

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE ESTATÍSTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

TEORIA DA INFORMAÇÃO APLICADA A
MODELOS ARMA: TESTES PARA IDENTIFICAÇÃO
E QUANTIFICAÇÃO DE OVERFITTING

PEDRO MONTEIRO DE ALMEIDA JUNIOR

DISSERTAÇÃO DE MESTRADO



Recife
2016

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE ESTATÍSTICA

PEDRO MONTEIRO DE ALMEIDA JUNIOR

TEORIA DA INFORMAÇÃO APLICADA A
MODELOS ARMA: TESTES PARA IDENTIFICAÇÃO
E QUANTIFICAÇÃO DE OVERFITTING

ORIENTADOR: PROF. DR. ABRAÃO D.C. NASCIMENTO

Área de Concentração: Séries Temporais

Dissertação apresentada ao programa de Pós-graduação em Estatística da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de **Mestre em Estatística**.

Recife
2016

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

A447t Almeida Junior, Pedro Monteiro de.
Teoria da informação aplicada a modelos Arma: testes para identificação e quantificação de *Overfitting* / Pedro Monteiro de Almeida Junior. – 2016.
173 f.: il., fig., tab.

Orientador: Abraão D. C. Nascimento.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN. Estatística, Recife, 2016.
Inclui referências e apêndices.

1. Estatística aplicada. 2. Teoria da informação - Divergência. 3. Processamento de imagens. I. Nascimento, Abraão D. C. (orientador). II. Título.

310

CDD (23. ed.)

UFPE- MEI 2016-043

PEDRO MONTEIRO DE ALMEIDA JUNIOR

**TEORIA DA INFORMAÇÃO APLICADA A MODELOS ARMA: TESTES PARA
IDENTIFICAÇÃO E QUANTIFICAÇÃO DE OVERFITTING**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Estatística.

Aprovada em: 26 de fevereiro de 2016.

BANCA EXAMINADORA

Prof. Dr. Abraão David Costa do Nascimento
UFPE

Prof. PhD. Getúlio José Amorim do Amaral (Examinador Interno)
UFPE

Prof. Dr. Eufrásio de Andrade Lima Neto (Examinador Externo)
UFPB

Aos meus pais e a minha irmã.

Agradecimentos

Primeiramente a Deus, o criador e sustentador da humanidade, pela fé concedida, força nos momentos difíceis e por tudo que me proporcionou na vida.

Aos meus pais, Lucinete e Pedro, pelo amor incondicional, pelos ensinamentos sobre a vida, incentivo e apoio nos momentos que foram precisos.

Aos meus queridos irmãos, Juliana e Felipe, pela amizade incondicional.

Ao meu orientador, Abraão Nascimento, por toda orientação, no qual nunca faltaram palavras de encorajamento e incentivo. Por toda paciência e dedicação no ensino de tópicos não triviais que foram fundamentais na construção deste trabalho. Também pelos conselhos e direcionamentos recebidos durante este período, sempre muito solícito e disposto a ajudar até fora do escopo da academia. Além de ser um profissional muito competente, sempre transmitindo muita seriedade naquilo que faz.

Aos Professores do Departamento de estatística da UFPB que contribuíram para minha formação de Bacharel em estatística, em especial aos professores: Eufrásio de Lima, Tatiane de Souza, Tarciana Liberal, Joab de Lima e Hemílio Coelho, por todo incentivo e ensinamentos compartilhados nas disciplinas e sob orientação.

Aos meus colegas de mestrado: Rodrigo Gonçalves, Rodrigo Alves, Telma, Jonas, Terezinha, Fernanda, Renata, Hildemar, Stênio, Wanessa, Laura, e Raphaela, por todos momentos de descontração, união e companheirismo que foi construída ao longo do curso.

Aos grandes amigos conquistados durante minha estadia em Recife no curso de mestrado: Hildemar, Jonas, Rodrigo e Telma, por todos bons e maus momentos vivenciados e compartilhados, levarei sempre a amizade de vocês.

Aos meus amigos, Camila, Ramon, Jodavid e Luana, por toda vivência compartilhada na "República Paraibana", em que se tornou uma verdadeira família durante essa temporada, amenizando um pouco a saudade de casa.

A todos os Professores do DE-UPPE, pelos ensinamentos que contribuíram para minha formação acadêmica e científica, em especial aos professores Cribari Neto, Gauss Cordeiro, Leandro Chaves, Abraão Nascimento e Audrey.

À Valéria, secretária do programa, sempre muito atenciosa e solicita para me orientar sobre as burocracias acadêmicas.

À CAPES, pelo apoio financeiro. E, finalmente, aos membros da banca examinadora pelas contribuições.

"Tente uma, duas, três vezes e se possível tente a quarta, a quinta e quantas vezes for necessário. Só não desista nas primeiras tentativas, a persistência é amiga da conquista. Se você quer chegar a onde a maioria não chega, faça o que a maioria não faz".

Bill Gates

Resumo

Dados registrados sequencialmente sobre um determinado período são frequentes em várias aplicações; tais como, em reconhecimento padrão (ISHI et al., 1998) e em processamento de imagens de radar (DUTRA, 1989). Diferentemente de contextos suportados pela independência das observações, dados desta natureza impõem uma estrutura de correlação que deve ser considerada nos modelos. Apresentamos inicialmente um estudo de desempenho via simulação de Monte Carlo para três métodos de estimação na classe de modelos ARMA: Máxima Verossimilhança Condicional, Não Condicional e Mínimos Quadrados Não Condicional. Vários estudos têm sido feitos neste sentido (vê, por exemplo, YU, 2004). Entretanto, faltam estudos que trabalhem sobre uma discretização razoável da região de estabilidade e que avaliem a influência da estimação sobre densidades ajustadas dos processos ARMA. Para vários pontos paramétricos da região de estabilidade, os três métodos são comparados à luz de medidas da Teoria da Informação, a saber as divergências de Kullback-Leibler, Jensen-Shannon, Battacharyya e Hellinger. Adicionalmente, as medidas utilizadas para quantificação do desempenho dos métodos de estimação são entendidas como elementos na classe (h, ϕ) de divergências proposta por Salicrú et al. (1994) e cinco novos testes de hipótese bilaterais (caso em que não há testes *uniformemente mais poderosos*) para identificação de modelos ARMA são propostos e comparados com o teste da razão entre verossimilhanças. Além disso, vários trabalhos têm avançado quanto à proposta de critérios para seleção de modelos na classe ARMA. McQuarrie e Tsai (1998) utilizaram a probabilidade de *overfitting* (isto é, escolher o modelo com ordem maior dado que o modelo verdadeiro tem ordem menor) como critério de comparação. A ideia de quantificar o erro da escolha entre modelos de probabilidade também foi estudada por Shannon (1948) que propôs os conceitos de informação, entropia e divergência. Nesta dissertação, propomos novos critérios para seleção de modelos em séries temporais baseados em medidas de distância. O desempenho dos novos critérios quantificados em termos da probabilidade de *overfitting* é comparado com outros métodos da literatura. Finalmente, as medidas derivadas são utilizadas como classificadores em imagens PolSAR.

Palavras-chave: *Divergência. Processo ARMA. Critérios de Seleção. Teoria da Informação. Imagens PolSAR*

Abstract

Data recorded sequentially on a given period are common on several applications; such as in pattern recognition (ISHI et al., 1998) and in radar image processing (DUTRA, 1989). Unlike contexts supported by independence assumption, data of this nature impose correlation patterns that should be considered in used models. In this dissertation, we first present a performance study of three estimation methods in the ARMA class: Conditional and Non-Conditional Maximum Likelihood and Least Squares. Several works have addressed the quantification of the performance on ARMA estimation methods. However, they neither present studies on several points of stability regions nor consider the effect of estimation procedures on ARMA process densities. These methods are compared by means of the Information Theory measures, say Kullback-Leibler, Jensen-Shannon, Battacharyya and Hellinger divergences. Further, the measures used to quantify the performance of the estimation methods are understood within the class (h, ϕ) -divergences proposed by Salicrú et al. (1994) and five new bilateral hypothesis tests (case on which there are not *uniformly most powerful* tests) for ARMA identification are proposed and compared with the likelihood ratio test. McQuarrie e Tsai (1998) considered the *overfitting* (to study and the performance of those methods) probability as a comparison criterion. The idea of quantifying the error of the model choose was introduced by Shannon (1948) who proposed the concepts of information, entropy and divergence. In this dissertation, we propose new selection criteria for time series models. The proposals are formulated taken into account distance measures between probability distributions. The performance of new criteria are quantified in terms of the *overfitting* probability and compared to those due to literature methods. Finally, the derived measures are employed as classifiers on PolSAR images.

Keywords: *Divergence. ARMA Process. Selection Criteria. Information Theory. Pol-SAR Images*

Lista de Figuras

4.1	Estimação dos parâmetros ϕ e σ^2 para AR(1)	56
4.2	EQM dos estimadores de ϕ e σ^2 para AR(1)	57
4.3	Média da divergência de Δ_{KL} para AR(1)	61
4.4	Média da divergência de Δ_{TKL} para um AR(1)	62
4.5	Média da divergência de Δ_{JS} para um AR(1)	63
4.6	Média da divergência de Δ_B para um AR(1)	64
4.7	Média da divergência de Δ_H para um AR(1)	65
4.8	Estimação dos parâmetros θ e σ^2 para MA(1)	67
4.9	EQM da estimação dos parâmetros θ e σ^2 para MA(1)	68
4.10	Média da divergência de Δ_{KL} para MA(1)	70
4.11	Média da divergência de Δ_{TKL} para um MA(1)	71
4.12	Média da divergência de Δ_{JS} para um MA(1)	72
4.13	Média da divergência de Δ_B para um MA(1)	73
4.14	Média da divergência de Δ_H para um MA(1)	74
4.15	Desempenho da estimação do parâmetro ϕ para ARMA(1,1)	76
4.16	Desempenho da estimação do parâmetro θ para ARMA(1,1)	77
4.17	Desempenho da estimação do parâmetro σ^2 para ARMA(1,1)	77
4.18	Desempenho do EQM da estimação do parâmetro ϕ para ARMA(1,1)	78
4.19	Desempenho do EQM da estimação do parâmetro θ para ARMA(1,1)	78
4.20	Desempenho do EQM da estimação do parâmetro σ^2 para ARMA(1,1)	79
4.21	Melhor método de estimação via Δ_B na forma reduzida	80
4.22	Melhor método de estimação via Δ_B na forma completa	80
4.23	Melhor método de estimação via Δ_{TKL} na forma reduzida	81
4.24	Melhor método de estimação via Δ_{TKL} na forma completa	81
4.25	Melhor método de estimação via Δ_{KL} na forma reduzida	82
4.26	Melhor método de estimação via Δ_{KL} na forma compelta	82

4.27	Melhor método de estimação via Δ_{JS} na forma reduzida	83
4.28	Melhor método de estimação via Δ_{JS} na forma completa	83
4.29	Poder do teste via EMNC para AR(1) com $n=35$	88
4.30	Poder do teste via EMNC para AR(1) com $n=100$	89
4.31	Poder do teste via EMNC para MA(1) com $n=35$	99
4.32	Poder do teste via EMNC para AR(1) com $n=100$	100
4.33	Poder do teste via EMNC para ARMA(1,1) com $n=35$	108
4.34	Poder do teste via EMNC para ARMA(1,1) com $n=150$	109
4.35	Cenário 1: Probabilidades de Overfitting dos critérios clássicos(Amostras pequenas)	124
4.36	Cenário 1: Probabilidades de Overfitting dos critérios clássicos(Amostras grandes)	125
4.37	Cenário 1: Probabilidades de Overfitting dos critérios propostos(Amostras pequenas)	126
4.38	Cenário 1: Probabilidades de Overfitting dos critérios propostos(Amostras grandes)	127
4.39	Cenário 1: Comparação das Probabilidades de Overfitting dos critérios (Amostras pequenas)	130
4.40	Cenário 1: Comparação das Probabilidades de Overfitting dos critérios (Amostras grandes)	131
4.41	Cenário 2: Probabilidades de Overfitting dos critérios clássicos(Amostras pequenas)	133
4.42	Cenário 2: Probabilidades de Overfitting dos critérios clássicos(Amostras grandes)	134
4.43	Cenário 2: Probabilidades de Overfitting dos critérios propostos (Amostras pequenas)	136
4.44	Cenário 2: Probabilidades de Overfitting dos critérios propostos(Amostras grandes)	137
4.45	Cenário 2: Comparação das Probabilidades de Overfitting dos critérios (Amostras pequenas)	139
4.46	Cenário 2: Comparação das Probabilidades de Overfitting dos critérios (Amostras grandes)	140
5.1	Geometria de obtenção da imagem SAR	146
5.2	Imagem real PolSAR	147
5.3	Séries transformadas e suas FAC e FACP para região de oceano.	150
5.4	Séries transformadas e suas FAC e FACP para região de floresta.	151

5.5	Séries transformadas e suas FAC e FACP para região de áreas urbanas.	. 152
5.6	Séries originais e suas FAC e FACP para região de áreas urbanas.	 153

Lista de Tabelas

2.1	Log-verossimilhanças e suas respectivas somas quadráticas do processo $AR(1)$	33
2.2	Log-verossimilhanças e suas respectivas somas quadráticas do processo $AR(p)$	33
2.3	Log-verossimilhanças e suas respectivas somas quadráticas do processo $MA(1)$	36
2.4	Log verossimilhança e Mínimos quadrados do processo $ARMA(p, q)$	37
3.1	Divergências (h, ϕ) e h e ϕ associadas	44
4.1	Estimativas dos parâmetros para $ARMA(1, 1)$ com $n = 15$	76
4.2	Estimativas dos parâmetros para $ARMA(1, 1)$ com $n = 100$	79
4.3	Poder do teste via medidas de divergências e RV para $AR(1)$ com $\phi = 0, 1$	90
4.4	Poder do teste via medidas de divergências e RV para $AR(1)$ com $\phi = 0, 5$	91
4.5	Poder do teste via medidas de divergências e RV para $AR(1)$ com $\phi = 0, 8$	92
4.6	Poder do teste via medidas de divergências e RV para $MA(1)$ com $\theta = 0.1$	96
4.7	Poder do teste via medidas de divergências e RV para $MA(1)$ com $\theta = 0.5$	97
4.8	Poder do teste via medidas de divergências e RV para $MA(1)$ com $\theta = 0.8$	98
4.9	Tamanhos empíricos para medidas de divergência e RV	101
4.10	Cenário a: Poder dos testes via medidas de divergências e RV para $ARMA(1, 1)$ com $\phi = \theta = 0.1$	105
4.11	Cenário a: Poder dos testes via medidas de divergências e RV para $ARMA(1, 1)$ com $\phi = \theta = 0.5$	106
4.12	Cenário a: Poder dos testes via medidas de divergências e RV para $ARMA(1, 1)$ com $\phi = \theta = 0.8$	107
4.13	Cenário b: Poder dos testes via medidas de divergências e RV para $ARMA(1, 1)$ com $\phi = 0.1 + \Delta, \theta = 0.1$	113
4.14	Cenário b: Poder dos testes via medidas de divergências e RV para $ARMA(1, 1)$ com $\phi = 0.5 + \Delta, \theta = 0.5$	114
4.15	Cenário b: Poder dos testes via medidas de divergências e RV para $ARMA(1, 1)$ com $\phi = 0.8 + \Delta, \theta = 0.8$	115

4.16	Cenário c: Poder dos testes via medidas de divergências e RV para ARMA(1,1) com $\phi = 0.1, \theta = 0.1 + \Delta$	119
4.17	Cenário c: Poder dos testes via medidas de divergências e RV para ARMA(1,1) com $\phi = 0.5, \theta = 0.5 + \Delta$	120
4.18	Cenário c: Poder dos testes via medidas de divergências e RV para ARMA(1,1) com $\phi = 0.8, \theta = 0.8 + \Delta$	121
4.19	Cenário 1: Probabilidade de Overfitting dos critérios clássicos	141
4.20	Cenário 2: Probabilidade de Overfitting dos critérios clássicos	141
4.21	Cenário 1: Probabilidade de Overfitting dos critérios propostos	142
4.22	Cenário 2: Probabilidade de Overfitting dos critérios propostos	142
5.1	P -valores dos testes de normalidade nos dados PolSAR.	149
5.2	Estimação dos parâmetros por MVC em séries PolSAR	154
5.3	Estimação dos parâmetros por MVNC em séries PolSAR	155
5.4	P -valores dos testes de <i>Shapiro-Wilker</i> e de <i>Ljung e Box</i> sobre os resíduos	155
5.5	Tamanhos dos testes para séries SAR assumindo estimação MVC	156
5.6	Tamanhos dos testes para séries SAR assumindo estimação MVNC	157
5.7	Poder dos testes de hipóteses para séries SAR assumindo estimação MVC	157
5.8	Poder dos testes de hipóteses para séries SAR assumindo estimação MVNC	157

Lista de abreviaturas e siglas

<i>MVC</i>	<i>Máxima Verossimilhança Condicional</i>
<i>MVNC</i>	<i>Máxima Verossimilhança Não Condicional</i>
<i>MQNC</i>	<i>Mínimos Quadrados Não Condicional</i>
<i>RVC</i>	<i>Razão de Verossimilhança Condicional</i>
<i>RVNC</i>	<i>Razão de Verossimilhança Não Condicional</i>
<i>RVC</i>	<i>Razão de Verossimilhança Condicional</i>
<i>RVNC</i>	<i>Razão de Verossimilhança Não Condicional</i>
<i>HH</i>	<i>Polarização dos pulsos emitidos na direção Horizontal(H) e recebido na direção Horizontal(H)</i>
<i>HV</i>	<i>Polarização dos pulsos emitidos na direção Horizontal(H) e recebido na direção Vertical(V)</i>
<i>VV</i>	<i>Polarização dos pulsos emitidos na direção Vertical(V) e recebido na direção Vertical(V)</i>

Lista de símbolos

ℓ_C	<i>Log verossimilhança condicional</i>
ℓ_{NC}	<i>Log verossimilhança não condicional</i>
Δ_{KL}	<i>Divergência de Kullback – Leibler</i>
Δ_{TKL}	<i>Divergência Tipo Kullback – Leibler</i>
Δ_{JS}	<i>Distância de Jensen – Shannon</i>
Δ_B	<i>Distância de Bhattacharyya</i>
Δ_H	<i>Distância de Hellinger</i>
$\pi(\cdot)$	<i>Função poder do teste em consideração</i>
S_{KL}	<i>Estatística do teste baseado na Divergência de Kullback – Leibler</i>
S_{TKL}	<i>Estatística do teste baseado na Divergência Tipo Kullback – Leibler</i>
S_{JS}	<i>Estatística do teste baseado na distância de Jensen – Shannon</i>
S_B	<i>Estatística do teste baseado na distância de Bhattacharyya</i>
S_H	<i>Estatística do teste baseado na distância de Hellinger</i>

