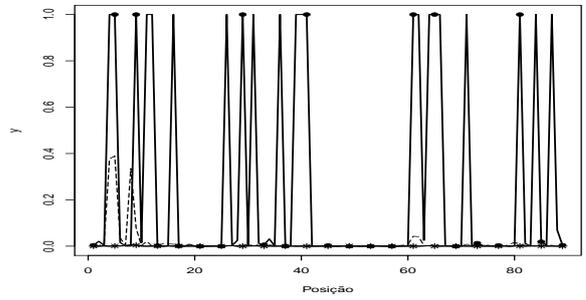
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



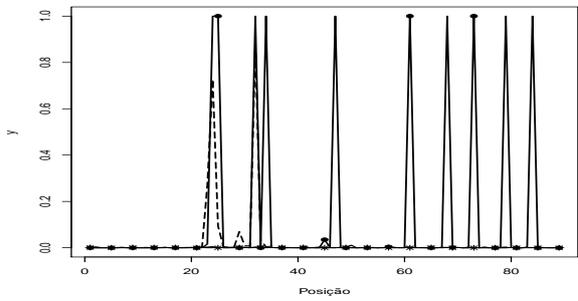
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



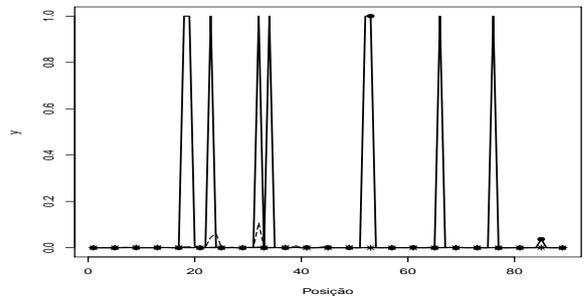
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



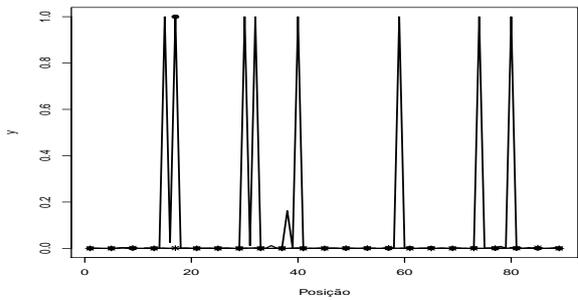
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



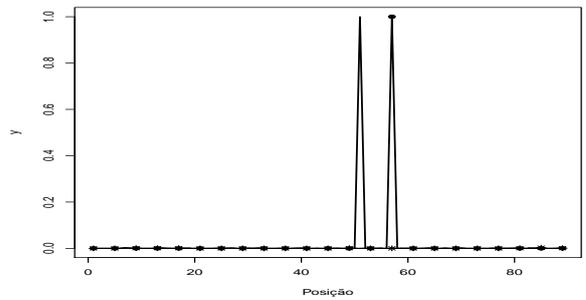
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

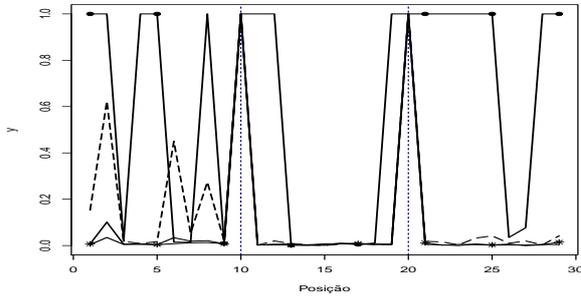


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

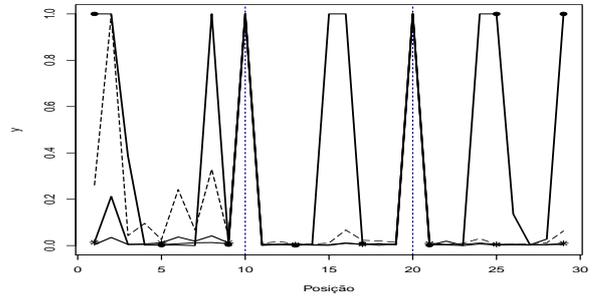


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

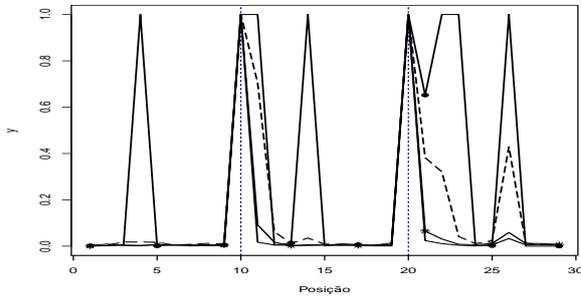
Figura D.56: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário sem mudança,  $n = 90$ .



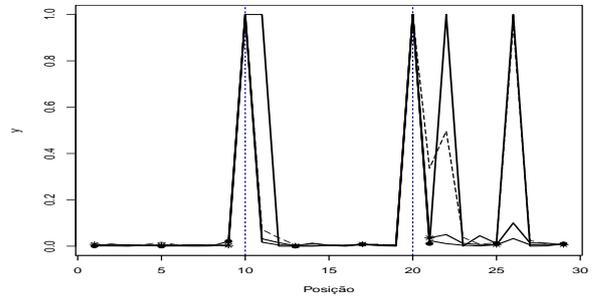
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



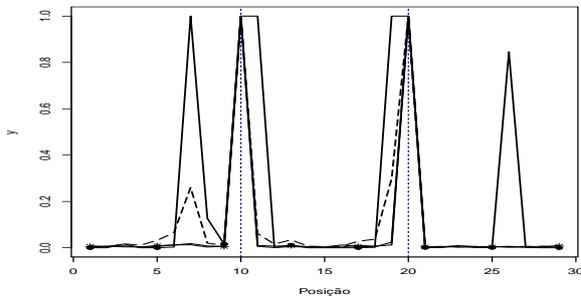
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



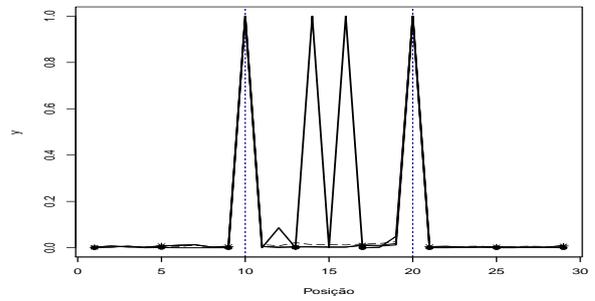
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



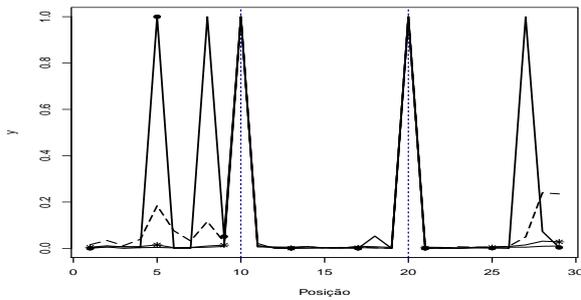
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