Sumário

1	Introdução	19
1.1	Considerações iniciais	19
1.2	Detalhamento das contribuições	21
1.2.1	Primeira contribuição	21
1.2.2	Segunda contribuição	21
1.2.3	Terceira contribuição	21
1.2.4	Quarta contribuição	22
1.3	Plataforma computacional e organização do trabalho	22
2	Processo ARMA	24
2.1	Conceitos preliminares	24
2.2	Estimação para modelos ARMA	27
2.2.1	Estimações preliminares	27
2.2.2	Máxima verossimilhança e mínimos quadrados	30
2.3	Critérios clássicos para identificação de modelos	37
3	Critérios de Informação em Modelos ARMA	42
3.1	Classe (h, ϕ) -divergência	44
3.1.1	Divergências utilizadas	44
3.1.2	Resultados assintóticos	46
3.2	Contribuições teóricas	47
3.2.1	Divergências entre processos $AR(p)$	47
3.2.2	Divergências entre processos $MA(1)$ e $ARMA(1,1)$	51
4	Resultados e discussões	53
4.1	Avaliação dos métodos de estimação clássicos	53
4.1.1	Estimação no modelo $AR(1)$	55

4.1.2	Estimação no modelo MA(1)	66
4.1.3	Estimação no modelo ARMA(1,1)	75
4.2	Testes para identificação de modelos	84
4.2.1	Testes de identificação para classe de modelo AR(1)	84
4.2.2	Testes de identificação para classe de modelo MA(1)	93
4.2.3	Identificação nos modelos ARMA(1,1)	101
4.3	Quantificação de <i>Overfitting</i>	122
4.3.1	<i>Overfitting</i> sobre o Cenário 1	122
4.3.2	<i>Overfitting</i> sobre o Cenário 2	130
4.4	Conclusão dos estudos computacionais	143
5	Classificadores baseados para dados PolSAR	145
5.1	Radar de abertura sintética e modelo espacial para imagem SAR	145
5.2	Esquema de classificação de imagem SAR	147
5.3	Aplicação a imagem AIRSAR de São Francisco	148
6	Considerações finais	159
	Referências	161
A	Derivadas da Δ_{KL} para o modelo ARMA	166
B	Derivadas da Δ_B para o modelo ARMA	170

Capítulo 1

Introdução

1.1 Considerações iniciais

Dados registrados sequencialmente sobre um determinado período são frequentes em várias aplicações; tais como, em reconhecimento padrão (ISHI et al., 1998) e em processamento de imagens de radar (DUTRA, 1989). Esses dados são chamados de *Séries Temporais* (ST). A *análise de séries temporais* consiste em um conjunto de ferramentas para trabalhar com dados indexado por um suporte (tempo, posição espacial, ...) ordenado. Nesta área, tem-se basicamente um duplo fim: modelar a estrutura de correlação inerente aos dados e prever de valores futuros. Nessa dissertação, consideramos a classe de modelos ARMA (*Autoregressive Moving-Average-ARMA*), que tem se mostrado uma alternativa eficiente para descrever fenômenos correlacionados complexos, como discutido por (DOOB, 1953, p. 495-506). Na prática, o uso de modelos ARMA pode ser sistematizado na realização de três estágios: (i) Identificação da ordem do modelo ARMA, (ii) Estimação dos coeficientes ARMA e (iii) Análise de resíduos e diagnóstico. Nesta dissertação, apresentamos avanços quanto às duas primeiras etapas com base em medidas da teoria da informação.

A ideia de usar ferramentas da Teoria Estatística da Informação – tais como medidas de entropias e divergências – em Inferência Estatística é relativamente nova, datando do final dos anos '40. A origem da Teoria da Informação se deu com Shannon (1948), motivado pelo problema de transmitir informações de maneira eficiente de um canal de comunicação com efeito do ruído. Verdú (1998) define a Teoria da Informação como: “

“ *Uma teoria unificadora com interseções entre Probabilidade, Estatística, Ciência da Computação e outros campos. A Teoria da Informação continua a definir a base para o desenvolvimento das comunicações, armazenamentos e processamentos de dados e outras tecnologias de informação.*”

Dada esta unificação, trabalhos têm objetivado responder à questão “Qual a relação entre Teoria da Informação e Estatística?” De acordo com Lindley (1956) referenciando o trabalho

de Shannon (1948), deve-se entender que (i) *informação* é um conceito estatístico e (ii) Shannon (1948) propôs uma função única para quantificar informação em uma distribuição de probabilidade. Este conceito foi também introduzido por outros pesquisadores. De acordo com Kullback (1966) a natureza estatística e matemática de teoria da informação foi enfatizada por três homens, Fisher (1956), Shannon (1948) e Wiener (1956). O conceito de informação foi estendido na forma da medida de *entropia*, que pode ser entendida como a quantidade média de informação fornecida por um sistema estocástico. Com base no conceito de informação, Kullback e Leibler (1951) introduziu uma ferramenta chamada de medida de *divergência*, como uma generalização da entropia de *Shannon*, denominada divergência de *Kullback-Leibler*. A partir das propostas da entropia de Shannon e divergência de Kullback-Leibler, vários trabalhos têm endereçado a proposta de classes de estatísticas de teste com base em entropias e divergências. Nesta dissertação, dar-se-á ênfase à aplicação da classe (h, ϕ) de divergências proposta por Menéndez et al. (1995) aos processos ARMA. Estes métodos se constituem alternativas às ferramentas clássicas, tal como à estatística da razão entre verossimilhanças.

Vários trabalhos têm avançado quanto ao segundo estágio (MAURICIO, 2002; MCLEOD; ZHANG, 2008; YU, 2004) apresentado no primeiro parágrafo desta introdução. Os principais métodos de estimação para a classe de modelo ARMA são baseados na Máxima Verossimilhança não Condicional (POLLOCK, 1999; HAMILTON, 1994), na Máxima Verossimilhança Condicional e em Mínimos Quadrados (POLLOCK, 1999, Cap. 21). Em geral, esses três métodos não apresentam expressões fechadas para os estimadores. Assim, métodos de otimização não linear devem ser utilizados; tais como *Broyden-Fletcher-Goldfarb-Shanno* (BFGS), *Newton-Raphson* (NR), *Nelder-Mead* (NM), entre outros. É conhecido que estes métodos são fortemente influenciados por atribuição dos valores iniciais. Na literatura de processos ARMA, as estimações preliminares dos parâmetros são comumente obtidas empregando o algoritmo de Yule-Walker (BROCKWELL; DAVIS, 2001, p. 139 – 142) para modelos autorregressivos $AR(p)$ e os algoritmos de Inovações (BROCKWELL; DAVIS, 2001, p. 150 – 156) ou Hannan e Rissanen (1982) para processos de médias móveis $MA(q)$ e $ARMA(p, q)$. Nesta dissertação, procuramos fornecer uma resposta à questão “*Qual dos três métodos de estimação é mais vantajoso para uma dada sub-região paramétrica em termos de viés, erro quadrático médio (EQM) e da distorção da verossimilhança conjunta via medidas de divergências?*”

Uma das principais questões em modelos de séries temporais é “*Quais variáveis são importantes no modelo?*” A resposta a esta pergunta permite através de um critério de seleção balancear a qualidade no ajuste e o aumento na complexidade dos modelos. Uma vez que o modelo probabilístico, diga-se $\mathcal{A}$ tem sido proposto para um experimento, os critérios de seleção de modelos dão suporte na escolha do submodelo de $\mathcal{A}$ mais adequado

aos dados. Os critérios de seleção de modelos são avaliados frequentemente sob estudos de simulação (SEGHOUANE; BEKARA, 2004; SEGHOUANE; AMARI, 2007). Nestes critérios, é fundamental identificar erros nos desempenhos dos critérios. Nesta dissertação, usamos os modelos *Autoregressives* (AR) para comparar o desempenho dos critérios de seleção usuais. Sob *consistência*—quando o modelo verdadeiro é por si só um modelo candidato—*overfitting* é definido como a escolha do modelo com variáveis extras (em séries temporais, escolher o modelo com ordem maior dado que o modelo verdadeiro tem ordem menor). Um modelo *overfitting* pode ser instável no sentido que a variabilidade de um processo com maior ordem pode levar a previsões ou predições equivocadas. Portanto, a proposta de critérios menos vulneráveis ao *overfitting* é procurada. Nesta dissertação, propomos quatro novas medidas de quantificação de *overfitting*.

1.2 Detalhamento das contribuições

1.2.1 Primeira contribuição

Uma contribuição dessa dissertação consiste em apresentar estudos comparativos entre três métodos de estimação na classe de modelos ARMA: Máxima Verossimilhança Condicional, Máxima Verossimilhança Não Condicional e Mínimos Quadrados Não Condicional. Os três métodos são comparados sob duas perspectivas: (i) medidas clássicas de viés e erro quadrático médio (EQM) e (ii) à luz de medidas da Teoria da Informação, a saber as divergências de Kullback-Leibler, tipo Kullback-Leibler, Jensen-Shannon, Battacharyya e Hellinger. Como contribuição teórica, derivamos cinco medidas de *performance* no Capítulo 3 utilizadas no item (ii). Especificamente, o capítulo 4 traz experimentos Monte Carlo realizados para avaliar os processos de estimação sobre discretizações das regiões de estabilidade dos modelos AR(1), MA(1) e ARMA(1,1).

1.2.2 Segunda contribuição

Como segunda contribuição, propomos cinco testes de identificação alternativos aos derivados da razão entre verossimilhanças (RV) para modelos ARMA. O desempenho deles é quantificado em termos de versões empíricas da função poder e comparado com o derivado dos testes da RV condicional e não condicional. Em particular, os cenários utilizados são dos modelos AR(1), MA(1) e ARMA(1,1).

1.2.3 Terceira contribuição

Outra contribuição apresentada no Capítulo 3 dessa dissertação é a proposta de novos critérios para seleção de modelos em séries temporais, formulada com base em medidas

de distâncias entre distribuições de probabilidade. Como contribuição teórica, propomos cinco critérios de seleção, discutidos no Capítulo 3. O desempenho dos novos critérios é quantificado em termos da probabilidade de *overfitting* e comparado com aquele devido a outros oito métodos da literatura discutidos no Capítulo 4. O foco desta contribuição está na comparação entre modelos autoregressivos; especificamente, AR(1), AR(2) e AR(3).

1.2.4 Quarta contribuição

As ideias das propostas teórico metodológicas dos três itens anteriores foram adaptadas a um caso do modelo ARMA espacial definido por Yao e Brockwell (2006). Alguns trabalhos têm avançado na descrição via teoria da informação de dados de radar de abertura sintética (Synthetic Aperture Radar-SAR) tais como (NASCIMENTO et al., 2013; 2012; 2010; FRERY; NASCIMENTO; CINTRA, 2013). Todos eles assumiram independência entre os modelos para os *pixels* de imagem SAR. Nesta dissertação, elaboramos classificadores para imagens SAR com base na combinação entre teoria da medida e modelo ARMA espacial.

Os resultados apresentam avanços com relação aqueles propostos por Dutra (1989).

1.3 Plataforma computacional e organização do trabalho

Este trabalho está dividido da seguinte forma. No Capítulo 2, apresentamos a classe de modelos ARMA e algumas de suas propriedades, tais como funções de autocorrelação, métodos de estimação e resultados assintóticos associados e testes de hipótese. As contribuições teóricas—os novos testes de hipóteses, medidas de *performance* e critérios de seleção—com base em medidas de divergências são fundamentadas e apresentadas no Capítulo 3. No Capítulo 4, apresentamos os estudos de simulação Monte Carlo comparando as propostas de ferramentas para processos ARMA com métodos da literatura com o mesmo fim. Uma aplicação a classificação de texturas em imagens SAR é realizada no Capítulo 5 utilizando as ferramentas de identificação ARMA sob discussão e assumindo que um *pixel* é ARMA espacial distribuído. Finalmente, no Capítulo 6, elencamos as principais contribuições desta dissertação e alguns tópicos de pesquisa futura.

Os resultados inerentes as simulações, gráficos gerados e análises estatísticas apresentadas nesta dissertação foram obtidos utilizando o ambiente de programação R em sua versão 3.0.2 para sistema operacional *Microsoft Windows*, que se encontra disponível gratuitamente através do site <http://www.R-project.org>. Este trabalho foi digitado utilizando o sistema de tipografia L^AT_EX desenvolvido por L^Amp^ort (1994). Para detalhes sobre o sistema de tipografia L^AT_EX acesse <http://www.tex.ac.uk/CTAN/latex>.

O *Software* PolSARpro foi utilizado neste trabalho para extrair as amostras para dados PolSAR(Polarimetric-SAR) aplicadas no Capítulo 5. Para mais detalhe acesse <https://earth.esa.int/web/polsarpro/home>

Capítulo 2

Inferência clássica em processos ARMA: Estimação, identificação e seleção

Nesta dissertação, assumiremos o processo ARMA para descrever conjuntos de dados associados a um suporte que segue uma ordenação. Este processo pode ser considerado como uma alternativa simples de modelar fenômenos complicados da natureza. Como suporte a esta assertiva, (DOOB, 1953, p. 495-506) tem mostrado que um processo estacionário pode ser aproximado por um processo ARMA adequado. A modelagem ARMA é usualmente feita em três estágios: (i) *Identificação do modelo* que objetiva determinar a ordem do modelo ARMA, (ii) *Estimação do modelo e testes de hipóteses* que visa determinar coeficientes ARMA a partir de uma série observada e (iii) *Análise de resíduo e diagnóstico* que se destina a examinar as suposições feitas para o modelo. Nesta dissertação, contribuições são apresentadas quanto aos estágios (i) e (ii). Além disso, algumas ferramentas propostas são aplicadas a detecção de texturas em processamento de imagens SAR.

Este capítulo será conduzido como segue: na Seção 2.1, apresentam-se conceitos importantes em séries temporais, tais como funções para caracterização de uma ST, definições de estacionaridade, modelagem ARMA, dentre outros conceitos. A Seção 2.2.1 trata de métodos de estimação preliminar, utilizados como entrada para a estimação por máxima verossimilhança e mínimos quadrados. Estes dois últimos métodos são discutidos na Seção 2.2.2. A Seção 2.3 finaliza com a apresentação dos critérios clássicos de seleção de modelos.

2.1 Conceitos preliminares

Ao lidar com um número finito de variáveis aleatórias, costuma-se computar a *matriz de covariância* associada ao modelo a fim de ganhar *insight* sobre a estrutura de dependência

entre as variáveis. Para lidar com uma série temporal $\{X_t; t \in T\}$ com média μ , necessita-se estender o conceito de matriz de covariância para tratar com conjuntos infinitos de variáveis aleatórias. Uma ferramenta importante é a função de autocovariância (fac) definida como segue:

Definição 1 (Função de autocovariância). *Seja $\{X_t; t \in T\}$ um processo com $\text{Var}(X_t) < \infty$ para todo $t \in T$. A fac $\gamma_X(\cdot, \cdot)$ de $\{X_t\}$ é definida por*

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - EX_r)(X_s - EX_s)], \quad \text{para } r, s \in T.$$

Com base na Definição 1, pode-se construir uma condição sobre a qual os processos são analiticamente mais tratáveis, chamada de *estacionaridade*.

Definição 2 (Estacionaridade). *A série temporal $\{X_t; t \in \mathbb{Z}\}$ é chamada de estacionária se*

- (i) $E|X_t| < \infty$ para todo $t \in \mathbb{Z}$;
- (ii) $EX_t = \mu$ para todo $t \in \mathbb{Z}$ e
- (iii) $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ para todo $r, s, t \in \mathbb{Z}$.

A estacionaridade é a primeira suposição adotada na dissertação. Outra hipótese importante é a estrutura do modelo. Assumiremos que as situações-problema desta dissertação são bem modeladas pela classe ARMA. O processo ARMA $\{X_t; t \in \mathbb{Z}\}$ de média μ e ordens p e q é denotado por $\dot{X}_t \triangleq \{X_t - \mu\} \sim \text{ARMA}(p, q)$ e pode ser definido como

$$\begin{aligned} \dot{X}_t - \phi_1 \dot{X}_{t-1} - \cdots - \phi_p \dot{X}_{t-p} &= Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \Leftrightarrow \\ \underbrace{\left(1 - \sum_{i=1}^p \phi_i B^i\right)}_{\phi(B)} \dot{X}_t &= \underbrace{\left(1 + \sum_{i=1}^q \theta_i B^i\right)}_{\theta(B)} Z_t \Leftrightarrow \phi(B) \dot{X}_t = \theta(B) Z_t, \end{aligned} \quad (2.1)$$

em que $B^j \{X_t\} = X_{t-j}$ para $j = 0, 1, \dots$ é o operador de defasagem, $\phi(B) = 1 - \sum_{i=1}^p \phi_i B^i$ é o polinômio autoregressivo, $\theta(B) = 1 + \sum_{i=1}^q \theta_i B^i$ é o polinômio de médias móveis, $\phi_p, \theta_q \neq 0$, $\{Z_t\} \text{RB}(0, \sigma^2)$ é o processo *ruído branco* ou *de inovações* com variância σ^2 ; isto é, um processo com média zero e de função de autocovariância $\gamma_Z(h) = \sigma^2 \mathbb{I}_0(h)$. Assumiremos ainda que o modelo (2.1) é estacionário e invertível – isto é, que as raízes de $\phi(z) = 0$ e $\theta(z) = 0$ estão fora do círculo unitário e que $\phi(z) = 0$ e $\theta(z) = 0$ não têm raízes em comum.

Um conceito importante tanto para caracterização de estabilidade quanto para derivações de propriedades de sub-modelos ARMA é o conceito de causalidade.

Definição 3 (Causalidade). *Um processo ARMA(p, q) é causal (isto é, pode ser expresso em função apenas das entradas Z_t) se existe uma sequência de constantes $\{\psi_j\}$ tal que $\sum_{j=0}^{\infty} |\psi_j| < \infty$ e*

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad t = 0, \pm 1, \pm 2, \dots \quad (2.2)$$

Derivar condições sobre as quais os processos ARMA são causal é de importância crucial para caracterização de STs estáveis ou explosivas (HAMILTON, 1994, p. 4) e/ou na formulação de estratégias de estimação restrita e testes de hipótese. Sob a satisfação da causalidade, pode-se derivar que $\{X_t; t = 0, \pm 1, \pm 2, \dots\} \sim \text{ARMA}(p, q)$ tem função de autocovariancia na defasagem h , $\gamma_X(h)$, dada por

$$\gamma_X(h) = \sigma^2 \sum_{j,k=-\infty}^{\infty} \psi_j \psi_k \mathbb{I}_0(h - j + k).$$

De um modelo geral, o seguinte teorema dá suporte a obtenção destas regiões de causalidade ou estabilidade.

Teorema 1 (Condições para Causalidade). *Seja $\{X_t\}$ um processo ARMA(p, q) para o qual $\phi(\cdot)$ e $\theta(\cdot)$ não tem raízes comum. Então $\{X_t\}$ é causal se só se $\phi(z) \neq 0$ para todo $z \in \mathbb{C}$ tal que $|z| \leq 1$. Os coeficientes ψ_j em (2.2) são determinados pela relação*

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1.$$

Do Teorema 1, as seguintes regiões de estabilidade para modelos ARMA podem ser verificadas:

- (i) O modelo AR(1), $X_t = \phi X_{t-1} + Z_t$, é estável para $\{\phi \in \mathbb{R}: -1 < \phi < 1\}$.
- (ii) O modelo AR(2), $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$, é estável para (SHUMWAY; STOFFER, 2005, p. 96)

$$\{(\phi_1, \phi_2) \in \mathbb{R}^2: |\phi_2| < 1, \phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1\}.$$

- (iii) O modelo MA(1), $X_t = Z_t + \theta Z_{t-1}$, é estável para $\{\theta \in \mathbb{R}: -1 < \theta < 1\}$.
- (iv) O modelo ARMA(1,1), $X_t = \phi X_{t-1} + \theta Z_{t-1} + Z_t$, é estável para $\{(\phi, \theta) \in \mathbb{R}^2: -1 < \phi < 1, -1 < \theta < 1\}$.

Nesta dissertação, estudos de simulação para avaliar os métodos estimação, teste de hipótese e seleção de modelos serão considerados sob discretizações das regiões de estabilidade.

2.2 Estimação para modelos ARMA

Iniciemos com uma breve discussão da evolução dos métodos de estimação em processos ARMA; especificamente dois deles: máxima verossimilhança (MV) e mínimos quadrados (MQ). Para o primeiro caso, são considerados ambos os estimadores com base nas estruturas condicionais (chamado de *Máxima Verossimilhança Condicional*–MVC) e não condicionais (chamado de *Máxima Verossimilhança Não Condicional*–MVNC) das log verossimilhanças. Adicionalmente, considera-se o estimador de mínimos quadrados não condicional (MQNC).

É conhecido que um problema auxiliar à estimação não condicional é a obtenção da inversa da matriz de covariância (Σ^{-1}). Neste sentido, vários avanços têm sido desenvolvidos na obtenção dos elementos de Σ^{-1} em processos ARMA. (CHAMPERNOWNE, 1948, p. 206) e (SIDDIQUI, 1958) apresentam a inversa exata para processos AR puros. Em modelos $MA(1)$, destacam-se as propostas de Uppuluri e Carperter (1969) e Shaman (1969), que obtiveram expressões exatas para Σ^{-1} . Estes resultados foram estendidos por Mentz (1972) e Shaman (1973) para processos $MA(q)$. Em meados dos anos 70, Tiao e Ali (1971) propuseram uma forma exata para Σ^{-1} em modelos $ARMA(1, 1)$. Posteriormente, Galbraith e Galbraith (1974) apresentam expressões para obter o determinante e matriz de covariância inversa de modelos ARMA. Nesta dissertação, utilizaremos alguns resultados de Galbraith e Galbraith (1974) para obter estimadores não condicionais para processos $AR(p)$ e $ARMA(1, 1)$ e derivação das medidas de divergência.

2.2.1 Estimações preliminares: Yule-Walker, Inovações e Hannan-Rissanen

Estimações preliminares dos parâmetros ARMA— $\phi = (\phi_1, \dots, \phi_p)^\top$, $\theta = (\theta_1, \dots, \theta_q)^\top$ e σ^2 —podem ser obtidas pelos métodos de Yule-Walker e Inovações. Estas estimativas costumam ser utilizadas como entradas para as estimações via MVNC e MQNC.

As estimativas de Yule-Walker são mais apropriadas para modelos autorregressivos $AR(p)$; enquanto que aquelas dos métodos de inovações ou Hannan-Rissanen (HANNAN; RISSANEN, 1982) são utilizadas em processos $MA(q)$ e $ARMA(p, q)$ (BROCKWELL; DAVIS, 2001).

Método de Yule-Walker:

Seja $\{X_t; t = 1, \dots, n\} \sim AR(p)$ um processo autorregressivo de ordem p com média zero causal:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + Z_t, \quad \text{e} \quad \{Z_t\} \sim RB(0, \sigma^2).$$

Considere que se objetiva estimar os coeficientes do vetor $\phi = (\phi_1, \dots, \phi_p)^\top$ e a variância ruído branco σ^2 com base nas observações $X_1, \dots, X_n$. De um algebrismo simples, tem-se

as equações de *Yule-Walker*:

$$\gamma(k) - \phi_1 \gamma(k-1) - \dots - \phi_p \gamma(k-p) = \begin{cases} 0, & k = 1, \dots, p, \\ \sigma^2, & k = 0. \end{cases}$$

As equações de *Yule-Walker* podem ser escrita como $\mathbf{\Gamma}_p \boldsymbol{\phi} = \boldsymbol{\gamma}_p$ e $\sigma^2 = \gamma(0) - \boldsymbol{\phi}^\top \boldsymbol{\gamma}_p$, em que

$$\mathbf{\Gamma}_p = \begin{pmatrix} \gamma(0) & \dots & \gamma(p-1) \\ \vdots & \ddots & \vdots \\ \gamma(p-1) & \dots & \gamma(0) \end{pmatrix} \quad e \quad \boldsymbol{\gamma}_p = \begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(p) \end{pmatrix}$$

Assim, substituindo $\mathbf{\Gamma}_p$ e $\boldsymbol{\gamma}_p$ por suas versões amostrais, obtém-se os estimadores de *Yule-Walker*: $\hat{\mathbf{\Gamma}}_p \hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\gamma}}_p$ e $\hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\boldsymbol{\phi}}^\top \hat{\boldsymbol{\gamma}}_p$, em que

$$\hat{\mathbf{\Gamma}}_p = \begin{pmatrix} \hat{\gamma}(0) & \dots & \hat{\gamma}(p-1) \\ \vdots & \ddots & \vdots \\ \hat{\gamma}(p-1) & \dots & \hat{\gamma}(0) \end{pmatrix} \quad e \quad \hat{\boldsymbol{\gamma}}_p = \begin{pmatrix} \hat{\gamma}(0) \\ \vdots \\ \hat{\gamma}(p) \end{pmatrix}.$$

Um importante resultado assintótico é dado no seguinte teorema.

Teorema 2. *Se $\{X_t\}$ é um processo $AR(p)$ causal com $\{Z_t\} \sim RB(0, \sigma^2)$ e $\hat{\boldsymbol{\phi}}$ é o estimador Yule-Walker para $\boldsymbol{\phi}$, então*

$$\sqrt{n}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} N(\mathbf{0}, \sigma^2 \mathbf{\Gamma}_p^{-1}),$$

em que “ $\xrightarrow{\mathcal{D}}$ ” denota convergência em distribuição. Além disso,

$$\hat{\sigma}^2 \xrightarrow[n \rightarrow \infty]{\mathcal{P}} \sigma^2,$$

“ $\xrightarrow{\mathcal{P}}$ ” denota convergência em probabilidade.

Algoritmo de inovações:

Seja $\{X_t; t = 1, \dots, n\} \sim MA(q)$ um processo MA(q), o modelo MA ajustado é dado por

$$X_t = Z_t + \hat{\theta}_{q1} Z_{t-1} + \hat{\theta}_{q2} Z_{t-2} + \dots + \hat{\theta}_{qq} Z_{t-q} \quad e \quad \{Z_t\} \sim RB(0, \hat{v}_q),$$

em que os vetores de coeficientes estimados $\hat{\boldsymbol{\theta}}_q = (\hat{\theta}_{q1}, \hat{\theta}_{q2}, \dots, \hat{\theta}_{qq})^\top$ e as variâncias dos ruídos branco $\hat{v}_q$, para $q = 1, 2, \dots$, são especificadas na seguinte definição.

Definição 1 (Estimativas de inovações em processos MA). *Se $\hat{\gamma}(0) > 0$, definem-se as estimativas de inovações $\hat{\boldsymbol{\theta}}_q$ e $\hat{v}_q$ para $q = 1, 2, \dots, n-1$ por relações recursivas como: $\hat{v}_0 = \hat{\gamma}(0)$,*

$$\hat{\theta}_{q,q-k} = \hat{v}_k^{-1} \left[\hat{\gamma}(q-k) - \sum_{j=0}^{k-1} \hat{\theta}_{q,q-j} \hat{\theta}_{k,k-j} \hat{v}_j \right], \quad k = 0, \dots, q-1$$

$$e \quad \hat{v}_q = \hat{\gamma}(0) - \sum_{j=0}^{q-1} \hat{\theta}_{q,q-j}^2 \hat{v}_j.$$

Este algoritmo de inovações também pode ser usado para processos ARMA causal e invertível. O seguinte teorema apresenta propriedades estatísticas assintóticas associadas às estimadores por inovações.

Teorema 3. *Seja $\{X_t\}$ um processo ARMA(p, q) causal e invertível, $\phi(B)X_t = \theta(B)Z_t$ e $\{Z_t\} \sim RB(0, \sigma^2)$, com $EX_t^4 < \infty$. Seja também $\Psi(z) = \sum_{j=0}^{\infty} \Psi_j z^j = \frac{\theta(z)}{\phi(z)}$, $|z| \neq 1$ (com $\Psi_0 = 1$ e $\Psi_j = 0$ para $j < 0$). Então para uma sequência de inteiros positivos $\{m(n), n = 1, 2, \dots\}$ tal que $m \rightarrow \infty$ e $m = o(n^{1/3})$ quando $n \rightarrow \infty$, temos para cada k*

$$\sqrt{n}(\theta_{m1} - \Psi_1, \theta_{m2} - \Psi_2, \dots, \theta_{mk} - \Psi_k)^\top \xrightarrow[n \rightarrow \infty]{\mathcal{D}} N_q(\mathbf{0}, \mathbf{A}).$$

em que $\mathbf{A} = (a_{ij})_{i,j=1,\dots,q}$ e

$$a_{ij} = \sum_{r=1}^{\min(i,j)} \Psi_{i-r} \Psi_j - r$$

e $\min(\cdot)$ representa o mínimo de um conjunto de valores. Além disso,

$$\hat{v}_q \xrightarrow[n \rightarrow \infty]{\mathcal{P}} \sigma^2.$$

Prova: Veja Brockwell e Davis (1985).

Algoritmo de Hannan-Rissanen:

Seja $\{X_t; t = 1, \dots, n\} \sim \text{ARMA}(p, q)$,

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_t - q \quad \text{e } \{Z_t\} \sim RB(0, \sigma^2).$$

Sejam $\phi = (\phi_1, \dots, \phi_p)^\top$ e $\theta = (\theta_1, \dots, \theta_q)^\top$. O algoritmo de Hannan e Rissanen (1982) pode ser definido nos seguintes passos:

Algoritmo 1: Etapas do Algoritmo de Hannan-Rissanen

Passo 1: Inicialmente, estima-se pelo método de *Yule-Walker* um modelo de alta ordem $AR(m)$ (com $m > \max(p, q)$). Assumindo $\hat{\phi}_{m1}, \dots, \hat{\phi}_{mm}$ como os coeficientes estimados, então Z_t é estimado por

$$\hat{Z}_t = X_t - \hat{\phi}_{m1} X_{t-1} - \dots - \hat{\phi}_{mm} X_{t-m}, \quad t = m+1, \dots, n.$$

Passo 2: O vetor $\beta = (\phi^\top, \theta^\top)^\top$ é estimado por regressar de X_t em termos que $X_{t-1}, \dots, X_{t-p}, \hat{Z}_{t-1}, \dots, \hat{Z}_{t-q}$, isto é, definindo $\hat{\beta}$ como o argumento de minimiza

$$S(\beta) = \sum_{t=m+1}^n (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} - \theta_1 \hat{Z}_{t-1} - \dots - \theta_q \hat{Z}_{t-q})^2.$$

Assim, o estimador *Hannan-Rissanen* pode ser dado por $\hat{\beta} = (\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{Z}^\top \mathbf{X}_n$, em que $\mathbf{X}_n = (X_{m+1}, \dots, X_n)^\top$ e

$$\mathbf{Z} = \begin{pmatrix} X_m & X_{m-1} & \dots & X_{m-p+1} & \hat{Z}_m & \hat{Z}_{m-1} & \dots & \hat{Z}_{m-p+1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ X_{n-1} & X_{n-2} & \dots & X_{n-p} & \hat{Z}_{n-1} & \hat{Z}_{n-2} & \dots & \hat{Z}_{n-q} \end{pmatrix}.$$

O estimador de *Hannan-Rissanen* da variância ruído branco σ^2 é

$$\hat{\sigma}_{HR}^2 = \frac{S(\hat{\beta})}{n - m}.$$

Os estimadores de *Hannan-Rissanen* têm a mesma eficiência assintótica dos estimadores de máxima verossimilhança (HANNAN; RISSANEN, 1982; BROCKWELL; DAVIS, 2001).

2.2.2 Máxima verossimilhança e mínimos quadrados

Sejam $x_1, x_2, \dots, x_n$ uma realização n -dimensional de $\{X_t; t = 1, 2, \dots, n\} \sim ARMA(p, q)$, a verossimilhança conjunta associada é dada por

$$L(\beta, \sigma^2) \triangleq L(\beta, \sigma^2; x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n; \phi, \theta, \sigma^2). \quad (2.3)$$

Nesta seção, discutiremos dois métodos de estimação para $(\beta^\top, \sigma^2)^\top$. Eles podem ser definidos em dois passos: (1) encontrar a função de verossimilhança da (2.3) e (2) maximizar $L(\beta, \sigma^2)$ no caso de MV ou minimizar uma soma de quadrados definida a partir de $L(\beta, \sigma^2)$ para MQ. Para este fim assumiremos que a entrada de X_t é $\{Z_t\} \sim RBG(0, \sigma^2)$.

Hamilton (1994, cap. 5) deriva algumas verossimilhanças exatas para processos ARMA gaussiano fatorando a densidade conjunta em

$$f(x_n, \dots, y_1; \beta, \sigma^2) = f(x_p, \dots, x_1; \beta, \sigma^2) \times \prod_{t=p+1}^n f(x_t | x_{t-1}, \dots, x_{t-p}; \beta, \sigma^2).$$

A função de log-verossimilhança pode então ser expressa como

$$\begin{aligned}\ell_{\text{NC}}(\boldsymbol{\beta}, \sigma^2) &= \log L(\boldsymbol{\beta}, \sigma^2; x_1, \dots, x_n) \\ &= \log f(x_p, \dots, x_1; \boldsymbol{\beta}, \sigma^2) + \sum_{t=p+1}^n \log f(x_t | x_{t-1}, \dots, x_{t-p}; \boldsymbol{\beta}, \sigma^2).\end{aligned}\quad (2.4)$$

A função $\ell_{\text{NC}}(\boldsymbol{\beta}, \sigma^2)$ em (2.4) é chamada de *log-verossimilhança exata* ou *não condicional*. A segunda parcela dessa expressão é conhecido como a log verossimilhança condicional, denotada como $\ell_{\text{C}}(\boldsymbol{\beta}, \sigma^2)$. Em modelos de séries temporais gaussianos, costuma-se considerar dois tipos de estimadores MV. O primeiro se baseia na maximização da função de verossimilhança condicional; isto é, seja $\boldsymbol{\delta} = (\boldsymbol{\beta}^\top, \sigma^2)^\top \in \Delta$ e Δ o espaço paramétrico associado,

$$\hat{\boldsymbol{\delta}}_{\text{MVC}} = \arg \max_{\boldsymbol{\delta} \in \Delta} \left[\underbrace{\sum_{t=p+1}^n \log f(x_t | x_{t-1}, \dots, x_{t-p}; \boldsymbol{\delta})}_{\triangleq \ell_{\text{C}}(\boldsymbol{\delta})} \right]$$

O segundo é definido pela maximização da função de verossimilhança não condicional:

$$\hat{\boldsymbol{\delta}}_{\text{MVNC}} = \arg \max_{\boldsymbol{\delta} \in \Delta} \left[\underbrace{\log f(x_p, \dots, x_1; \boldsymbol{\delta}) + \sum_{t=p+1}^n \log f(x_t | x_{t-1}, \dots, x_{t-p}; \boldsymbol{\delta})}_{\triangleq \ell_{\text{NC}}(\boldsymbol{\delta})} \right]$$

A estimação de MQ condicional coincide com o método MVC. Já a estimação de MQ não condicional, denotada por MQNC, pode ser definida como o argumento que minimiza a soma de quadrados não condicional-definida como $S_{\text{NC}}(\boldsymbol{\delta})$ – obtida de $\ell_{\text{NC}}(\boldsymbol{\delta})$:

$$\hat{\boldsymbol{\delta}}_{\text{MQNC}} = \arg \max_{\boldsymbol{\delta} \in \Delta} [S_{\text{NC}}(\boldsymbol{\delta})].$$

Na discussão subsequente, as log-verossimilhanças e soma de quadrados dos modelos AR(1), MA(1) e ARMA(1,1) são explicitadas.

Segundo Shumway e Stoffer (2005, p. 133), sob certas condições, os estimadores $\hat{\boldsymbol{\delta}}_{\text{MVNC}}$, $\hat{\boldsymbol{\delta}}_{\text{MVC}}$ e $\hat{\boldsymbol{\delta}}_{\text{MQNC}}$ satisfazem a seguinte convergência:

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} N_q(\mathbf{0}, \sigma^2 \boldsymbol{\Gamma}_{p,q}^{-1}),$$

em que a matriz de ordem $p + q$ $\boldsymbol{\Gamma}_{p,q}$ é a matriz informação de Fisher associada ao modelo ARMA(p,q). Nesta dissertação, checamos os desempenhos destes estimadores para tamanhos pequenos e moderados de amostra.