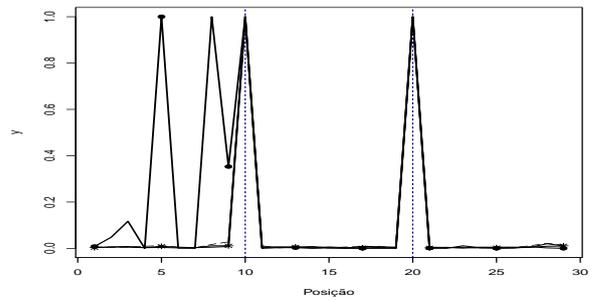
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

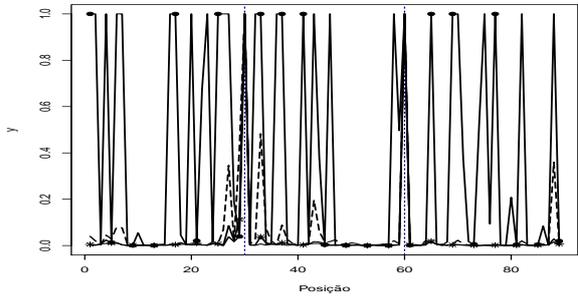


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

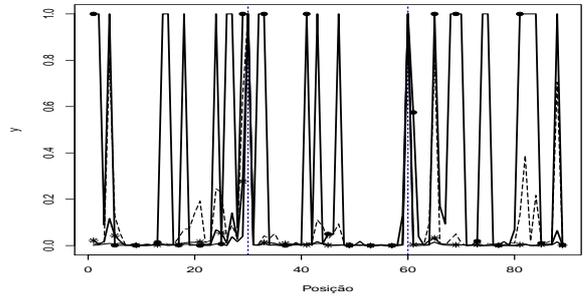


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

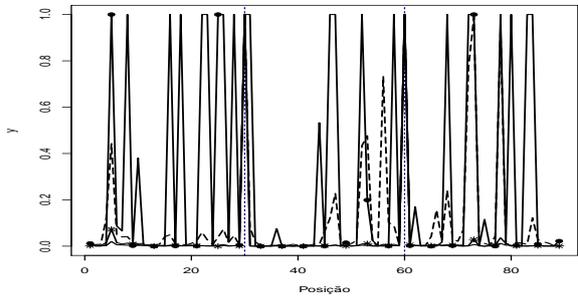
Figura D.57: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 30$ .



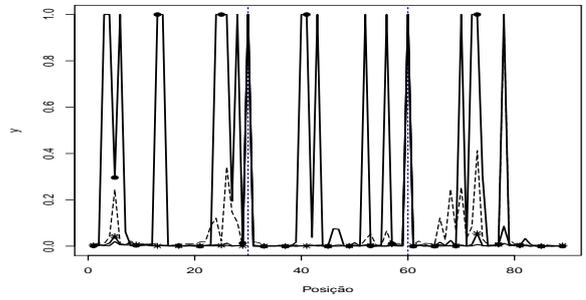
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



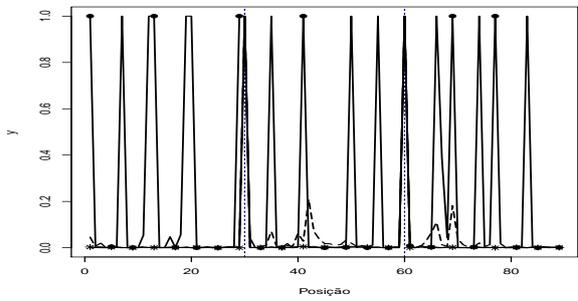
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



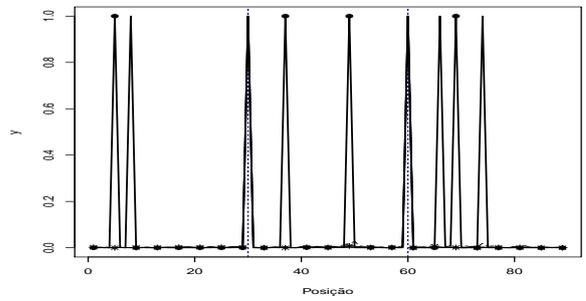
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



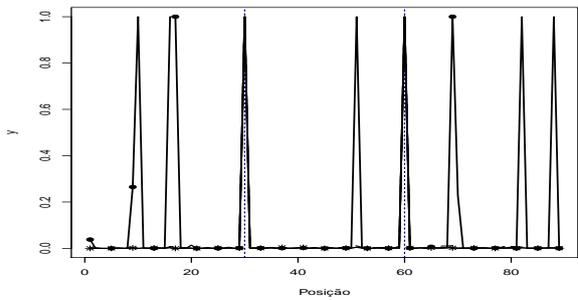
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



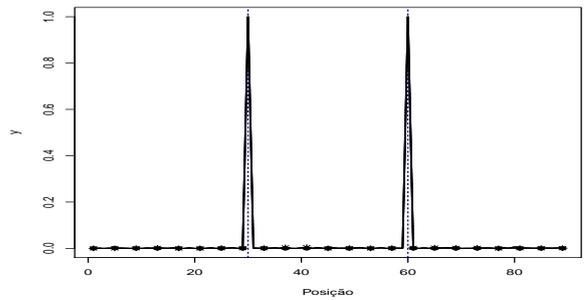
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

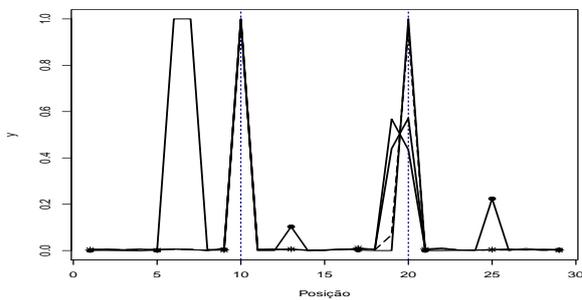


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

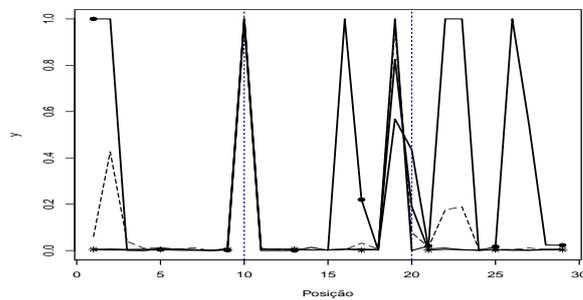


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

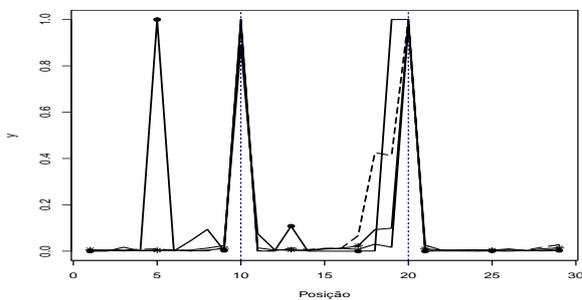
Figura D.58: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau não simétrico,  $n = 90$ .