Verossimilhança para processos AR(p) gaussianos:

Seja $\{\dot{X}_t; t = 1 \dots, n\} \sim \text{AR}(p)$ com média $\mathbb{E}(X_t) = \mu = c(1 - \phi_1 - \dots - \phi_p)^{-1}$ e entrada $\{Z_t\} \sim \text{RBG}(0, \sigma^2)$. Neste caso, $\boldsymbol{\delta} = (c, \phi_1 \dots \phi_p, \sigma^2)^\top$. Seja também $\mathbf{x}_p = (X_1, \dots, X_p)^\top$,

sua função densidade conjunta é dada por

$$\begin{aligned} f(x_1, \dots, x_p; \boldsymbol{\delta}) &= (2\pi)^{-\frac{p}{2}} |\sigma^2 \mathbf{V}_p|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\dot{\mathbf{x}}_p - \boldsymbol{\mu}_p)^\top \sigma^{-2} \mathbf{V}_p^{-1} (\dot{\mathbf{x}}_p - \boldsymbol{\mu}_p) \right\} \\ &= (2\pi \sigma^2)^{-\frac{p}{2}} |\mathbf{V}_p^{-1}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma^2} (\dot{\mathbf{x}}_p - \boldsymbol{\mu}_p)^\top \mathbf{V}_p^{-1} (\dot{\mathbf{x}}_p - \boldsymbol{\mu}_p) \right\}, \end{aligned}$$

em que $\dot{\mathbf{x}}_p$ é uma possível realização de $\mathbf{x}_p$ e a matriz de covariâncias de $\mathbf{x}_p$ é dada por

$$\boldsymbol{\Sigma}_p = \sigma^2 \mathbf{V}_p = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 \end{bmatrix}_{p \times p}$$

Adicionalmente, tem-se que $\hat{X}_{n+1} = \mathbb{E}(X_{n+1} | X_n, X_{n-1}, \dots, X_{n-p+1}) = \sum_{i=1}^p \phi_i X_{n+1-i}$ (BROCKWELL; DAVIS, 2001). Segue-se que $(X_n | X_{n-1}, \dots, X_1) \sim N(c + \phi_1 x_{n-1} + \dots + \phi_p x_{n-p}, \sigma^2)$ e, portanto, a log-verossimilhança condicional é dada por

$$\ell_C(\boldsymbol{\delta}) = -\left(\frac{n-p}{2}\right) \log(2\pi) - \left(\frac{n-p}{2}\right) \log(\sigma^2) - \frac{1}{2\sigma^2} S_C(c, \boldsymbol{\phi}).$$

em que,

$$S_C(c, \boldsymbol{\phi}) = \sum_{t=p+1}^n (x_t - c - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p})^2.$$

A log-verossimilhança não condicional é expressa como

$$\begin{aligned} \ell_{\text{NC}}(\boldsymbol{\delta}) &= -\frac{p}{2} \log(2\pi) - \frac{p}{2} \log(\sigma^2) + \frac{1}{2} \log |\mathbf{V}_p^{-1}| - \frac{1}{2\sigma^2} (\dot{\mathbf{x}}_p - \boldsymbol{\mu}_p)^\top \mathbf{V}_p^{-1} (\dot{\mathbf{x}}_p - \boldsymbol{\mu}_p) \times \\ &\quad - \left(\frac{n-p}{2}\right) \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=p+1}^n (x_t - c - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p})^2 \\ &= -\left(\frac{n}{2}\right) \log(2\pi) - \left(\frac{n}{2}\right) \log(\sigma^2) + \frac{1}{2} \log |\mathbf{V}_p^{-1}| - \frac{1}{2\sigma^2} S_{\text{NC}}(c, \boldsymbol{\phi}). \end{aligned}$$

em que

$$S_{\text{NC}}(c, \boldsymbol{\phi}) = (\dot{\mathbf{x}}_p - \boldsymbol{\mu}_p)^\top \mathbf{V}_p^{-1} (\dot{\mathbf{x}}_p - \boldsymbol{\mu}_p) + \sum_{t=p+1}^n (x_t - c - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p})^2.$$

Note que o uso da log-verossimilhança não condicional requer a inversão da matriz $\mathbf{V}_p$. Neste sentido, denotando os elementos da matriz $\mathbf{V}_p^{-1}$ por $v_{ij}(p)$, Galbraith e Galbraith (1974, p.70) mostrou que

$$v_{ij}(p) = \left[\sum_{k=0}^{i-1} \phi_k \phi_{k+j-i} - \sum_{k=p+1-j}^{p+i-j} \phi_k \phi_{k+j-i} \right] \quad (2.5)$$

, para $1 \leq i \leq j \leq p$. Os principais resultados desta seção são organizados nas Tabelas 2.1 e 2.2.

Tabela 2.1: Log-verossimilhanças e suas respectivas somas quadráticas do processo AR(1)

Log verossimilhanças condicional e não condicional
$\ell_C(c, \phi, \sigma^2) = -\left(\frac{n-1}{2}\right) \log(2\pi) - \left(\frac{n-1}{2}\right) \log(\sigma^2) - \frac{1}{2\sigma^2} S_C(c, \phi)$ $\ell_{NC}(c, \phi, \sigma^2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) + \frac{1}{2} \log(1 - \phi^2) - \frac{1}{2\sigma^2} S_{NC}(c, \phi)$
Soma dos quadrados condicional e não condicional
$S_C(c, \phi) = \sum_{t=2}^n (x_t - c - \phi x_{t-1})^2$ $S_{NC}(c, \phi) = (1 - \phi^2) \left(x_1 - \frac{c}{1-\phi}\right)^2 + \sum_{t=2}^n (x_t - c - \phi x_{t-1})^2$

Tabela 2.2: Log-verossimilhanças e suas respectivas somas quadráticas do processo AR(p)

Log verossimilhanças condicional e não condicional
$\ell_C(c, \phi, \sigma^2) = -\left(\frac{n-p}{2}\right) \log(2\pi) - \left(\frac{n-p}{2}\right) \log(\sigma^2) - \frac{1}{2\sigma^2} S_C(c, \phi)$ $\ell_{NC}(c, \phi, \sigma^2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) + \frac{1}{2} \log(\mathbf{V}_p^{-1}) - \frac{1}{2\sigma^2} S_{NC}(c, \phi)$
Soma dos quadrados condicional e não condicional
$S_C(c, \phi) = \sum_{t=p+1}^n (x_t - c - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p})^2$ $S_{NC}(c, \phi) = (\hat{\mathbf{x}}_p - \boldsymbol{\mu})^\top \mathbf{V}_p^{-1} (\hat{\mathbf{x}}_p - \boldsymbol{\mu}) + \sum_{t=p+1}^n (x_t - c - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p})^2$

Verossimilhança para processos MA(1) gaussianos:

Seja $\{\dot{X}_t; t = 1 \dots, n\} \sim \text{MA}(1)$ com média $\mathbb{E}(X_t) = \mu$ e entrada $\{Z_t\} \sim \text{RBG}(0, \sigma^2)$. Neste caso, $\boldsymbol{\delta} = (\mu, \theta, \sigma^2)^\top$. Note também que, se os valores de Z_{t-1} são conhecidos, então

$$[X_t | Z_{t-1}] \sim N(\mu + \theta Z_{t-1}, \sigma^2),$$

Assim, assumindo $z_0 = 0$, vale-se (para $t = 1, \dots, n$)

$$f(x_t | z_{t-1}; \boldsymbol{\delta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (x_t - \mu - \theta z_{t-1})^2 \right\}.$$

Note que a sequência $\{Z_1, \dots, Z_{t-1}\}$ pode ser calculada por $Z_t = X_t - \mu - \theta Z_{t-1}$. A função de verossimilhança condicional é dada por

$$L(\boldsymbol{\delta}) = \prod_{t=2}^n f(x_t | z_{t-1}, z_0 = 0; \boldsymbol{\delta}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=2}^n z_{t-1}^2 \right\}$$

a log-verossimilhança condicional é dada por

$$\ell_C(\boldsymbol{\delta}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=2}^n Z_{t-1}.$$

Por outro lado, para o cálculo da função de verossimilhança exata ou não condicional, a fatoração triangular da matriz de covariância posso ser utilizada nesse processo. A matriz de covariância de um processo MA(1) é

$$\boldsymbol{\Sigma} = \sigma^2 \begin{bmatrix} 1 + \theta^2 & \theta & 0 & \dots & 0 & 0 \\ \theta & 1 + \theta^2 & \theta & \dots & 0 & 0 \\ 0 & \theta & 1 + \theta^2 & \dots & \theta & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \theta & 1 + \theta^2 \end{bmatrix}_{n \times n}$$

A função de verossimilhança não condicional é dada por

$$f(x_1, \dots, x_n; \boldsymbol{\delta}) = (2\pi)^{-\frac{n}{2}} |\sigma^2|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$

em que $\mathbf{x} = (x_1, \dots, x_n)$ e $\boldsymbol{\mu} = (\mu, \dots, \mu)$. Utilizando a fatoração triangular temos (HAMILTON, 1994), $\boldsymbol{\Sigma} = \mathbf{A} \mathbf{D} \mathbf{A}^\top$, em que,

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ \frac{\theta}{1+\theta^2} & 1 & 0 & \dots & 0 & 0 \\ 0 & \frac{\theta(1+\theta^2)}{1+\theta^2+\theta^4} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{\theta(1+\theta^2+\theta^4+\dots+\theta^{2(n-2)})}{1+\theta^2+\theta^4+\dots+\theta^{2(n-1)}} & 1 \end{bmatrix}_{n \times n}$$

$$\mathbf{D} = \sigma^2 \begin{bmatrix} 1 + \theta^2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \frac{1 + \theta^2 + \theta^4}{1 + \theta^2} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \frac{1 + \theta^2 + \theta^4 + \theta^6}{1 + \theta^2 + \theta^4} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \frac{1 + \theta^2 + \cdots + \theta^{2n}}{1 + \theta^2 + \cdots + \theta^{2(n-1)}} \end{bmatrix}_{n \times n}$$

Substituindo as matrizes $\mathbf{A}$ e $\mathbf{D}$, a densidade conjunta pode ser reescrita como

$$f(x_1, \dots, x_n; \boldsymbol{\delta}) = (2\pi)^{-\frac{n}{2}} |\mathbf{A} \mathbf{D} \mathbf{A}^\top|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top (\mathbf{A}^\top)^{-1} \mathbf{D}^{-1} \mathbf{A}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$

Como $\mathbf{A}$ é uma matriz triangular formada pela diagonal principal de 1's, então $|\mathbf{A}| = 1$. Definindo $\mathbf{x}_* = \mathbf{A}^{-1}(\mathbf{x} - \boldsymbol{\mu})$,

$$f(x_1, \dots, x_n; \boldsymbol{\delta}) = (2\pi)^{-\frac{n}{2}} |\mathbf{D}^\top|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \mathbf{x}_*^\top \mathbf{D}^{-1} \mathbf{x}_* \right\}.$$

Note que o sistema $\mathbf{A} \mathbf{x}_* = \mathbf{x}$ pode ser escrito explicitamente como

$$\begin{aligned} x_{*1} &= x_1 - \mu \\ \frac{\theta}{1 + \theta^2} x_{*1} + x_{*2} &= x_2 - \mu \\ &\vdots \\ \frac{\theta[1 + \theta^2 + \cdots + \theta^{2(n-2)}]}{1 + \theta^2 + \cdots + \theta^{2(n-1)}} x_{*(t-1)} + x_{*t} &= x_t - \mu \end{aligned} \tag{2.6}$$

Assim,

$$x_{*t} = x_t - \mu - \frac{\theta[1 + \theta^2 + \cdots + \theta^{2(t-2)}]}{1 + \theta^2 + \cdots + \theta^{2(t-1)}} x_{*(t-1)} \tag{2.7}$$

Além disso, o t -ésimo elemento da diagonal principal da matriz $\mathbf{D}$ é

$$d_{tt} = \sigma^2 \frac{1 + \theta^2 + \cdots + \theta^{2t}}{1 + \theta^2 + \cdots + \theta^{2(t-1)}}.$$

Pela propriedade do determinante de matrizes diagonais, vale-se então

$$|\mathbf{D}| = \prod_{t=1}^n d_{tt} \quad \text{e} \quad \mathbf{x}_*^\top \mathbf{D}^{-1} \mathbf{x}_* = \sum_{t=1}^n \frac{\mathbf{x}_{*t}^2}{d_{tt}}$$

Finalmente, a log-verossimilhança não condicional pode ser dada por

$$\ell_{\text{NC}}(c, \theta, \sigma^2) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n \log(d_{tt}) - \frac{1}{2\sigma^2} \sum_{t=1}^n \frac{x_{*t}^2}{d_{tt}}$$

Na Tabela 2.3, apresenta-se os resultados finais das expressões das log-verossimilhanças condicionais e não condicionais e suas respectivas somas quadráticas para um processo MA(1).

Tabela 2.3: Log-verossimilhanças e suas respectivas somas quadráticas do processo MA(1)

Log verossimilhanças condicional e não condicional
$\ell_C(c, \theta, \sigma^2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} S_C(c, \theta)$
$\ell_{NC}(c, \theta, \sigma^2) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n \log(d_{tt}) - \frac{1}{2\sigma^2} S_{NC}(c, \theta)$
Soma dos quadrados condicional e não condicional
$S_C(c, \theta) = \sum_{t=1}^n z_t^2 = \sum_{t=1}^n (x_t - c - \theta z_{t-1})^2$
$S_{NC}(c, \theta) = \sum_{t=1}^n \frac{\hat{x}_t^2}{d_{tt}}$

Verossimilhança para processos ARMA(1,1) gaussianos:

Seja $\{\dot{X}_t; t = 1 \dots, n\} \sim \text{ARMA}(p, q)$ com média $\mathbb{E}(X_t) = \mu = c(1 - \phi_1 - \dots - \phi_p)^{-1}$ e entrada $\{Z_t\} \sim \text{RBG}(0, \sigma^2)$. Neste caso, $\boldsymbol{\delta} = (c, \boldsymbol{\phi}^\top, \boldsymbol{\theta}^\top, \sigma^2)^\top$. Dados os valores iniciais $\mathbf{x}_0 = (x_0, x_{-1}, \dots, x_{1-p})^\top$ e $\mathbf{z}_0 = (z_0, z_{-1}, \dots, z_{1-q})^\top$, a sequência $\{Z_1, Z_2, \dots, Z_n\}$ pode ser calculada de $\{X_1, \dots, X_n\}$ pela identidade

$$Z_t = X_t - c - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} - \theta_1 Z_{t-1} - \dots - \theta_q Z_{t-q}. \quad (2.8)$$

Segundo BOX e JENKINS (1976, p. 211), recomenda-se que o conjunto de valores iniciais para Z 's seja o vetor nulo e para X 's seja os próprios atuais valores. Assim, a iteração 2.8 é inicializada em $t = p + 1$ com o conjunto de valores observados $x_1, \dots, x_p$ e $z_p = z_{p-1} = \dots = z_{p-q+1} = 0$. Então a log-verossimilhança condicional é calculada da seguinte forma

$$\ell_C(\boldsymbol{\delta}) = \log f(x_n, \dots, x_{p+1} | z_p = z_{p+1-q} = 0) = -\left(\frac{n-p}{2}\right) (\log 2\pi - \log \sigma^2) - \frac{1}{2\sigma^2} \sum_{t=p+1}^n z_t^2.$$

A versão não condicional para a log-verossimilhança foi proposta por Galbraith e Galbraith (1974):

$$\ell_{NC}(\boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2) = -\left(\frac{n}{2}\right) (\log 2\pi - \log \sigma^2) + \frac{1}{2} \log |\mathbf{M}| - \frac{1}{2\sigma^2} \mathbf{x}^\top \mathbf{M} \mathbf{x}.$$

Em particular, para processos ARMA(1,1), a log-verossimilhança não condicional fica definida pelas seguintes expressões:

$$|\mathbf{M}| = \left[\frac{(1 - \phi\theta)^2 - (\theta - \phi)^2 \theta^{2n}}{(1 - \phi^2)(1 - \theta^2)} \right]^{-1}$$

e $\mathbf{M} = \{m_{r,s}\}_{r,s=1,\dots,n}$ tal que (para $1 \leq r < s \leq n$)

$$m_{rr} = \frac{(1 - \phi\theta)^2(\theta - \phi)^2 [1 - \theta^{2(r-1)}] [1 - \theta^{2(n-r)}] + (1 - \theta^2)[(1 - \phi\theta)^2 - (\theta - \phi)^2\theta^{2(n-1)}]}{(1 - \theta^2)[(1 - \phi\theta)^2 - (\theta - \phi)^2\theta^{2n}]} \quad (2.9)$$

e

$$m_{rs} = \frac{\theta^{s-r-1}(1 - \phi\theta)(\theta - \phi) [1 - \phi\theta - (\theta - \phi)\theta^{2(r-1)}] [(1 - \phi\theta) - (\theta - \phi)^2\theta^{2(n-s)+1}]}{(1 - \theta^2)[(1 - \phi\theta)^2 - (\theta - \phi)^2\theta^{2n}]} \quad (2.10)$$

Na Tabela 2.4, apresentam-se as log-verossimilhanças e suas somas quadráticas nas versões condicionais e não condicionais.

Tabela 2.4: Log verossimilhança e Mínimos quadrados do processo $ARMA(p, q)$

Log verossimilhança condicional e não condicional
$\ell_C(c, \phi, \theta, \sigma^2) = -\frac{(n-p)}{2} \log(2\pi) - \frac{(n-p)}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} S_C(c, \theta, \phi)$ $\ell_{NC}(c, \phi, \theta, \sigma^2) = -\left(\frac{n}{2}\right) (\log 2\pi - \log \sigma^2) + \frac{1}{2} \log \mathbf{M} - \frac{1}{2\sigma^2} \mathbf{x}^\top \mathbf{M} \mathbf{x}$
Soma dos quadrados condicional e não condicional
$S_C(c, \theta, \phi) = \sum_{t=p+1}^n (x_t - c - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p} - \theta_1 z_{t-1} - \dots - \theta_q z_{t-q})^2$ $S_{NC}(c, \theta, \phi) = \mathbf{x}^\top \mathbf{M} \mathbf{x}$

2.3 Critérios clássicos para identificação de modelos

Iniciemos com uma breve apresentação da evolução dos critérios clássicos de seleção de modelos em regressão múltipla e séries temporais. O primeiro critério de seleção de modelos a ser amplamente utilizado foi o R^2 ajustado (R_{adj}^2), como uma correção do R^2 , que aumenta quando uma variável é adicionada, recomendado sempre o aumento na complexidade de um modelo. Outros critérios de seleção de modelos foram propostos em meados das décadas de 60 e 70, com destaque para FEP (*Final Prediction Error*-FPE) proposto por Akaike (1969). Depois, abordagens relativas à teoria da informação também apareceram na década de 70. Um dos mais importantes foi o AIC (*Akaike Information Criterion*-AIC) também proposto por Akaike em (AKAIKE, 1973) e (AKAIKE, 1974) e fundamentado na divergência de *Kullback-Leibler*. É importante citar que o critério AIC é comumente usado para seleção de modelos em séries temporais. A partir deste trabalho, houve um crescimento de trabalhos colaborando na combinação “seleção de modelos e teoria da informação”. Como exemplos, tem-se:

- o BIC (*Bayesian Information Criterion*–BIC) ou equivalentemente o SIC (*Schwarz Information Criterion*–SIC) introduzido por Schwarz (1978).
- o HQ (*Hannan e Quinn Criterion*–HQ) proposto por Hannan e Quinn (1979).
- o FEPu–FEP corrigido–proposto por Bhansali e Downham (1977).

Por volta da década de 80, Hurvich e Tsai (1989) adaptou os resultados de Sugiura (1978) para desenvolver um estimador não viesado da divergência de *Kullback-Leibler*. Este resultados foi usado para contornar problemas de inconsistências do critério AIC, resultando no AICc.

Até então, existiam várias medidas de seleção dispersas. Shibata (1980) organizou algumas medidas em um grupo, estabelecendo a noção de *eficiência assintótica* como um paradigma para selecionar modelos mais apropriados. A partir de então se fundamentou o conceito de melhor critério sob duas perspectivas: *consistência* (Definição 5) e *eficiência* (Definição 4).

Definição 4 (Critério de eficiência). *Uma suposição comum em regressão e séries temporais é que o modelo verdadeiro ou gerador tem dimensão infinita ou que o conjunto de modelos candidatos não contém o modelo verdadeiro, diga-se X_t^* . O objetivo então é selecionar um modelo que melhor se aproxima de X_t^* de um conjunto de modelos candidatos finitos dimensionais. De acordo com Shibata (1980), para amostras suficientemente grande, o critério que escolhe o modelo com erro quadrático médio mínimo é assintoticamente eficiente. FPE, FEPu, AIC, AICc e AICu são alguns exemplos de critérios eficientes.*

Definição 5 (Critério de eficiência). *Outra suposição aceitável é de que o modelo verdadeiro é finito dimensional e que está incluso no conjunto de modelos candidatos. Neste caso, objetiva-se escolher o modelo da lista de modelos candidatos. O critério que identifica o modelo correto assintoticamente com probabilidade um é chamado de consistente. SIC, HQ e HQc são exemplos de critérios consistentes.*

A seguir, apresentam-se brevemente os critérios clássicos assumidos nesta dissertação que serão comparados com algumas medidas propostas no próximo capítulo. Esta comparação será feita com base na probabilidade de *Overfitting*; isto é, na probabilidade de escolher um modelo com mais variáveis ao modelo verdadeiro.

Inicialmente, seja $\{X_t; t = 1, \dots, n\} \sim \text{AR}(p)$, isto é,

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + Z_t \quad \text{e } Z_t \sim \text{RBG}(0, \sigma^2), \quad (2.11)$$

em que n é o número de observações do processo e $Z_t \sim \text{RBG}$ representa um ruído branco gaussiano ou normal. Neste caso, tem-se que os estimadores de máxima verossimilhança não viesados para a variância do ruído branco σ^2 são dados, respectivamente, por:

$$\hat{\sigma}_p^2 = \sum_{t=p+1}^n \frac{\left(X_t - \hat{\phi}_1 X_{t-1} - \cdots - \hat{\phi}_p X_{t-p} \right)^2}{n-p} \quad (2.12)$$

e

$$S_p^2 = \sum_{t=p+1}^n \frac{\left(X_t - \hat{\phi}_1 X_{t-1} - \cdots - \hat{\phi}_p X_{t-p} \right)^2}{n-2p}. \quad (2.13)$$

Com base neste dois estimadores, segue-se uma rápida explanação de alguns métodos de identificação de ordem da literatura que serão utilizados no estudo de comparação com as novas propostas de critério de seleção de modelos.

AIC

O critério AIC pode ser adaptado para um modelo autoregressivo de acordo com $\hat{\sigma}^2$ dada em (2.12) como a seguir:

$$AIC = \log(\hat{\sigma}_p^2) + \frac{2(p+1)}{n-p}.$$

AICc

Hurvich e Tsai (1989) derivou o critério AIC corrigido obtendo o valor esperado da divergência de *Kullback-Leibler* (KL) para modelos autorregressivos gaussianos. Assumindo que o modelo verdadeiro seja $\text{AR}(p_*)$, em que p_* é a ordem verdadeira do modelo. Então, aplicando o valor esperado na divergência de *Kullback-Leibler* (discutida no próximo capítulo) e assumindo que $p > p_*$, em que p é a ordem do modelo candidato, tem-se

$$\begin{aligned} \mathbb{E}_*[\text{Divergência KL}] &= \mathbb{E}_* \left[\log \left(\frac{\hat{\sigma}_p^2}{\sigma_*^2} \right) + \frac{\sigma_*^2}{\hat{\sigma}_p^2} + \frac{L_2}{\hat{\sigma}_p^2} - 1 \right] \\ &= \mathbb{E}_*[\log(\hat{\sigma}_p^2)] - \log(\sigma_*^2) + \frac{n}{n-2p-2} - 1, \end{aligned}$$

em que $\mathbb{E}_*$ é o valor esperado sob o modelo verdadeiro e $L_2 = \frac{1}{n-p} \sum_{t=p+1}^n (\hat{y}_t - \mu_{*t})^2$. Seja $\log(\hat{\sigma}_p^2)$ um estimador não viesado para $\mathbb{E}_*[\log(\hat{\sigma}_p^2)]$ e descartando o termo constante $\log(\sigma_*^2) - 1$, então

$$AICc = \log(\hat{\sigma}_p^2) + \frac{n}{n-2p-2}.$$

AICu

O critério AICu pode ser adaptado para modelos autorregressivos utilizando s_p^2 em (2.13). Deste modo,

$$AICu = \log(\hat{s}_p^2) + \frac{n}{n - 2p - 2}.$$

FEP

Considere uma realização n dimensional do processo AR(p), $\{Y_t; t = 1, \dots, n\}$ tal que

$$Y_t = \sum_{i=1}^p \phi_i Y_{t-i} V_t \text{ e } V_t \sim \text{RBG}(0, \sigma^2),$$

independente de X_t em (2.11), isto é (Y_t, V_t) é independente de (X_t, Z_t) . Assim, um preditor natural para X_t é dado por

$$\hat{X}_t = \hat{\phi}_1 X_{t-1} + \dots + \hat{\phi}_p X_{t-p+1},$$

em que $\hat{\phi}_1, \dots, \hat{\phi}_p$ são as estimativas de Yule-Walker obtidas por ajustar um modelo AR(p) às observações $\{Y_1, \dots, Y_n\}$. Usando o Teorema 2 e a independência entre $\{X_t\}$ e $\{\hat{\phi}_1, \dots, \hat{\phi}_p\}$, Akaike (1969) mostrou que-para n suficientemente grande-o erro quadrático médio de predição aproximado é dado por

$$\mathbb{E} \left[(\hat{X}_{n+1} - X_{n+1})^2 \right] \approx \sigma^2 \left(1 + \frac{p}{n - p} \right). \quad (2.14)$$

Assumindo que o modelo verdadeiro é AR(p_*) e $p > p_*$, então

$$\mathbb{E}[\hat{\sigma}_p^2] \approx \left(\frac{n - 2p}{n - p} \right) \sigma^2.$$

Desta forma, substituindo a estimativa não viesada para σ^2 , $\left(\frac{n - 2p}{n - p} \right)^{-1} \hat{\sigma}^2$, na equação (2.14), obtem-se o critério FEP de Akaike dado por

$$\text{FEP} = \hat{\sigma}_p^2 \left(\frac{n}{n - 2p} \right). \quad (2.15)$$

FEPu

Usando S_p^2 dado em (2.13) em (2.15), pode-se adaptar o critério FEPu para modelos autorregressivos como

$$\text{FEPu} = S_p^2 \left(\frac{n}{n - 2p} \right).$$

SIC

O critério SIC proposto por Schwarz (1978) é equivalente ao critério BIC de Akaike (1978) para modelos autorregressivos (PRIESTLEY, 1981, p. 376). O SIC pode ser definido como

$$\text{SIC} = \log(\hat{\sigma}_p^2) + \left[\frac{\log(n-p)p}{n-p} \right].$$

HQ

O critério consistente HQ proposto por Hannan e Quinn (1979) para modelos autorregressivos é definido como

$$HQ = \log(\hat{\sigma}_p^2) + \left(\frac{2 \log \log(n-p)p}{n-p} \right).$$

HQu

A versão corrigida do critério de seleção HQ para modelos autorregressivos utilizando $\hat{\sigma}_p^2$ fornecido na equação 2.12 é dada por

$$HQ_u = \log(\hat{\sigma}_p^2) + \left(\frac{2 \log \log(n-p)p}{n-2p-2} \right).$$

Capítulo 3

Critérios de Informação em Modelos ARMA

Mahalanobis (1936) introduziu uma das medidas mais importantes para a área de distâncias em espaços de probabilidade. A formalização de conceitos neste campo é um dos objetivos da *Teoria Estatística da Informação*. Nesta seção, dar-se-á ênfase a testes de hipóteses sob medidas de divergências em processos ARMA.

A *Teoria da informação* pode ser definida como a formalização matemática da transmissão de informações entre canais estocásticos. Nyquist (1924), Nyquist (1928) e Hartley (1928) apresentaram os primeiros estudos nesse sentido. Entretanto, foi com Shannon (1948) que os trabalhos de Nyquist e Hartley ganharam fundamentação. Em particular, três conceitos definidos sobre uma medida de probabilidade,

$$f(x; \boldsymbol{\theta}) = \frac{dP_{\boldsymbol{\theta}}}{d\mu}(x) = \begin{cases} f(x; \boldsymbol{\theta}), & \text{se } \mu \text{ é uma medida de Lebesgue (ML),} \\ \Pr_{\boldsymbol{\theta}}(X = x), & \text{se } \mu \text{ é uma medida de contagem (MC)} \end{cases}$$

para $\boldsymbol{\theta} \in \Theta$, foram introduzidos por Shannon:

- Informação: Seja X uma variável aleatória munida com $f(x; \boldsymbol{\theta})$, a informação é definida como

$$I(X) \triangleq -\log f(X; \boldsymbol{\theta}). \quad (3.1)$$

Em particular, para medidas de contagem, as informações associadas aos eventos $[X = \xi_1]$ e $[X = \xi_2]$ tais que $P(X = \xi_1) > P(X = \xi_2)$ obedecem à desigualdade $I(\xi_2) > I(\xi_1)$; isto é, eventos menos prováveis são mais informativos. Para ilustrar considere a relação dos conceitos probabilidade e informação com o evento “O sol vai nascer amanhã”. É notável que se podessemos quantificar a probabilidade desse evento ela seria próxima de um; por outro lado, a informação associada à este evento seria próxima de zero, o que é bastante intuitivo pois o evento é uma redundância ao que se espera.

- Entropia: Subsequentemente, o conceito (3.1) foi utilizado para definir a quantidade média de informação provida por um sistema na forma da definição de *entropia de Shannon* dada por

$$H(X) \triangleq \int_{\mathcal{X}} -f(x; \boldsymbol{\theta}) \log f(x; \boldsymbol{\theta}) d\mu(x). \quad (3.2)$$

A entropia de Shannon tem sido estendida por Burbea e Rao (1982) com uma classe de medidas chamada “ ϕ -entropia” e definida como

$$\begin{aligned} H_{\phi}(X) &\triangleq H_{\phi}(P_{\boldsymbol{\theta}}) \triangleq H_{\phi}(\boldsymbol{\theta}) \triangleq \int_{\mathcal{X}} \phi(f(x; \boldsymbol{\theta})) d\mu(x) \\ &= \begin{cases} \int_{\mathcal{X}} \phi(f(x; \boldsymbol{\theta})) dx, & \text{se } \mu \text{ ML,} \\ \sum_{x \in S_{\mathcal{X}}} \phi(\Pr_{\boldsymbol{\theta}}(X = x)), & \text{se } \mu \text{ MC,} \end{cases} \end{aligned} \quad (3.3)$$

em que $\phi : (0, \infty) \mapsto \mathbb{R}$ é uma função contínua e concava. Em particular, (3.3) se degenera em (3.2) para $\phi(x) = -x \log(x)$.

- Divergência: Finalmente, o conceito (3.2) foi utilizado para definir o erro de se escolher um modelo $P_{\boldsymbol{\theta}_1}$ quando outro, $P_{\boldsymbol{\theta}_2}$, deveria ter sido escolhido na forma da divergência de Kullback-Leibler

$$\Delta_{\text{KL}}(\boldsymbol{\theta}_1 || \boldsymbol{\theta}_2) \triangleq \int_{\mathcal{X}} \log \frac{f(x; \boldsymbol{\theta}_1)}{g(x; \boldsymbol{\theta}_2)} f_{\boldsymbol{\theta}_1}(x) d\mu(x). \quad (3.4)$$

A medida (3.4) também tem sido estendida em classes de divergências (NASCIMENTO, 2012). Considere as variáveis aleatórias $X \sim f(x; \boldsymbol{\theta}_1)$ e $Y \sim g(x; \boldsymbol{\theta}_2)$ com suporte comum: $\mathcal{X}$, para μ como uma medida de Lebesgue, e $S_{\mathcal{X}}$, quando μ é uma medida de contagem. As classes de divergências podem ser representadas pela seguinte medida:

$$\begin{aligned} \Delta_{\varphi}(X || Y) &\triangleq \Delta_{\varphi}(P_{\boldsymbol{\theta}_1} || P_{\boldsymbol{\theta}_2}) \triangleq \Delta_{\varphi}(\boldsymbol{\theta}_1 || \boldsymbol{\theta}_2) \triangleq \int_{\mathcal{X}} \varphi(f(x; \boldsymbol{\theta}_1), g(x; \boldsymbol{\theta}_2)) d\mu(x) \\ &= \begin{cases} \int_{\mathcal{X}} \varphi(f(x; \boldsymbol{\theta}_1), g(x; \boldsymbol{\theta}_2)) dx, & \text{se } \mu \text{ ML,} \\ \sum_{x \in S_{\mathcal{X}}} \varphi(f(x; \boldsymbol{\theta}_1), g(x; \boldsymbol{\theta}_2)), & \text{se } \mu \text{ MC,} \end{cases} \end{aligned} \quad (3.5)$$

em que $\varphi(x, y)$ é uma função real das variáveis $x, y > 0$. Neste caso, $\varphi(\cdot, \cdot)$ representa qualquer função definida em $(x, y) \in [0, \infty) \times [0, \infty)$ tal que $\varphi(0, 0) = 0$, $\varphi(0, y) = \lim_{x \downarrow 0} \varphi(x, y)$ e $\varphi(x, 0) = \lim_{y \downarrow 0} \varphi(x, y)$.

A seguir, apresenta-se uma discussão da classe de divergências utilizada nesta dissertação.

3.1 Classe (h, ϕ) -divergência: Definição e resultados assintóticos

Csiszár (1967) e Ali e Silvey (1996) propuseram uma classe de divergências denotada como “ ϕ -divergência” com base na família de classes (3.5). Nesta caso, especificou-se o integrando de (3.5) como $\varphi(x, y) = y \phi(x/y)$, resultando na classe:

$$\Delta_\phi(X||Y) \triangleq \Delta_\phi(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) \triangleq \int_{\mathcal{X}} g(x; \boldsymbol{\theta}_2) \phi\left(\frac{f(x; \boldsymbol{\theta}_1)}{g(x; \boldsymbol{\theta}_2)}\right) d\mu(x). \quad (3.6)$$

Entretanto, esta classe tem a desvantagem de não abordar medidas importantes de divergências, tais como Bhattacharyya (KAILATH, 1967), Rényi (RÉNYI, 1961) e Sharma and Mittal (LIESE; VAJDA, 2006). Como solução a este problema, Menéndez et al. (1995) introduziram a classe de divergências (h, ϕ) :

$$\Delta_\phi^h(X||Y) \triangleq \Delta_\phi^h(P_{\boldsymbol{\theta}_1}||P_{\boldsymbol{\theta}_2}) \triangleq \Delta_\phi^h(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) \triangleq h(\Delta_\phi(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2)), \quad (3.7)$$

em que $h: (0, \infty) \rightarrow [0, \infty)$ é uma função estritamente crescente e $\phi: (0, \infty) \rightarrow [0, \infty)$ é uma função convexa tal que $0 \phi(0/0) = 0$, $0 \phi(x/0) = \lim_{x \rightarrow \infty} \phi(x)/x$ e $\phi(1) = h(0) = 0$ e $\phi''(1) > 0$.

3.1.1 Divergências utilizadas

A classe de divergências (h, ϕ) entre dois processos n -dimensionais $\mathbf{x} = \{X_t; t = 1, \dots, n\} \sim f(\mathbf{z}; \boldsymbol{\theta}_1)$ e $\mathbf{y} = \{Y_t; t = 1, \dots, n\} \sim g(\mathbf{z}; \boldsymbol{\theta}_2)$ para $\mathbf{z} \in \mathcal{X} \subset \mathbb{R}^n$ é definida por

$$\Delta_\phi^h(\mathbf{x}||\mathbf{y}) = h\left(\mathbb{E}_{\mathbf{y}} \phi\left(\frac{f(\mathbf{y}; \boldsymbol{\theta}_1)}{g(\mathbf{y}; \boldsymbol{\theta}_2)}\right)\right) = h\left(\int_{\mathcal{X}} \phi\left(\frac{f(\mathbf{z}; \boldsymbol{\theta}_1)}{g(\mathbf{z}; \boldsymbol{\theta}_2)}\right) g(\mathbf{z}; \boldsymbol{\theta}_2) d\mathbf{z}\right), \quad (3.8)$$

em que o elemento diferencial $d\mathbf{z}$ é dado por $d\mathbf{z} = \prod_{i=1}^n dz_i$ e z_i é o i -ésimo elemento de $\mathbf{z}$ para $i = 1, \dots, n$, em que $\mathbf{z}$ representa a variável que é definida sobre um suporte comum. Divergências consolidadas na literatura podem ser obtidas por escolhas de h e ϕ em (3.8). Nesta dissertação, trabalhamos com cinco divergências cujas funções h e ϕ são dadas na Tabela 3.1.