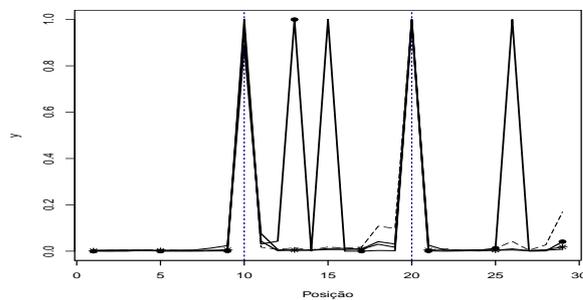
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



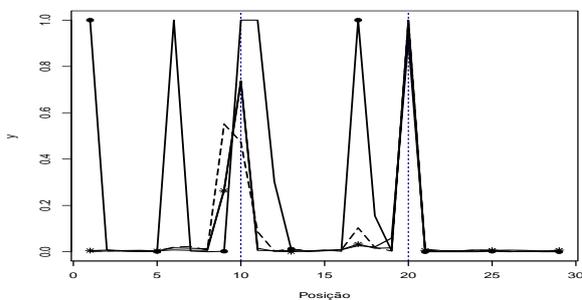
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



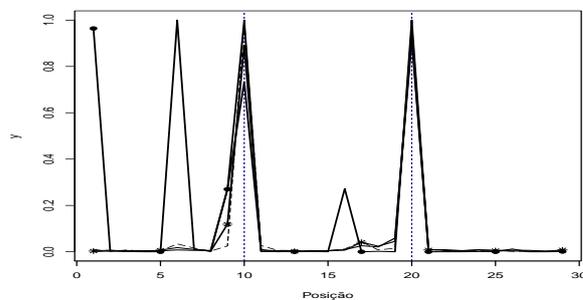
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



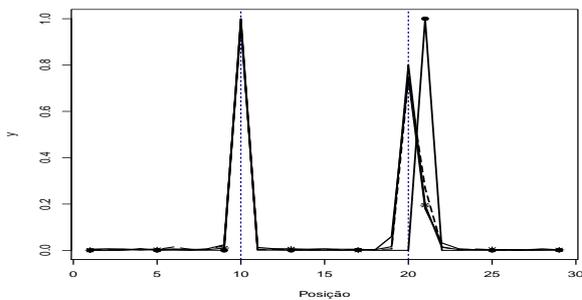
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



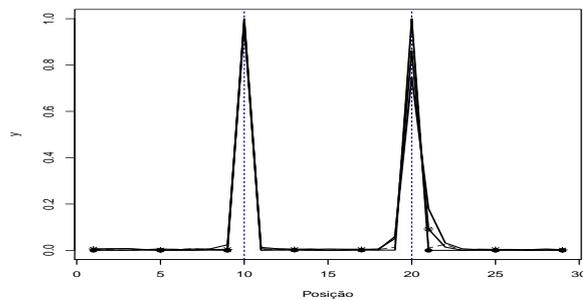
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

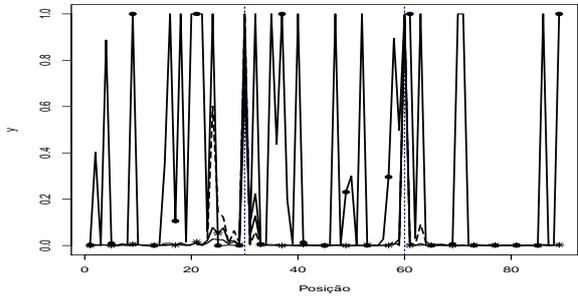


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

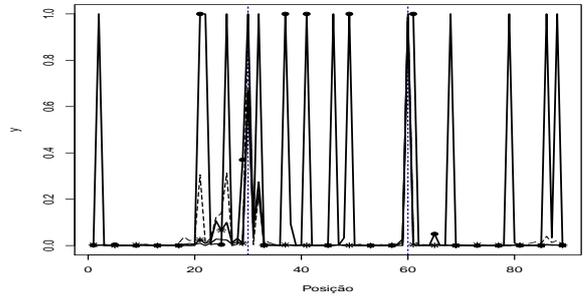


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

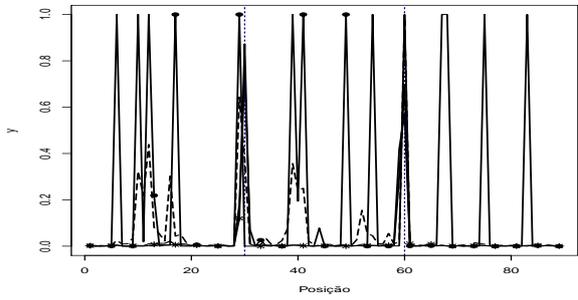
Figura D.59: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 30$ .



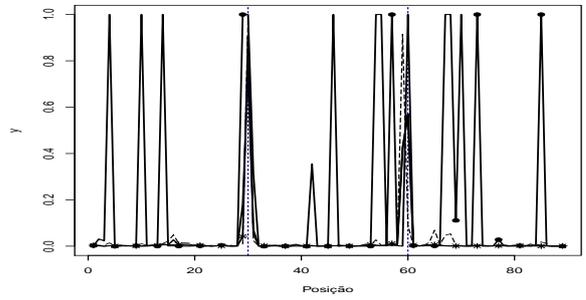
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



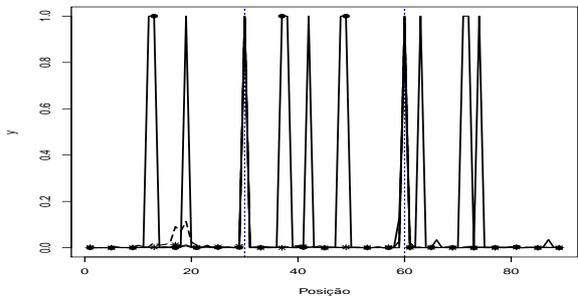
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



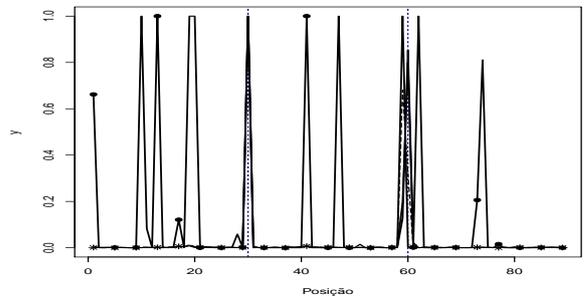
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



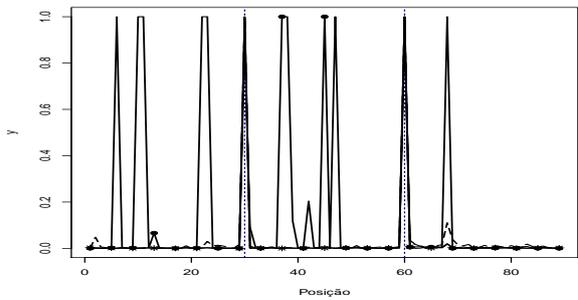
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