Tabela 3.1: Divergências (h, ϕ) e h e ϕ associadas

(h, ϕ) -divergência	$h(y)$	$\phi(x)$
Kullback-Leibler	y	$x \log(x) - x + 1$
Tipo Kullback-Leibler	y	$-\log(x) + x - 1$
Jenssen-Shannon	$y/2$	$(x - 1) \log(x)$
Bhattacharyya	$-\log(-y + 1), 0 \leq y < 1$	$-\sqrt{x} + \frac{x+1}{2}$
Hellinger	$y/2, 0 \leq y < 2$	$(\sqrt{x} - 1)^2$

No que segue, a fórmula de cada divergência é explicitada e aplicada ao caso em que, para $\boldsymbol{\theta}_i = (\boldsymbol{\mu}_i^\top, \text{vec}(\boldsymbol{\Sigma}_i)^\top)^\top$ em $i = 1, 2$ e $\text{vec}(\cdot)$ como o operador de vetorização,

$$\mathbf{x}(\boldsymbol{\theta}_1) \sim N_n(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) : f(\mathbf{z}; \boldsymbol{\theta}_1) = |2\pi\boldsymbol{\Sigma}_1|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{z} - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_1^{-1} (\mathbf{z} - \boldsymbol{\mu}_1) \right\}$$

e

$$\mathbf{y}(\boldsymbol{\theta}_2) \sim N_n(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) : g(\mathbf{z}; \boldsymbol{\theta}_2) = |2\pi\boldsymbol{\Sigma}_2|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{z} - \boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}_2^{-1} (\mathbf{z} - \boldsymbol{\mu}_2) \right\}.$$

Divergência de Kullback-Leibler (Δ_{KL}):

$$\begin{aligned} \Delta_{\text{KL}}(\mathbf{x}||\mathbf{y}) &= \int_{\mathcal{X}} f(z_1, \dots, z_n; \boldsymbol{\theta}_1) \log \frac{f(z_1, \dots, z_n; \boldsymbol{\theta}_1)}{g(z_1, \dots, z_n; \boldsymbol{\theta}_2)} dz_1 \cdots dz_n \\ &= \frac{1}{2} \left[(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}_2^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) + \text{tr}(\boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 - \mathbb{I}_n) + \log \left(\frac{|\boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_1|} \right) \right]. \end{aligned}$$

Esta divergência tem uma relação com o lema de Neyman-Pearson (EGUCHI; COPAS, 2006) e com base nesta medida se propôs o AIC como discutido capítulo anterior.

Divergência Tipo Kullback-Leibler (Δ_{TKL}):

$$\begin{aligned} \Delta_{\text{TKL}}(\mathbf{x}||\mathbf{y}) &= \int_{\mathcal{X}} g(z_1, \dots, z_n; \boldsymbol{\theta}_1) \log \frac{g(z_1, \dots, z_n; \boldsymbol{\theta}_1)}{f(z_1, \dots, z_n; \boldsymbol{\theta}_2)} dz_1 \cdots dz_n \\ &= \frac{1}{2} \left[(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}_1^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) + \text{tr}(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_2 - \mathbb{I}_n) + \log \left(\frac{|\boldsymbol{\Sigma}_1|}{|\boldsymbol{\Sigma}_2|} \right) \right]. \end{aligned}$$

Distância Jensen-Shannon (Δ_{JS}):

$$\begin{aligned} \Delta_{\text{JS}}(\mathbf{x}, \mathbf{y}) &= \Delta_{\text{KL}}(\mathbf{x}||\mathbf{y}) + \Delta_{\text{TKL}}(\mathbf{x}||\mathbf{y}) \\ &= \frac{1}{2} \left[(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top (\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1}) (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) + \text{tr}(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_2 + \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1) - 2n \right]. \end{aligned}$$

Em teoria da probabilidade e estatística, a divergência *Jensen-Shannon* é utilizada para medir a similaridade entre duas distribuições de probabilidade. Esta medida baseia-se na divergência de *Kullback-Leibler*, no qual pode ser obtida pela soma das divergências Δ_{KL} e Δ_{TKL} .

Distância Battacharyya (Δ_B):

$$\begin{aligned} \Delta_B(\mathbf{x}, \mathbf{y}) &= -\log \int_{\mathcal{X}} \sqrt{g(z_1, \dots, z_n; \boldsymbol{\theta}_1) f(z_1, \dots, z_n; \boldsymbol{\theta}_2)} dz_1 \cdots dz_n \\ &= \frac{1}{2} \left[\log |\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2| - n \log 2 + \frac{\log |\boldsymbol{\Sigma}_1^{-1}| + \log |\boldsymbol{\Sigma}_2^{-1}|}{2} + \frac{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top [\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2]^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}{2} \right]. \end{aligned}$$

Goudail, Réfrégier e Delyon (2004) mostraram que essa distância é uma ferramenta mais precisa do que a distância de *Kullback-Leibler* quanto a definição de contraste em algoritmos para processamento de imagens PolSAR.

Distância Hellinger (Δ_H):

$$\Delta_H(\mathbf{x}, \mathbf{y}) = 1 - \int_{\mathcal{X}} \sqrt{g(z_1, \dots, z_n; \boldsymbol{\theta}_1) f(z_1, \dots, z_n; \boldsymbol{\theta}_2)} dz_1 \cdots dz_n = 1 - \exp[-\Delta_B(\mathbf{x}, \mathbf{y})].$$

Giet e Lubrano (2008) empregaram métodos de estimação baseados na minimização de Δ_H em contextos de equações diferenciais estocásticas. Dentre as distâncias consideradas, a distância de *Hellinger* é a única limitada, apresentando valores no intervalo $[0, 1]$. Deste modo, essa medida funciona como uma taxa.

3.1.2 Resultados assintóticos: Divergência vs razão entre verossimilhanças

De acordo com Kupperman (1957), a estatística de teste para a hipótese nula $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ com base na divergência de Kullback-Leibler tem o seguinte comportamento assintótico

$$2n \Delta_{KL}(\hat{\boldsymbol{\theta}} || \boldsymbol{\theta}_0) = 2n \mathbb{E}_{\hat{\boldsymbol{\theta}}} \left[\log \frac{f_{\hat{\boldsymbol{\theta}}}(X)}{f_{\boldsymbol{\theta}_0}(X)} \right] = 2n \int_{\mathcal{X}} \log \frac{f_{\hat{\boldsymbol{\theta}}}(x)}{f_{\boldsymbol{\theta}_0}(x)} dF_{\hat{\boldsymbol{\theta}}}(x) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \chi_M^2, \quad (3.9)$$

em que $\mathcal{X}$ é o suporte de uma variável aleatória X em $(\sigma_{\mathcal{X}}, \text{Pr}_{\boldsymbol{\theta}})$ com função distribuição acumulada $F_{\boldsymbol{\theta}}(x)$, M é a dimensão de $\boldsymbol{\theta} = (\theta_1, \dots, \theta_M)^\top$ e $\hat{\boldsymbol{\theta}}$ é o estimador de máxima verossimilhança (EMV) para $\boldsymbol{\theta}$.

Salicrú et al. (1994) estendeu o resultado (3.9) para qualquer estatística definida a partir de (3.7) na forma do seguinte corolário.

Corolário 1. *Assuma a satisfação das condições de regularidade dadas em Salicrú et al. (1994). Se $\hat{\boldsymbol{\theta}}$ é o EMV para $\boldsymbol{\theta}$ e $\boldsymbol{\theta}_0$ é conhecido, então*

(i)

$$\sqrt{n} [\Delta_{\phi}^h(\hat{\boldsymbol{\theta}} || \boldsymbol{\theta}_0) - \Delta_{\phi}^h(\boldsymbol{\theta} || \boldsymbol{\theta}_0)] \xrightarrow[n \rightarrow \infty]{\mathcal{D}} N(0, \mathbf{t}^\top K(\boldsymbol{\theta}) \mathbf{t}),$$

em que $K(\boldsymbol{\theta}) = \mathbb{E}[(\partial \log f(\mathbf{x}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta})(\partial \log f(\mathbf{x}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta}^\top)]$ é a matriz informação de Fisher e $\mathbf{t} = (t_1, \dots, t_M)$ é tal que sua i -ésima entrada é

$$t_i = \left(\int_{\mathcal{X}} \frac{\partial f(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta_i} \phi' \left(\frac{f(\mathbf{x}; \boldsymbol{\theta})}{g(\mathbf{x}; \boldsymbol{\theta}_0)} \right) d\mu(\mathbf{x}) \right) h' \left(\int_{\mathcal{X}} g(\mathbf{x}; \boldsymbol{\theta}_0) \phi \left(\frac{f(\mathbf{x}; \boldsymbol{\theta})}{g(\mathbf{x}; \boldsymbol{\theta}_0)} \right) d\mu(\mathbf{x}) - \phi(1) \right).$$

(ii) Se $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ é verdadeira e $h'(0) \phi''(1) \neq 0$, então

$$S_\phi^h \triangleq S_\phi^h(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) \triangleq \frac{2n \Delta_\phi^h(\hat{\boldsymbol{\theta}} || \boldsymbol{\theta}_0)}{h'(0) \phi''(1)} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \chi_M^2,$$

em que χ_M^2 denota a distribuição qui-quadrado com M graus de liberdade.

Nesta dissertação, consideraremos cinco estatísticas de teste para $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ no contexto de processos ARMA com base nas distâncias Δ_{KL} , Δ_{TKL} , Δ_{JS} , Δ_{B} e Δ_{H} dadas por S_{KL} , S_{TKL} , S_{JS} , S_{B} e S_{H} . Estes testes serão comparados com o teste clássico da razão entre verossimilhança, que tem seu resultado assintótico baseado no teorema de Wilks.

Teorema 4 (Teorema de Wilks). *Assuma que sejam satisfeitas para $X \sim P_{\boldsymbol{\theta}}$ as mesmas condições de regularidade do Corolário 1. Então, assumindo que $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ é verdadeira e que $n \rightarrow \infty$, a estatística da razão entre verossimilhanças S_{RV} é*

$$S_{\text{RV}} \triangleq 2[\ell(\hat{\boldsymbol{\theta}}) - \ell(\tilde{\boldsymbol{\theta}})] \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \chi_M^2,$$

em que $\ell(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$ é a função de log-verossimilhança, $\hat{\boldsymbol{\theta}}$ é o EMV sem restrição, $\tilde{\boldsymbol{\theta}}$ é o EMV sob restrição de $\mathcal{H}_0$ e $M = \nu - l$ ($l < \nu$) é diferença entre as quantidades de parâmetros livres nas hipóteses alternativa (ν) e nula (l).

Para modelos ARMA, tem-se duas log-verossimilhanças (condicional e não condicional) como discutido anteriormente. Nesta dissertação, denotaremos S_{RVC} e S_{RVNC} para as estatísticas da RV condicional e não condicional, respectivamente.

3.2 Contribuições teóricas

Nesta seção, apresentam-se resultados analíticos para explicitar as medidas Δ_{KL} , Δ_{TKL} , Δ_{JS} , Δ_{B} e Δ_{H} em contextos de interesse nos capítulos de resultados numéricos por simulação Monte Carlo e de aplicação.

3.2.1 Divergências entre processos AR(p)

Teorema 5. *Sejam $\{X_t^*; t = 1, \dots, n\} \sim \text{AR}(k) : \boldsymbol{\beta}^* = (\phi_1^*, \dots, \phi_k^*, \sigma_*^2)$ com matriz de covariâncias $\boldsymbol{\Omega}(\boldsymbol{\beta}^*)$ e $\{X_t; t = 1, \dots, n\} \sim \text{AR}(p) : \boldsymbol{\beta} = (\phi_1, \dots, \phi_p, \sigma^2)$ com matriz de covariâncias $\boldsymbol{\Omega}(\boldsymbol{\beta})$ para $k \leq p$ os modelos verdadeiro e aproximado, respectivamente. Assuma também que $\mu_t^* = \mathbb{E}(X_t | X_{t-1}^*, \dots, X_{t-k+1}^*)$ para $t \geq k$ e $\mu_t = \mathbb{E}(X_t | X_{t-1}, \dots, X_{t-p+1})$*

para $t \geq p$ como flexibilização da condição de estacionaridade dada em (2) afim de incluir descrição de modelos econométricos. A divergência de Kullback-Leibler entre os processos X_t^* e X_t , $\Delta_{KL}(\beta^* \parallel \beta)$, é dada por

$$\begin{aligned} & \frac{1}{2} \left\{ \underbrace{\frac{\sigma^2}{\sigma_*^2} \left[1 - \sum_{i=1}^k \phi_i^* \rho_X^*(i) \right]}_{\nu(\beta^*, \beta)} \left[\sum_{r=1}^n K^{*,r,r} + 2 \sum_{1 \leq r < s \leq n} K_{r,s} K^{*,r,s} \right] - n \right. \\ & \quad \left. + \nu(\beta^*, 1) \left\{ \sum_{r=p+1}^n K^{*,r,r} [\hat{X}_r(\beta^*) - \hat{X}_r(\beta)]^2 + 2 \sum_{1 \leq r < s \leq n} K_{r,s} K^{*,r,s} [\hat{X}_r(\beta^*) - \hat{X}_r(\beta)] [\hat{X}_s(\beta^*) - \hat{X}_s(\beta)] \right\} \right. \\ & \quad \left. + n \log \left(\frac{\sigma_*^2}{\sigma^2} \right) + \sum_{r,s=1}^p \log(1 - w_i(\beta) w_j(\beta)) - \sum_{r,s=1}^k \log(1 - w_i(\beta^*) w_j(\beta^*)) \right\}, \end{aligned} \quad (3.10)$$

em que $\nu(\beta^*, \beta) \triangleq \text{Var}(X_t)/\text{Var}(X_t^*)$ e $\nu(\beta^*, 1) \triangleq 1/\text{Var}(X_t^*)$, $\rho_X^*(i)$ e $\rho_X(i)$ são as FACs dos processos $\{X_t^*\}$ e $\{X_t\}$, $w_1(\beta), \dots, w_p(\beta)$ são as raízes de $\sum_{i=0}^p \phi_i w^{p-i} = 0$ com $\phi_0 = 1$, $w_1(\beta^*), \dots, w_p(\beta^*)$ são as raízes de $\sum_{i=0}^k \phi_i^* w^{p-i} = 0$ com $\phi_0^* = 1$, $\hat{X}_r(\beta^*) = \sum_{i=1}^k \phi_i^* X_{r-i}^*$ é o melhor preditor linear do processo $\{X_t^*\}$, $\hat{X}_r(\beta) = \sum_{i=1}^p \phi_i X_{r-i}^*$ é o melhor preditor linear do processo $\{X_t\}$, $K_{r,s}$ é a entrada (r, s) da matriz de correlação de $\{X_t\}$ e $K^{*,r,s}$ é a entrada (r, s) da matriz de correlação inversa de $\{X_t^*\}$. Baseando-se sobre o resultado (3.10), as divergências tipo KL, $\Delta_{TKL}(\beta^* \parallel \beta)$, e segundo Jensen-Shannon (JS), $\Delta_{JS}(\beta, \beta^*) = \Delta_{KL}(\beta^* \parallel \beta) + \Delta_{TKL}(\beta \parallel \beta^*)$, são facilmente obtidas.

Um esboço da prova do Teorema 5 pode ser encontrado no Apêndice A.