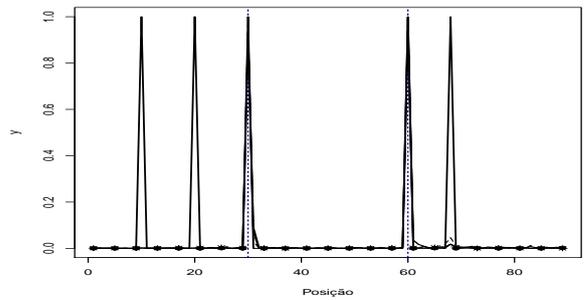
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

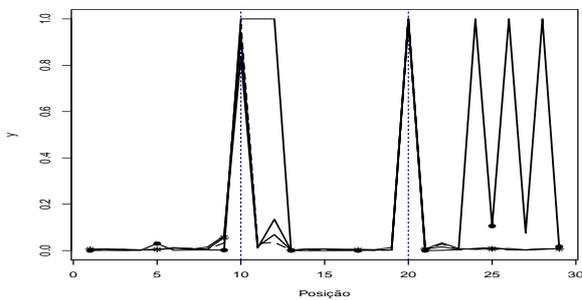


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

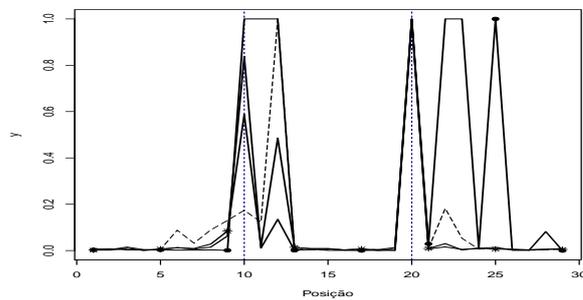


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

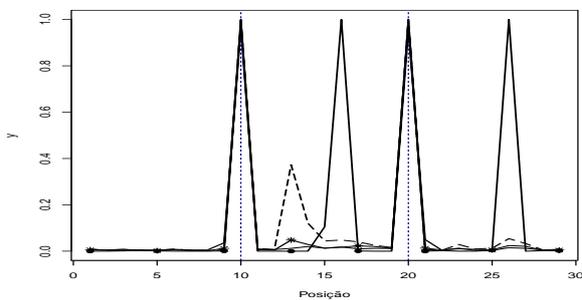
Figura D.60: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário degrau positivo,  $n = 90$ .



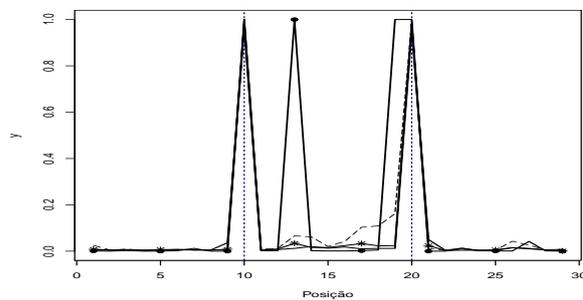
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



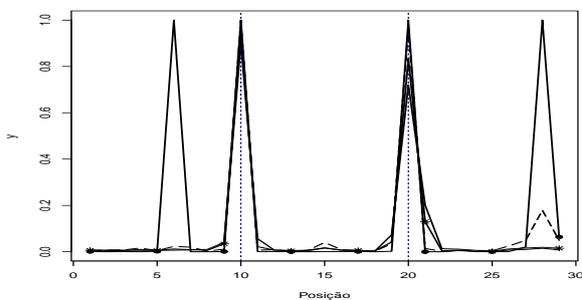
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



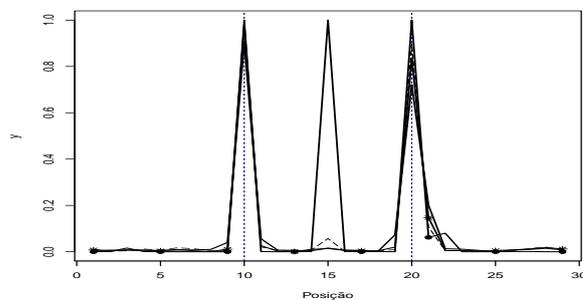
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



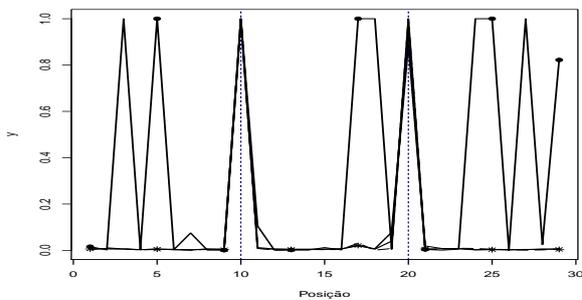
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



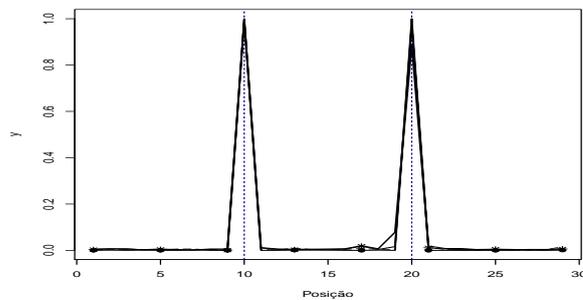
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

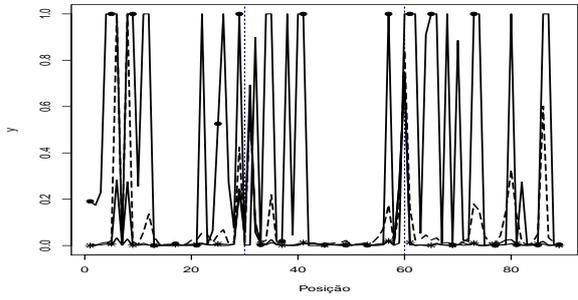


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

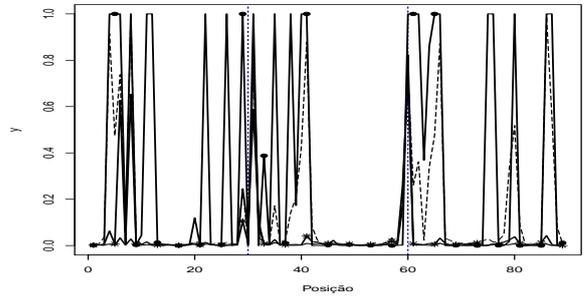


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

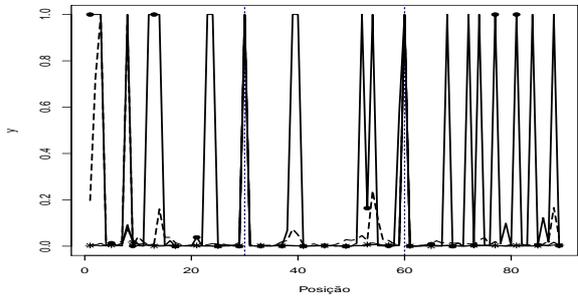
Figura D.61: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 30$ .