Corolário 2. As seguintes expressões para divergências entre processos autoregressivos seguem:

(a) Para $X_t^* \sim AR(1) : \beta^* = (\phi_1^*, \sigma_*^2)$ e $X_t \sim AR(1) : \beta = (\phi_1, \sigma^2)$,

$$\begin{aligned} \Delta_{KL}(\beta^* \parallel \beta) &= \frac{\sigma^2 (1 - \phi_1^{*2})}{\sigma_*^2 (1 - \phi_1^2)} \left[\sum_{r=1}^n K^{*,r,r} + 2 \sum_{1 \leq r < s \leq n} \phi_1^{s-r} K^{*,s,r} \right] - \log \left(\frac{1 - \phi_1^{*2}}{1 - \phi_1^2} \right) + n \log \left(\frac{\sigma_*^2}{\sigma^2} \right) \\ &+ \frac{1 - \phi_1^{*2}}{\sigma_*^2} \left\{ K^{*,1,1} X_1^{*2} + \sum_{t=2}^n K^{*,t,t} (X_t^* - \phi_1 X_{t-1}^*)^2 + 2 \sum_{1 \leq r < s \leq n} K^{*,s,r} (X_s^* - \phi_1 X_{s-1}^*) (X_r^* - \phi_1 X_{r-1}^*) \right\}. \end{aligned}$$

A versão amostral associada é dada por

$$\begin{aligned} \Delta_{KL}(\phi_1, \sigma^2) &= \frac{\sigma^2}{1 - \phi_1^2} \left[\sum_{r=1}^n \tilde{\Omega}^{r,r} + 2 \sum_{1 \leq r < s \leq n} \phi_1^{s-r} \tilde{\Omega}^{s,r} \right] + \log |\tilde{\Omega}| + \log (1 - \phi_1^2) - n \log (\sigma^2) \\ &+ \left\{ \tilde{\Omega}^{1,1} X_1^{*2} + \sum_{t=2}^n \tilde{\Omega}^{t,t} (X_t^* - \phi_1 X_{t-1}^*)^2 + 2 \sum_{1 \leq r < s \leq n} \tilde{\Omega}^{s,r} (X_s^* - \phi_1 X_{s-1}^*) (X_r^* - \phi_1 X_{r-1}^*) \right\}. \end{aligned} \quad (3.11)$$

(b) Para $X_t^* \sim AR(1) : \beta^* = (\phi_1^*, \sigma_*^2)$ e $X_t \sim AR(2) : \beta = (\phi_1, \phi_2, \sigma^2)$, Δ_{KL} é a expressão resultante de aplicar os seguintes termos em (3.10):

$$\begin{aligned} \nu(\beta^*, \beta) &= \frac{\sigma^2 [1 - \phi_1^{*2}]}{\sigma_*^2 \left[1 - \frac{\phi_1^2}{1-\phi_2} - \frac{(\phi_1^2 + \phi_2 - \phi_2^2)\phi_2}{1-\phi_2} \right]}, \\ K_{r,s} &= \begin{cases} \frac{\phi_1}{1-\phi_2}, & |s-r| = 1, \\ \frac{\phi_1^2 + \phi_2 - \phi_2^2}{1-\phi_2}, & |r-s| = 2, \\ \phi_1 \rho_X(|s-r|-1) + \phi_2 \rho_X(|s-r|-2), & |r-s| > 2, \end{cases} \\ K^{*r,s} &= \begin{cases} \frac{(1-\phi_1^{*2})\mathbb{I}_1(r) + \phi_1^{*2}}{1-\phi_1^{*2}}, & s = r, \\ \frac{-\mathbb{I}_1(|r-s|)\phi_1^*}{1-\phi_1^{*2}}, & s \neq r, \end{cases} \\ [\hat{X}_r(\beta^*), \hat{X}_r(\beta)] &= (\phi_1^* X_{r-1}, \phi_1 X_{r-1} + \phi_2 X_{r-2}) \end{aligned}$$

e

$$\log \frac{|\Omega(\beta^*)|}{|\Omega(\beta)|} = n \log \frac{\sigma_*^2}{\sigma^2} + 2 \log(1 + \phi_2) + \log[(1 - \phi_2)^2 - \phi_1^2] - \log(1 - \phi_1^{*2}).$$

O Corolário 2 decorre da aplicação de propriedades tipo momentos e das equações de predição dos processos AR(1), AR(2) no Teorema 5. Em particular, a Equação (3.11) e as medidas $\Delta_{TKL}(\phi_1, \sigma^2)$ e $\Delta_{JS}(\phi_1, \sigma^2)$ definidas sobre ela sugerem medidas alternativas às medidas de seleção clássicas e serão utilizadas na computação da probabilidade de *Overfitting* dos novos critérios propostos na Seção 4.3 do Capítulo 4.

Teorema 6. *Sejam $X_t^* \sim AR(k) : \beta^* = (\phi_1^*, \dots, \phi_k^*, \sigma_*^2)$ e $X_t \sim AR(p) : \beta = (\phi_1, \dots, \phi_p, \sigma^2)$ para $k \leq p$ os modelos verdadeiro e aproximado, respectivamente. Assuma também que $\mu_t^* = E(X_t^* | X_{t-1}^*, \dots, X_{t-k+1}^*)$ para $t \geq k$ e $\mu_t = E(X_t | X_{t-1}, \dots, X_{t-p+1})$ para $t \geq p$. A distância de Bhattacharyya entre os processos X_t^* e X_t , $\Delta_B(\beta^*, \beta)$, é dada por*

$$\begin{aligned} & (\mu_k - \mu_p)^\top \left[\frac{\Omega(\beta^*) + \Omega(\beta)}{2} \right]^{-1} (\mu_k - \mu_p) \\ & \frac{1}{4} \left[B(\beta)^\top \mu_{11} - B_{11}(\beta, \beta^*)^\top \mu_{11} \right]^\top \left[B(\beta)^\top \mu_{11} - B_{11}(\beta, \beta^*)^\top \mu_{11} \right] + \frac{1}{2} \left\{ -n \log 2 \right. \\ & \left. - \frac{1}{2} \left[n [\log(\sigma_*^2) + \log(\sigma^2)] - \sum_{r,s=1}^p \log(1 - w_i(\beta) w_j(\beta)) - \sum_{r,s=1}^k \log(1 - w_i(\beta^*) w_j(\beta^*)) \right] \right. \\ & \quad \left. \underbrace{\log |\Omega(\beta^*)| + \log |\Omega(\beta)|}_{\log |\Omega(\beta^*) + \Omega(\beta)|} \right\} \\ & + \underbrace{\sum_{r,s=1}^p \log(1 - w_i(\beta) w_j(\beta)) + n \log(\sigma^2) + \log |\mathbb{I}_n + \Omega(\beta^*) \Omega(\beta)^{-1}|}_{\log |\Omega(\beta^*) + \Omega(\beta)|} \left. \right\}, \end{aligned} \quad (3.12)$$

em que $A(\beta^*)$ e $B(\beta)$ são resultantes das fatorações de Cholesky $\Omega(\beta^*) = A(\beta^*) A(\beta^*)^\top$ e $\Omega(\beta) = B(\beta) B(\beta)^\top$, $\mu_{11} = [\mu(\beta) - \mu(\beta^*)]$ com $\mu(\beta) = [\dots | \hat{X}_r(\beta) | \dots]^\top$ e $\mu(\beta^*) = [\dots | \hat{X}_r(\beta^*) | \dots]^\top$ tal como definido no Teorema 5 e $B_{11}(\beta, \beta^*) = \mathcal{H}^\top [\mathbb{I}_n + \mathcal{H} \mathcal{H}^\top]^{-1} \mathcal{H}$

$A(\beta^*)$ com $\mathcal{H} = A(\beta^*)^{-1}B(\beta)$. Baseando-se sobre o resultado (3.12), a distância de Hellinger (H), $\Delta_H(\beta^*, \beta) = 1 - \exp(-\Delta_B)$ é facilmente obtida.

A prova deste teorema pode ser encontrada no Apêndice B. Com base neste teorema decorre o seguinte corolário.

Corolário 3. *As seguintes divergências entre processos autoregressivos são válidas:*

(a) Para $X_t^* \sim AR(1) : \beta^* = (\phi_1^*, \sigma_*^2)$ e $X_t \sim AR(1) : \beta = (\phi_1, \sigma^2)$,

$$\Delta_B(\beta^*, \beta) = -\frac{1}{2} \{ \log(1 - \phi^{*2}) + 3 \log(1 - \phi^2) - n[\log(\sigma^{*2}) - \log(\sigma^2)] \} + \frac{1}{2} \sum_{i=1}^n \log \lambda_{11}^{(i)} \\ + \frac{1}{4} [B(\phi_1, \sigma^2)^\top \mu_{11} - B_{11}(\phi_1, \phi_1^*, \sigma^2, \sigma_*^2)^\top \mu_{11}]^\top [B(\phi_1, \sigma^2)^\top \mu_{11} - B_{11}(\phi_1, \phi_1^*, \sigma^2, \sigma_*^2)^\top \mu_{11}],$$

em que $\lambda_{11}^{(1)} > \dots > \lambda_{11}^{(n)}$ são os autovalores de $\mathbb{I}_n + \Omega(\phi_1^*, \sigma_*^2)\Omega(\phi_1, \sigma^2)^{-1}$.

A versão amostral associada é dada por

$$\Delta_B(\phi_1, \sigma^2) = \frac{1}{4} [B(\phi_1, \sigma^2)^\top \tilde{\mu} - \tilde{B}(\phi_1, \sigma^2)^\top \tilde{\mu}]^\top [B(\phi_1, \sigma^2)^\top \tilde{\mu} - \tilde{B}(\phi_1, \sigma^2)^\top \tilde{\mu}] \\ + \frac{3}{2} \left[n \log(\sigma^2)/3 + \log(1 - \phi^2) - \log |\tilde{\Omega}|/3 \right] + \frac{1}{2} \sum_{i=1}^n \log \tilde{\lambda}_1^{(i)}, \quad (3.13)$$

em que $\tilde{\mu} = [\mu(\phi_1) - \mathbf{x}^*]^\top$ com $\mu(\phi_1) = [0, \phi_1 X_1^*, \dots, \phi_1 X_{n-1}^*]^\top$ e $\mathbf{x}^* = [X_1^*, X_2^*, \dots, X_n^*]^\top$ tal como definido no Teorema 5 e $\tilde{\lambda}_1^{(1)} > \dots > \tilde{\lambda}_1^{(n)}$ são os autovalores de $\mathbb{I}_n + \tilde{\Omega}\Omega(\phi_1, \sigma^2)^{-1}$.

(b) Para $X_t^* \sim AR(1) : \beta^* = (\phi_1^*, \sigma_*^2)$ e $X_t \sim AR(2) : \beta = (\phi_1, \phi_2, \sigma^2)$, a expressão resultante de aplicar os seguintes termos em (3.12):

$$\log |\Omega(\beta^*)^{-1}| + \log |\Omega(\beta)^{-1}| = -n [\log(\sigma^2) + \log(\sigma^2)] + \log(1 - \phi_1^{*2}) + 2 \log(1 + \phi_2) \\ + \log[(1 - \phi_2)^2 - \phi_1^2]$$

e

$$\log |\Omega(\beta^*) + \Omega(\beta)| = 2 \log(1 + \phi_2) + \log[(1 - \phi_2^2) - \phi_1^2] - n \log(\sigma^2) \\ - \sum_{i=1}^n \log \lambda_{12}^{(i)},$$

em que $\lambda_{12}^{(1)} > \dots > \lambda_{12}^{(n)}$ são os autovalores de $\mathbb{I}_n + \Omega(\phi_1^*, \sigma_*^2)\Omega(\phi_1, \phi_2, \sigma^2)^{-1}$.

Em particular, a Equação (3.13) e a medida $\Delta_H(\phi_1, \sigma^2)$ definida sobre ela sugerem novos critérios de seleção e serão utilizadas na computação da probabilidade de *Overfitting* dos novos critérios propostos na Seção 4.3 do Capítulo 4. .

3.2.2 Divergências entre processos MA(1) e ARMA(1,1)

Teorema 7. *Sejam*

$$\{X_t^*; t = 1, \dots, n\} \sim MA(1) : \beta^* = (\theta_1^*, \sigma_*^2) \text{ ou } \sim ARMA(1, 1) : \beta^* = (\phi_1^*, \theta_1^*, \sigma_*^2)$$

com matriz de covariâncias $\Omega(\beta^*)$ e

$$\{X_t; t = 1, \dots, n\} \sim MA(1) : \beta = (\theta_1, \sigma^2) \text{ ou } \sim ARMA(1, 1) : \beta = (\phi_1, \theta_1, \sigma^2)$$

com matriz de covariâncias $\Omega(\beta)$. Assuma também que $\mu_t^* = E(X_t | X_{t-1}^*, \dots, X_{t-k+1}^*)$ e $\mu_t = E(X_t | X_{t-1}, \dots, X_{t-p+1})$ como flexibilização da condição de estacionaridade dada em (2) afim de incluir descrição de modelos econométricos. A divergência de Kullback-Leibler entre os processos X_t^* e X_t , $\Delta_{KL}(\beta^* || \beta)$, é dada por

$$\begin{aligned} & \frac{1}{2} \left\{ \overbrace{\nu(\beta^*, \beta) \left[\sum_{r=1}^n K^{*r,r} + 2 \sum_{1 \leq r < s \leq n} K_{r,s} K^{*r,s} \right]}^{tr[\Omega^{-1}(\beta^*)\Omega(\beta)]} - n \right. \\ & \left. \frac{[\mu(\beta) - \mu(\beta^*)]^\top \Omega^{-1}(\beta^*) [\mu(\beta) - \mu(\beta^*)]}{+ \nu(\beta^*, 1) \left\{ \sum_{r=p+1}^n K^{*r,r} [\hat{x}_r(\beta^*) - \hat{x}_r(\beta)]^2 + 2 \sum_{1 \leq r < s \leq n} K_{r,s} K^{*r,s} [\hat{x}_r(\beta^*) - \hat{x}_r(\beta)] [\hat{x}_s(\beta^*) - \hat{x}_s(\beta)] \right\}} \right. \\ & \left. + \log \frac{|\Omega(\beta^*)|}{|\Omega(\beta)|} \right\}, \end{aligned} \quad (3.14)$$

em que

$$\begin{aligned} \nu(\beta^*, \beta) &= \begin{cases} \frac{\sigma_*^2 (1+\theta_1^2)}{\sigma_*^2 (1+\theta_1^{*2})}, & \text{para } MA(1), \\ \frac{\sigma^2 \left[1 + \frac{(\theta_1 + \phi_1)^2}{1 - \phi_1^2} \right]}{\sigma_*^2 \left[1 + \frac{(\theta_1^* + \phi_1^*)^2}{1 - \phi_1^{*2}} \right]}, & \text{para } ARMA(1, 1), \end{cases} \\ \log \frac{|\Omega(\beta^*)|}{|\Omega(\beta)|} &= \begin{cases} \log \frac{(1 - \theta_1^{*2(n-1)})}{(1 - \theta_1^{2(n-1)})} + \log \frac{(1 - \theta_1^2)}{(1 - \theta_1^{*2})}, & \text{para } MA(1), \\ \log \frac{[(1 - \theta_1 \phi_1)^2 - (\theta_1 - \phi_1)^2 \theta_1^{2n}]}{[(1 - \theta_1^* \phi_1^*)^2 - (\theta_1^* - \phi_1^*)^2 \theta_1^{*2n}]}, & \text{para } ARMA(1, 1), \end{cases} \end{aligned}$$

$\hat{X}_{r+1}(\beta) = \theta_1 [X_r^* - \hat{X}_r(\beta)]/v_{r-1}$ com $v_0 = 1 + \theta_1^2$ e $v_{r+1} = 1 + \theta_1^2 - (\theta_1^2/v_r)$ (para o caso $MA(1)$), $\hat{X}_{r+1}(\beta) = \phi_1 X_r^* + \theta_{r,1} [X_r^* - \hat{X}_r(\beta)]$ com $v_0 = (1 + 2\phi_1\theta_1 + \theta_1^2)/(1 - \phi_1^2)$, $\theta_{r,1} = \theta_1/v_{r-1}$ e $v_r = 1 + \theta_1^2 - (\theta_1^2/v_{r-1})$. $K_{r,s}$ é a entrada (r, s) da matriz de correlação de $\{X_t\}$ representada por

$$K_{r,s} = \begin{cases} \text{Para } MA(1) : & \frac{\theta}{1+\theta^2}, \text{ se } h = 1, \\ & 0, \text{ se } h > 1, \\ \text{Para } ARMA(1, 1) : & \frac{(1+\phi\theta)(\phi+\theta)}{1+2\theta\phi} \phi^{h-1} \text{ se } h \geq 1, \end{cases}$$

e

$$K^{r,s} = \begin{cases} \frac{\theta_1^{s-r} (1 - \theta_1^{2r} (1 - \theta_1^{2(n+1-s)}))}{(1 + \theta_1^2) (1 - \theta_1^2) (1 - \theta_1^{2(n+1)})}, & \text{para } MA(1), \\ \frac{m_{rs}}{\left[1 + \frac{(\theta_1 + \phi_1)^2}{1 - \phi_1^2} \right]}, & \text{para } ARMA(1, 1), \end{cases}$$

com m_{rs} definido nas Equações (2.9)-(2.10), é a entrada (r, s) da matriz de correlação inversa de $\{X_t\}$. Baseando-se sobre o resultado acima, As divergências tipo KL, $\Delta_{TKL}(\beta^* \parallel \beta)$, e segundo Jensen-Shannon (JS), $\Delta_{JS}(\beta, \beta^*) = \Delta_{KL}(\beta^* \parallel \beta) + \Delta_{TKL}(\beta \parallel \beta^*)$, são facilmente obtidas.

Teorema 8. *Sejam*

$$\{X_t^*; t = 1, \dots, n\} \sim MA(1) : \beta^* = (\theta_1^*, \sigma_*^2) \text{ ou } \sim ARMA(1, 1) : \beta^* = (\phi_1^*, \theta_1^*, \sigma_*^2)$$

com matriz de covariâncias $\Omega(\beta^*)$ e

$$\{X_t; t = 1, \dots, n\} \sim MA(1) : \beta = (\theta_1, \sigma^2) \text{ ou } \sim ARMA(1, 1) : \beta = (\phi_1, \theta_1, \sigma^2)$$

com matriz de covariâncias $\Omega(\beta)$. Assuma também que $\mu_t^* = E(X_t | X_{t-1}^*, \dots, X_{t-k+1}^*)$ e $\mu_t = E(X_t | X_{t-1}, \dots, X_{t-p+1})$. A distância de Bhattacharyya entre os processos X_t^* e X_t , $\Delta_B(\beta^*, \beta)$, é dada por

$$\begin{aligned} & \frac{1}{4} [B(\beta)^\top \mu_{11} - B_{11}(\beta, \beta^*)^\top \mu_{11}]^\top [B(\beta)^\top \mu_{11} - B_{11}(\beta, \beta^*)^\top \mu_{11}] \\ & + \frac{1}{2} \left\{ -n \log 2 - \frac{n}{2} [\log(\sigma_*^2) - \log(\sigma^2)] + \xi_{\mathcal{M}}(\beta, \beta^*) + \sum_{i=1}^n \log \lambda_{\mathcal{M}}^{(i)} \right\}, \end{aligned} \quad (3.15)$$

em que $A(\beta^*)$ e $B(\beta)$ são resultantes das fatorações de Cholesky $\Omega(\beta^*) = A(\beta^*)A(\beta^*)^\top$ e $\Omega(\beta) = B(\beta)B(\beta)^\top$, $\mathcal{H} = A(\beta^*)^{-1}B(\beta)$, $\mu_{11} = [\mu(\beta) - \mu(\beta^*)]$ com $\mu(\beta) = [\dots | \hat{X}_r(\beta) | \dots]^\top$ e $\mu(\beta^*) = [\dots | \hat{X}_r(\beta^*) | \dots]^\top$ tal como definido no Teorema 7,

$$B_{11}(\beta, \beta^*) = \mathcal{H}^\top [\mathbb{I}_n + \mathcal{H} \mathcal{H}^\top]^{-1} \mathcal{H} A(\beta^*) \text{ e}$$

$$\xi_{\mathcal{M}} = \begin{cases} 3 \log \frac{(1-\theta_1^{2(n+1)})}{(1-\theta_1^2)} + \log \frac{(1-\theta_1^{*2(n+1)})}{(1-\theta_1^{*2})}, & \text{para } \mathcal{M} = MA(1) \\ 3 \log \frac{(1-\phi_1^2)(1-\theta_1^2)}{((1-\theta_1\phi_1)^2 - (\theta_1-\phi_1)^2 \theta_1^{2n})} + \log \frac{(1-\phi_1^{*2})(1-\theta_1^{*2})}{((1-\theta_1^*\phi_1^*)^2 - (\theta_1^*-\phi_1^*)^2 \theta_1^{*2n})}, & \text{para } \mathcal{M} = ARMA(1,1) \end{cases}$$

Baseando-se sobre o resultado acima, A distância de Hellinger (H), $\Delta_H(\beta^*, \beta) = 1 - \exp(-\Delta_B)$ é facilmente obtida.