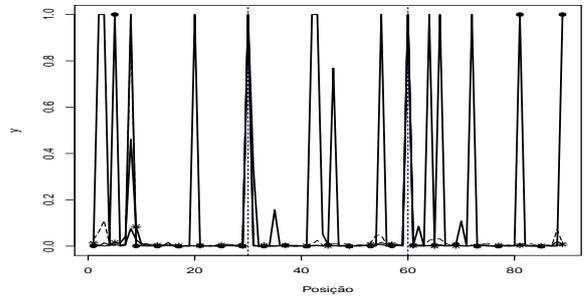
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



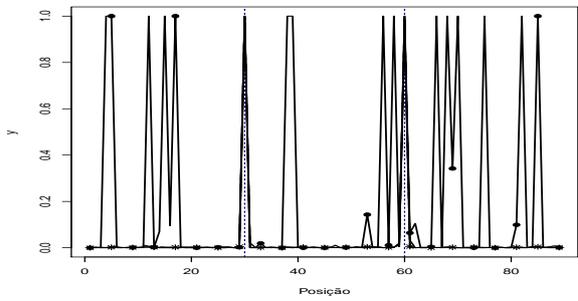
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



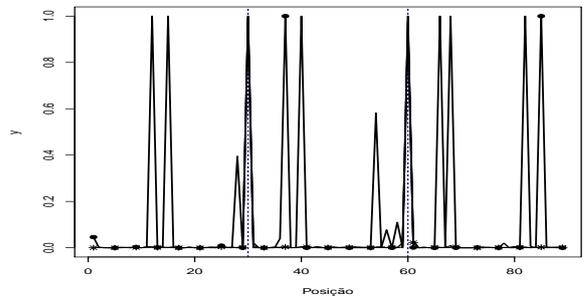
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



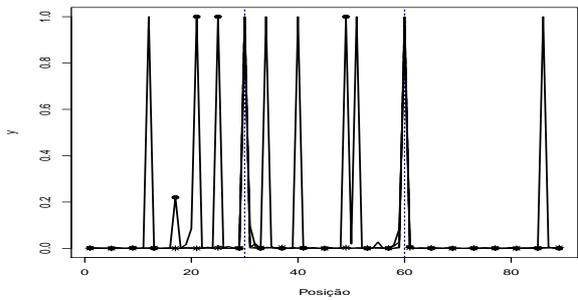
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



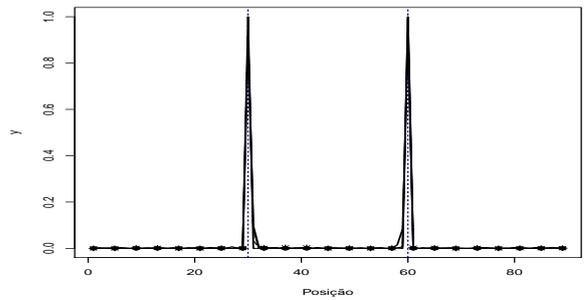
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

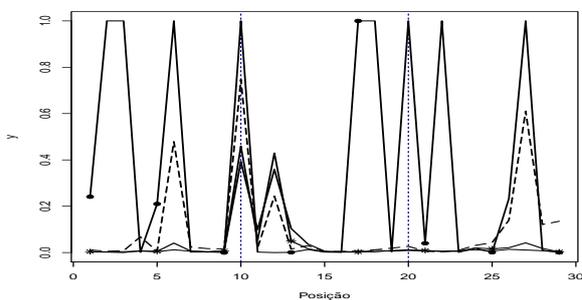


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

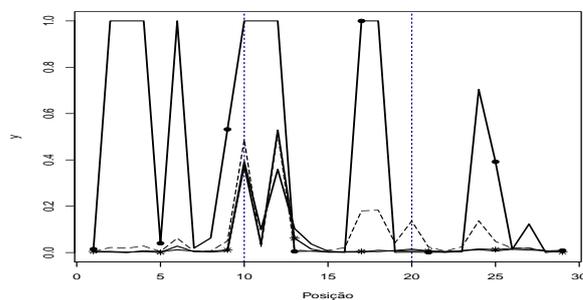


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

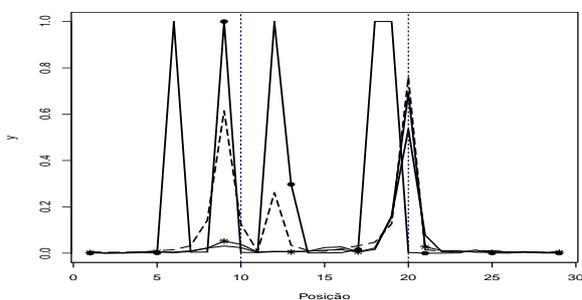
Figura D.62: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário escada crescente,  $n = 90$ .



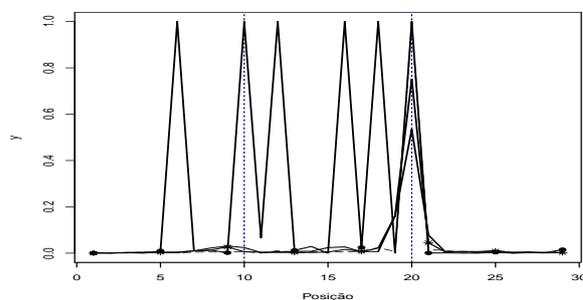
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



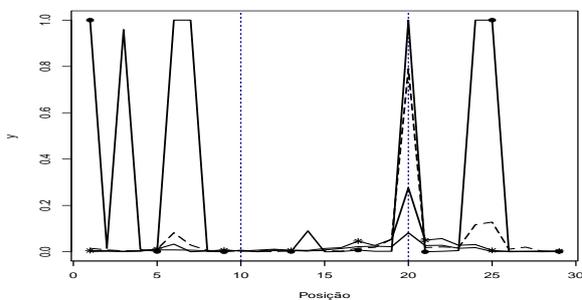
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



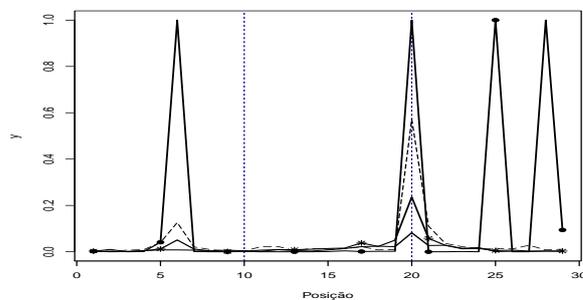
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



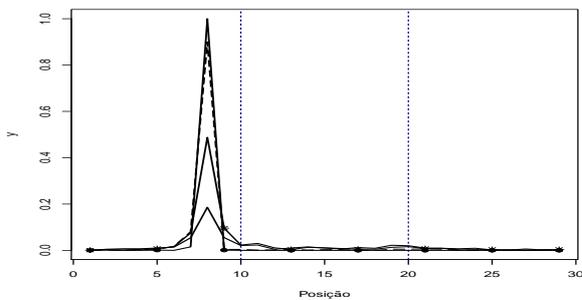
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



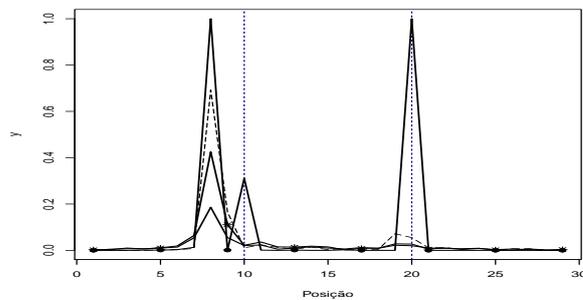
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

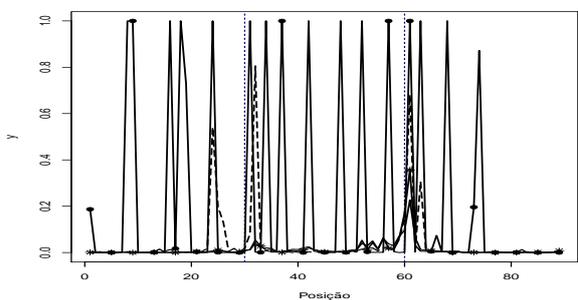


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

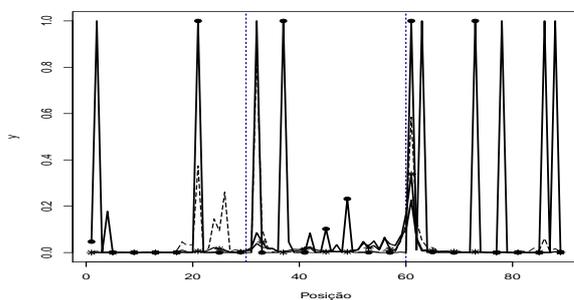


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

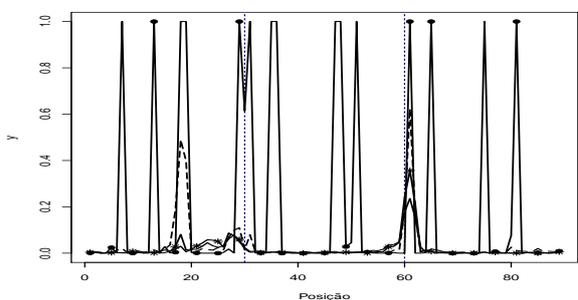
Figura D.63: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 30$ .



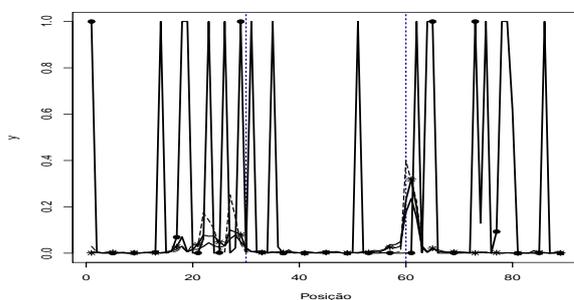
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



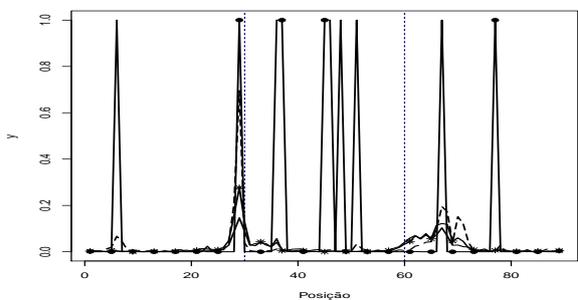
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



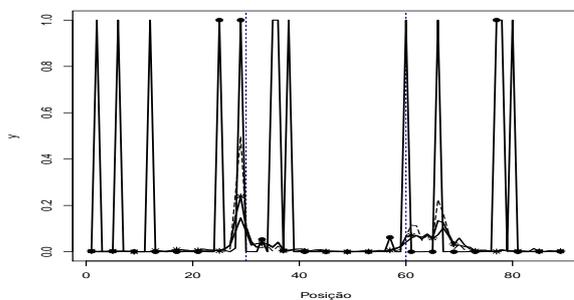
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



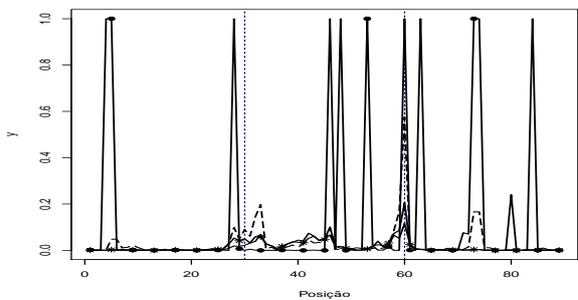
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



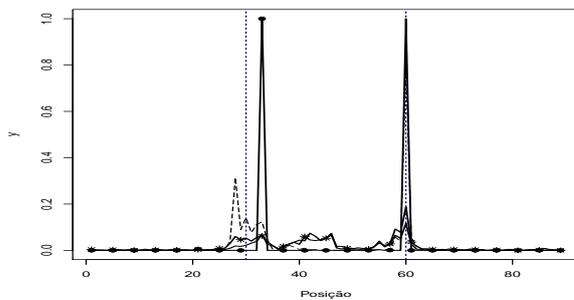
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

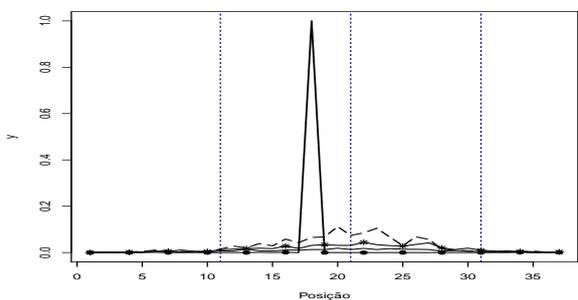


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

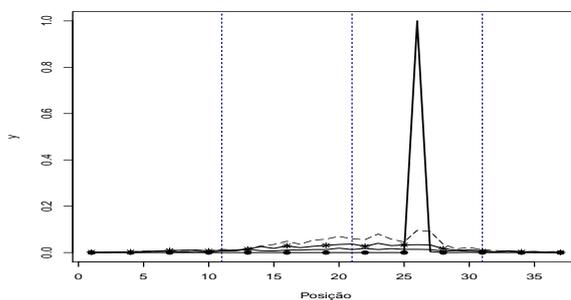


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

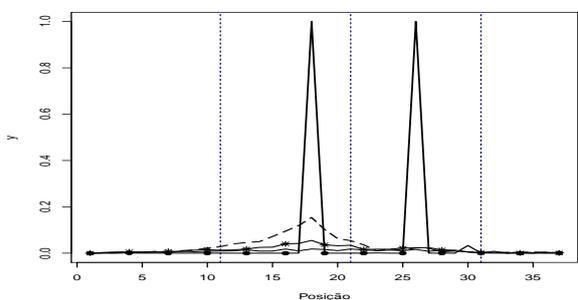
Figura D.64: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário epsilon,  $n = 90$ .



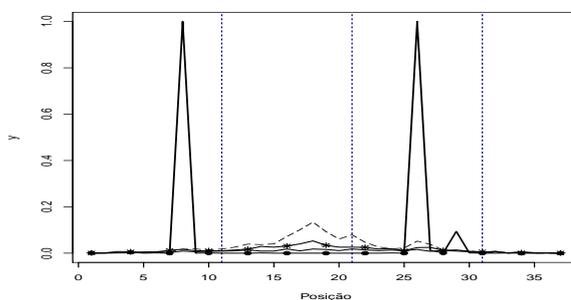
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



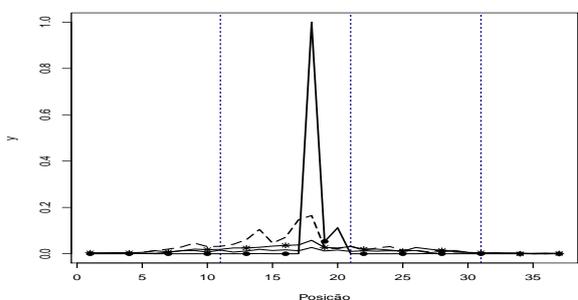
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



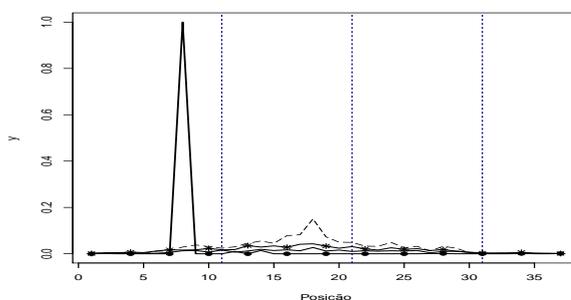
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



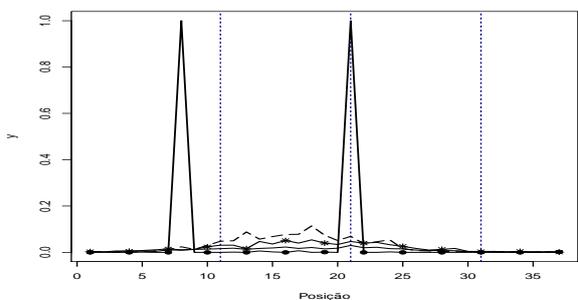
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



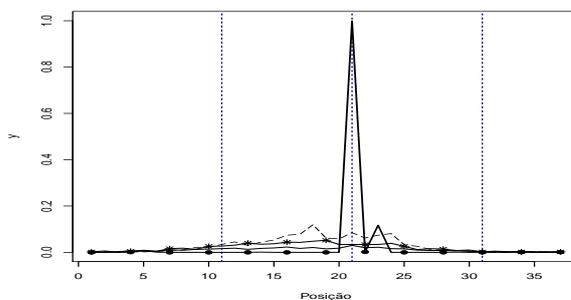
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$

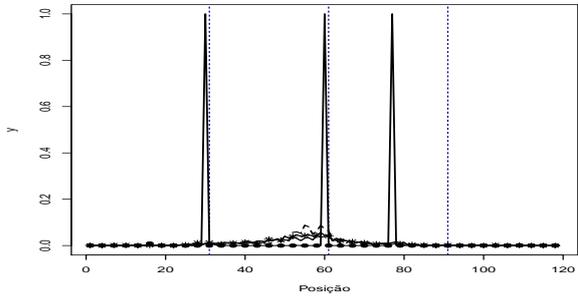


(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$

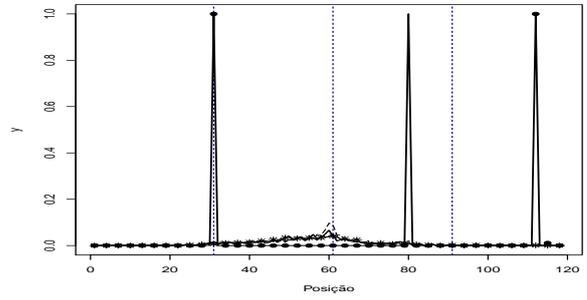


(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

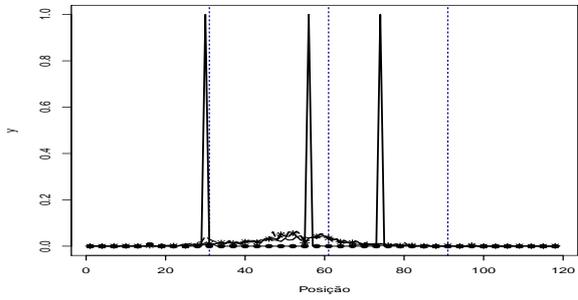
Figura D.65: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 30$ .



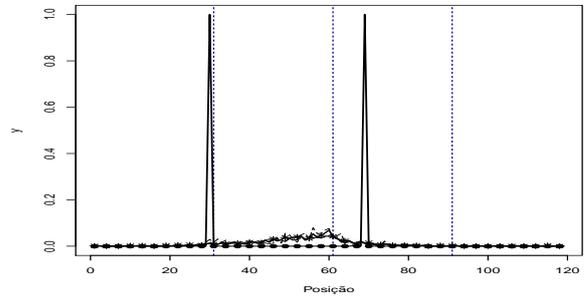
(a) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \leq 0$



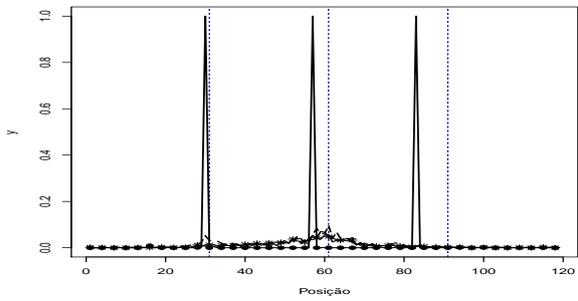
(b) Mod. gerado  $\lambda = 0$  - Mod. ajust.  $\lambda \geq 0$



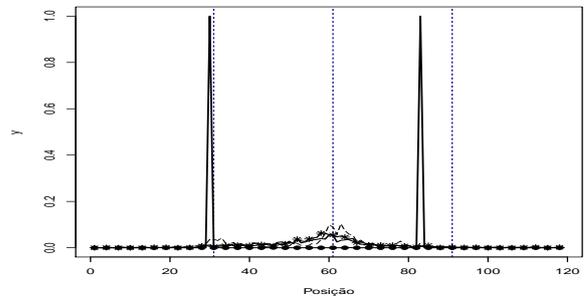
(c) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \leq 0$



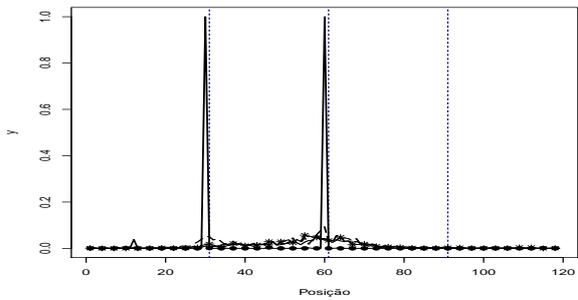
(d) Mod. gerado  $\lambda = 0.7$  - Mod. ajust.  $\lambda \geq 0$



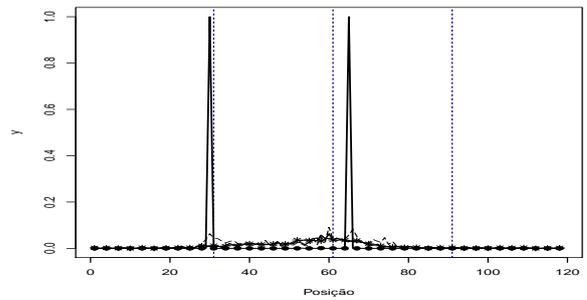
(e) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \leq 0$



(f) Mod. gerado  $\lambda = 2$  - Mod. ajust.  $\lambda \geq 0$



(g) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \leq 0$



(h) Mod. gerado  $\lambda = 100$  - Mod. ajust.  $\lambda \geq 0$

Figura D.66: Estimativa *a posteriori* da probabilidade de cada instante ser um ponto de mudança para  $\lambda = \pm 100$  (linha com  $\bullet$ ),  $\lambda = \pm 2$  (linha tracejada),  $\lambda = \pm 0.7$  (linha com  $*$ ) e  $\lambda = 0$  (linha cheia), cenário tendência,  $n = 90$ .

---

## Referências Bibliográficas

---

- Arellano-Valle, R. B., Castro, L. M. & Loschi, R. L. (2013), ‘Change point detection in the skew-normal model parameters’, *Communications in Statistics-Theory and Methods* **42**(4), 603–618.
- Arnold, B. C. & Beaver, R. J. (2002), ‘Skewed multivariate models related to hidden truncation and/or selective reporting’, *Test* **11**(1), 7–54.
- Azzalini, A. (1985), ‘A class of distribution which includes the normal ones’, *Scandinavian Journal of Statistics* **12**, 171–178.
- Azzalini, A. & Capitanio, A. (2003), ‘Distributions generated by perturbation of symmetry emphasis on a multivariate skew-t distribution’, *Journal of the Royal Statistical Society: Series B* **65**(2), 367–389.
- 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.
- Berger, J. O., Bernardo, J., David, A. & Smith, A. (1996), *Bayesian Statistics*, Oxford University Press, London.

- Berger, J. O., Insua, D. R. & Ruggeri, F. (2000), *Bayesian Robustness*, 1 edn, Springer-Verlag, Nova York, U.S.A.
- Berger, J. O., Moreno, E., Pericchi, L. R., Bayarri, M. J., Bernardo, J. M., Cano, J. A., la Horra, J. D., Martín, J., Ríos-Insúa, D., Betrò, B., Dasgupta, A., Gustafson, P., Wasserman, L., Kadane, J. B., Srinivasan, C., Lavine, M., O'Hagan, A., Polasek, W., Robert, C. P., Goutis, C., Ruggeri, F., Salinetti, G. & Sivaganesan, S. (1994), 'An overview of robust bayesian analysis', *Test* **3**(1), 5–124.
- BM&FBOVESPA (2014), Metodologia do Índice bovespa, Technical report, BM&FBOVESPA.
- Carota, C. & Ruggeri, F. (1994), 'Robust bayesian analysis given priors on partition sets', *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research* **3**(2), 73–86.
- Cifarelli, D. M. & Regazzini, E. (1987), 'On a general definition of concentration function', *The Indian Journal of Statistics, Series B (1960-2002)* **49**(3), 307–319.
- Crowley, E. M. (1997), 'Product partition models for normal means', *Journal of the American Statistical Association* **92**(437), 192–198.
- de Godoi, L. G. (2011), Análise bayesiana de sensibilidade sob distribuições *a priori* assimétricas, PhD thesis, IME-USP, São Paulo - Brasil.
- de Godoi, L. G. & Branco, M. D. (2014), 'Bayesian robustness under a skew-normal class of prior distribution', *International Journal of Approximate Reasoning* **55**(5), 1235–1251.
- Demarqui, F. N., Loschi, R. H. & Colosino, E. A. (2008), 'Estimating the grid of time-points for the piecewise exponential model', *Lifetime Data Analysis* **14**(3), 333–356.
- Demarqui, F. N., Loschi, R. H., Dey, D. K. & Colosino, E. A. (2012), 'A class of dynamic piecewise exponential models with random time grid', *Journal of Statistical Planning and Inference* **142**(3), 728–742.
- Fearnhead, P. (2006), 'Exact and efficient bayesian inference for multiple changepoint problems', *Statistics and Computing* **16**(2), 203–213.

- Ferreira, J. A., Loschi, R. H. & Costa, M. A. (2010), Product partition models with across cluster correlation.
- Fortini, S. & Ruggeri, F. (1995), ‘Concentration function and sensitivity to the prior’, *Journal of the Italian Statistical Society* **4**(3), 283–297.
- 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.
- Genton, M. G. (2004), *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Chapman and Hall, Boca Raton, FL.
- Green, P. J. (1995), ‘Reversible jump markov chain monte carlo computation and bayesian model determination’, *Biometrika* **82**(4), 711–732.
- Gupta, A. K., González-Farías, G. & Domínguez-Molina, J. (2004), ‘A multivariate skew normal distribution’, *Journal of Multivariate Analysis* **89**(1), 181–190.
- Gustafson, P. & Wasserman, L. (1995), ‘Local sensitivity diagnostics for bayesian inference’, *Annals of Statistics* **23**(6), 2153–2167.
- 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.
- Henze, N. (1986), ‘A probabilistic representation of the ‘skew-normal’ distribution’, *Scandinavian Journal of Statistics* **13**(4), 271–275.
- Loschi, R. H. & Cruz, F. (2002), ‘Applying the product partition model to the identification of multiple change points’, *Advances in Complex Systems* **5**(4), 371–387.
- 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.
- Loschi, R. H., Pontel, J. G. & Cruz, F. R. B. (2010), ‘Multiple change-point analysis for linear regression models’, *Chilean Journal of Statistics* **1**(2), 93–112.
- Marchenko, Y. V. & Genton, M. G. (2009), ‘Multivariate log-skew-elliptical distributions with applications to precipitation data’, *Environmetrics* **21**, 318–340.
- Monteiro, J. V. D., Loschi, R. H. & Assunção, R. (2011), ‘Product partition models with correlated parameters’, *Bayesian Analysis* **6**, 691–726.
- 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. (2012), *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>.
- Ruggeri, F., Insua, D. R. & Martín, J. (2005), *Robust Bayesian Analysis*, Vol. 25 of *Handbook of Statistics*, 1 edn, Elsevier North-Holland, amsterdam.
- Ruggeri, F. & Sivaganesan, S. (2000), ‘On a global robustness diagnostic for bayesian inference’, *Sankhya: The Indian Journal oof Statistics* **62**(1), 110–127. Serie A.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2002), ‘Bayesian measures of model complexity and fit’, *Journal of the Royal Statistical Society, Series B* **64**(4), 583–639.
- Tarantola, C., Consoni, G. & Dallaportas, P. (2008), ‘Bayesian clustering for row effects models’, *Journal of Statistical Planning and Inference* **138**, 2223–2235.
- van de Linde, A. (2007), ‘Local influence on posterior distributions under multiplicative modes of perturbation’, *Baysian Analysis* **2**(2), 319–332.

Yao, Y. (1984), 'Estimation of a noisy discrete-time step function: Bayes and empirical bayes approaches', *The Annal of Statistics* **12**(4), 1434–1447.

Zantedeschi, D., Damien, P. & Polson, N. G. (2011), 'Predictive macro-finance with dynamic partition models', *Journal of the American Statistical Association* **106**(494), 427–439.