Capítulo 4

Resultados e discussões

4.1 Avaliação dos métodos de estimação clássicos

Nesta seção, apresenta-se um estudo comparativo via simulação de Monte Carlo com 1000 réplicas, do desempenho de três métodos de estimação:

- Máxima Verossimilhança Condicional - MVC;
- Máxima Verossimilhança Não Condicional - MVNC;
- Mínimos Quadrados Não Condicional - MQNC.

Quantificação foi realizada sob três sub modelos da classe ARMA: Para $Z_t \sim \text{RBG}(0, \sigma^2)$,

- AR(1) : $\beta = [\phi, \sigma^2]$: $X_t = \phi X_{t-1} + Z_t$;
- MA(1) : $\beta = [\theta, \sigma^2]$: $X_t = Z_t + \theta Z_{t-1}$;
- ARMA(1,1) : $\beta = [\phi, \theta, \sigma^2]$: $X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1}$;

A discretização da região de estabilidade referentes aos estudos de simulações nesta dissertação para submodelos na classe ARMA:

- (a) AR(1): Assumindo $\phi = i$, com $i = -0.99, -0.98, \dots, 0.98, 0.99$, totalizando 199 pontos satisfazendo $|\phi| < 1$.
- (b) MA(1): Assumindo $\theta = i$, com $i = -0.99, -0.98, \dots, 0.98, 0.99$, totalizando 199 pontos satisfazendo $|\theta| < 1$.
- (c) ARMA(1,1): Assumindo $\phi = (-0.9, -0.7, -0.5, -0.3, -0.1, 0.1, 0.3, 0.5, 0.7, 0.9)$ e $\theta = (-0.9, -0.7, -0.5, -0.3, -0.1, 0.1, 0.3, 0.5, 0.7, 0.9)$.

Como figuras de mérito, utilizamos:

critérios clássicos: viés e erro quadrático médio (EQM);

critérios de informação: medidas de divergências.

Sobre este último item, como uma avaliação complementar aos desempenhos dos estimadores MVC, MNC e MQNC, foi utilizado os resultados derivados nos Corolários 2 e 3— Δ_{KL} , Δ_{TKL} , Δ_{JS} , Δ_B e Δ_H —para avaliar o impacto das estimativas sobre a densidade de probabilidade conjunta dos processos $AR(1)$, $MA(1)$ e $ARMA(1,1)$. As medidas foram utilizadas de duas formas:

- *forma reduzida*, em que os termos

$$T_1(\beta, \beta^*) = (\mu(\beta) - \mu(\beta^*))^\top \Omega(\beta^*) (\mu(\beta) - \mu(\beta^*))$$

e

$$T_2(\beta, \beta^*) = (\mu(\beta) - \mu(\beta^*))^\top \left[\frac{\Omega(\beta^*) + \Omega(\beta)}{2} \right]^{-1} (\mu(\beta) - \mu(\beta^*)).$$

são desconsiderados de Δ_{KL} (e suas variação, Δ_{TKL} e Δ_{JS}) e de Δ_B (e sua variação Δ_H), respectivamente.

- *forma natural*, em que as expressões são utilizadas tais como nos corolários.

Nesta dissertação, apresentamos as divergências nas duas formas, reduzida e natural(completa). O uso da forma reduzida está no contexto de processos estacionários, no qual será admitida esta forma para os estudos de simulação das seções seguintes deste capítulo. Entretanto, pode-se utilizar a forma completa das medidas para modelos econométricos, assumindo que as médias dos processos não são iguais ao longo do tempo.

Trabalhos separados têm sido publicados considerando a comparação naqueles submodelos por EQM e viés. Neste aspecto, trabalhos bem conhecidos ((CRAIG; NEWBOLD, 1980); (KANG, 1975)) apontam os estimadores de MVNC como mais adequado em relação a outros critérios de estimação, dentre eles os estimadores de máxima verossimilhança condicional e mínimos quadrados, em particular quando as séries tem poucas observações e/ou os parâmetros aproximam-se da fronteira de estabilidade. Trabalhos recentes tem sido propostos com o objetivo de obter aproximações para a verossimilhança exata, em que seus desempenhos são comparados com métodos de estimação clássicos via viés e EQM (MAURICIO, 1995, 2002; MCLEOD; ZHANG, 2008).

Outra importante contribuição no surgimento de métodos alternativos de estimação na classe de modelos ARMA gaussianos é proposta por Yu (2004), que apresentou um procedimento baseado em Mínima Função Característica e adicionalmente realiza estudos de comparação entre os métodos clássicos. Estes estudos intencionam capturar o efeito das estimativas sobre os ajustes das densidades nos sub-modelos. Um estudo neste sentido é

de fundamental importância nas etapas de inferência estatística, identificação de modelos e diagnóstico, pois estas etapas assumem que a estimação é adequada e envolvem medidas baseadas nas densidades ajustadas.

4.1.1 Estimação no modelo $AR(1)$

Inicialmente, a comparação é feita em modelos autoregressivos. Na Figura 4.1, apresenta-se a média das estimativas dos parâmetros ϕ e σ^2 para os métodos MQNC, MVC e MVNC nos tamanhos de amostras $n = 15, 35$ e 100 . Para $n = 15$, segundo a estimação de ϕ , verifica-se que o método MQNC tem melhor aproximação da reta $y = f(x) = x$ se mostrando menos viesada para estimar ϕ tal que $|\phi| \uparrow 1$ quando comparado com os métodos MVNC e MVC. Para $|\phi| \downarrow 0$, os métodos não apresentam diferença significativa. Com o aumento da amostra, os métodos de estimação são assintoticamente equivalentes em viés. A seguinte relação entre vieses, $B(\hat{\theta}) = \hat{\theta} - \theta$ para θ um parâmetro escalar qualquer, pode ser verificada:

$$|B(\hat{\phi}_{MQNC})| \leq |B(\hat{\phi}_{MVNC})| \leq |B(\hat{\phi}_{MVC})|.$$

Sobre o comportamento da média das estimativas do parâmetro σ^2 . Em geral a centralidade assintótica, é satisfeita para n grande como esperado. Para $|\phi| \downarrow 0$, não há diferença significativa entre as estimações do parâmetro σ^2 . Por outro lado, para $|\phi| \uparrow 1$ (mais fortemente para tamanhos pequenos), vê-se os métodos satisfazem a seguinte relação entre vieses:

$$|B(\hat{\sigma}^2_{MQNC})| \leq |B(\hat{\sigma}^2_{MVNC})| \leq |B(\hat{\sigma}^2_{MVC})|.$$

O $EQM - EQM(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = \text{Var}(\hat{\theta}) + B(\hat{\theta})^2$ —dos estimadores para ϕ e σ^2 em diferentes tamanhos de amostras é apresentado na Figura 4.2. Para amostras pequenas $n = 15$ e 35 (como apresentado na Figura 4.2), o método de estimação MVNC apresentou o menor EQM das estimativas de ϕ , enquanto que MQNC obteve o pior EQM para ϕ próximos de zero. Na região próxima a fronteira de estabilidade, o menor EQM foi do método MQNC. É importante citar que o EQM das estimativas diminuem a medida que o tamanho da amostra aumenta e não existem diferenças significativas entre os métodos para tamanho de amostra $n = 100$. Os resultados encontrados obedecem à seguinte relação de desigualdades:

$$\begin{cases} EQM(\hat{\phi}_{MVNC}) \leq EQM(\hat{\phi}_{MVC}) \leq EQM(\hat{\phi}_{MQNC}), & \text{para } |\phi| \leq 0.675 \\ EQM(\hat{\phi}_{MQNC}) \leq EQM(\hat{\phi}_{MVNC}) \leq EQM(\hat{\phi}_{MVC}), & \text{para } |\phi| > 0.675. \end{cases}$$

Em relação ao EQM do parâmetro σ^2 como visto na Figura 4.2, os métodos de estimação não apresentam diferenças significativas, considerando tamanhos de amostras $n = 15, 35$ e 100 .

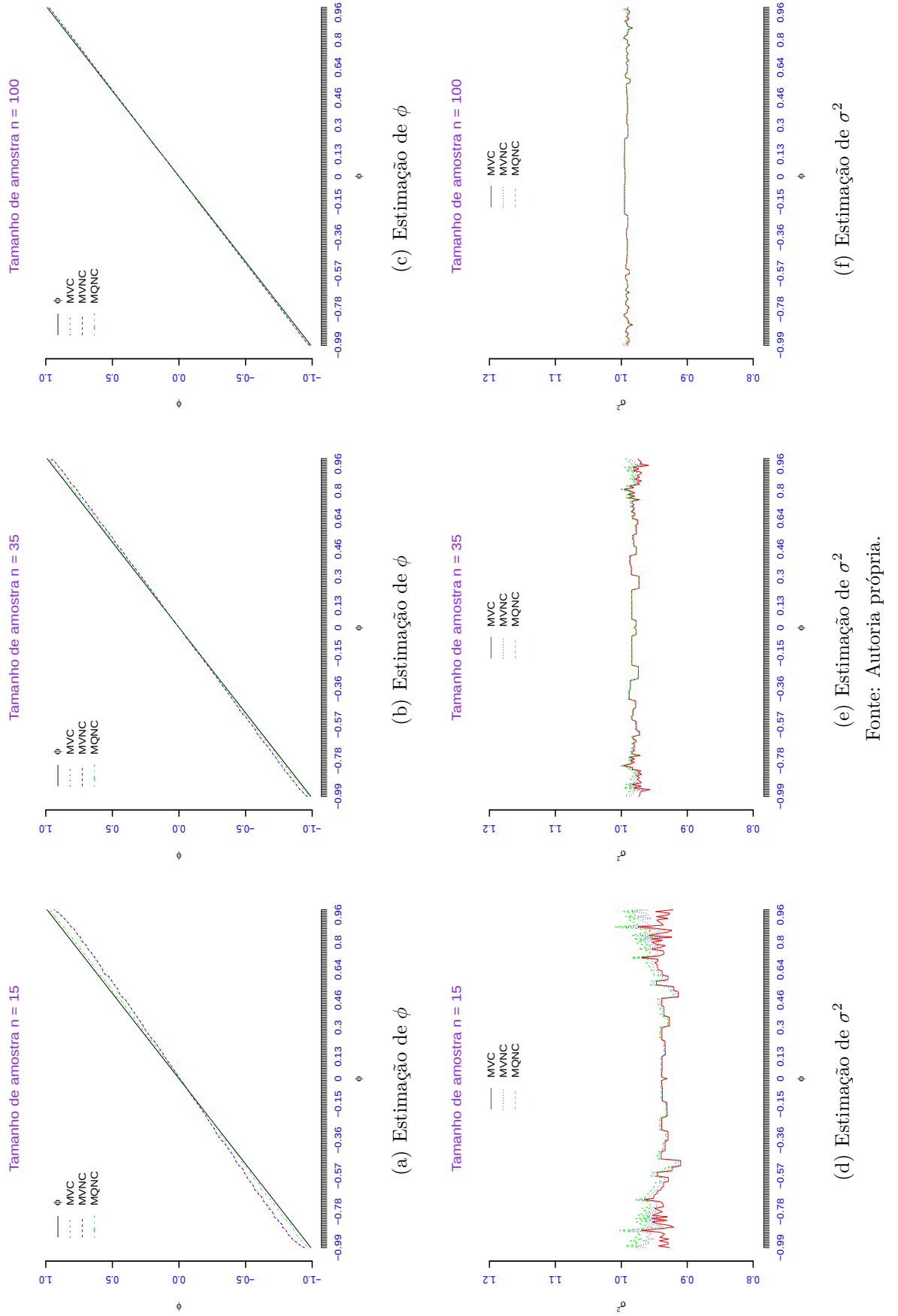
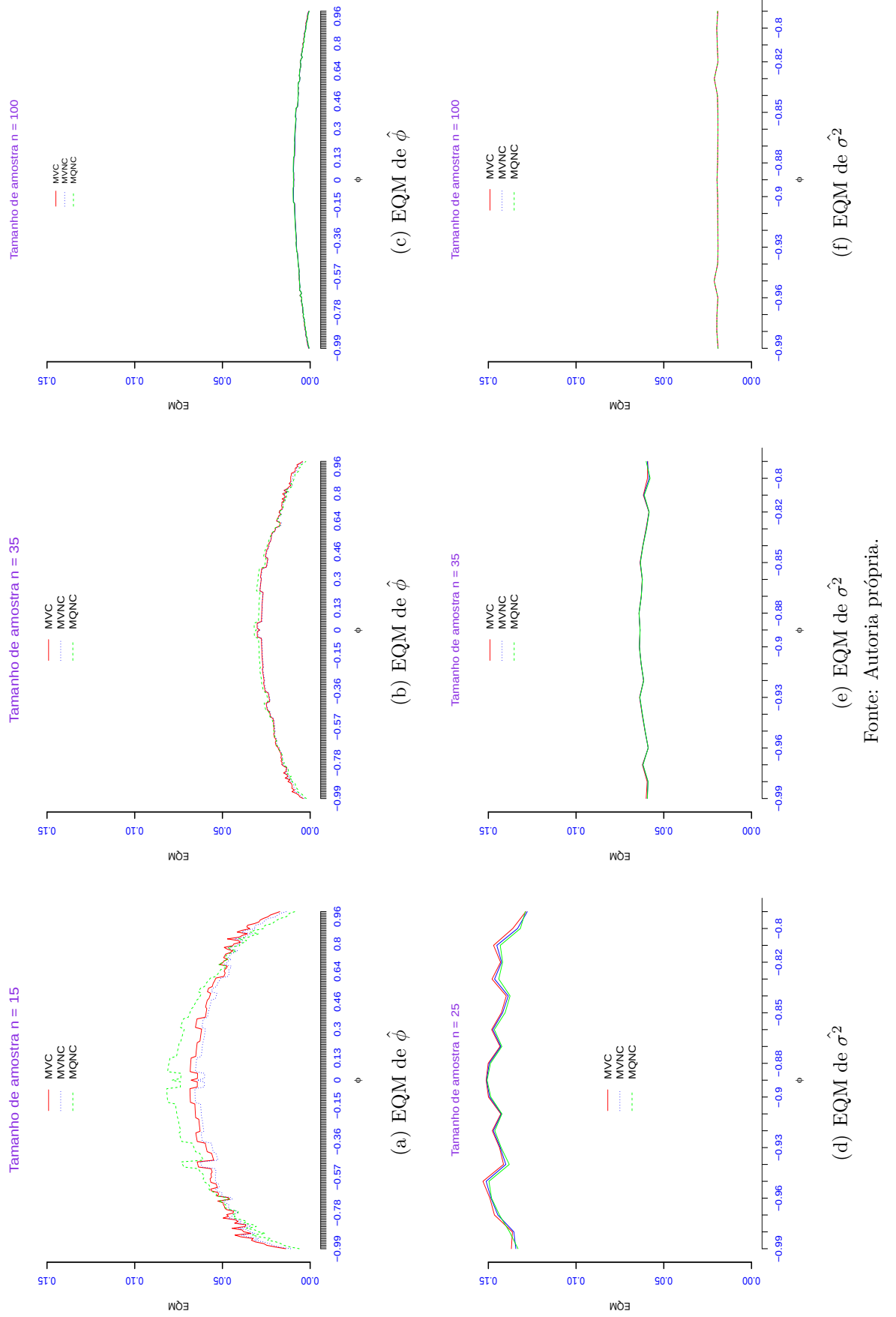
Figura 4.1: Estimação dos parâmetros ϕ e σ^2 sob a região de estabilidade do AR(1).

Figura 4.2: EQM dos estimadores de ϕ e σ^2 sob a região de estabilidade do AR(1).

Fonte: Autoria própria.

Avaliação via Δ_{KL} e Δ_{JS} :

Forma Reduzida: Nas Figuras 4.3 e 4.5, apresenta-se respectivamente, os valores das medidas de *Kullback-Leibler* e de *Jensen-Shannon* com formas reduzidas para diferentes tamanhos de amostras. Vale salientar que o método que apresenta o menor valor das medidas de distância é considerado o melhor em termos da distorção da verossimilhança conjunta. Os valores das divergências para os métodos de MQNC e MVC crescem com o aumento em módulo de $|\phi|$ para $\phi > 0.35$ e $\phi < -0.25$, isto implica que nessa região estes métodos tem desempenho insatisfatório, pois apresentam valores altos. Entretanto, esse crescimento é atenuado com o aumento do tamanho amostral; por exemplo, para $n = 100$, a relação de crescimento da sentença anterior é verificada para $|\phi| > -0.9$ enquanto que, no intervalo $-0.9 < \phi < 0.9$, os métodos de estimação são equivalentes (como apresentado nas Figuras 4.3 e 4.5). Por outro lado, O método de MVNC apresentou um comportamento robusto na fronteira de estabilidade para todos os tamanhos amostrais estudados, em que os valores das divergência foram bem inferiores em comparação com os métodos MVC e MQNC. Este resultado apresenta evidências em dois sentidos:

(1) as divergências Δ_{KL} e Δ_{JS} combinadas à estimação por MVNC podem ser ferramentas robustas para processos que se aproximem de um passeio aleatório, $X_t = X_{t-1} + Z_t$, e (2) o método de MVNC é o mais indicado entre termos de influência da estimação sobre a densidade conjunta de um processo autoregressivo gaussiano via Δ_{KL} e Δ_{JS} .

Forma Completa: Os resultados considerando a estrutura completa da *Kullback-Leibler* e *Jensen-Shannon*, são apresentados, respectivamente, nas Figuras 4.3 e 4.5. Nota-se que esta forma tem comportamento semelhante com a forma reduzida. Para $|\phi|$ próximo de 1, os métodos sempre obedecem a seguinte relação para tamanhos de amostras $n = 15$, $n = 35$ e $n = 100$:

$$\Delta_{KL \text{ ou } JS}(\hat{\phi}_{MVNC}) \leq \Delta_{KL \text{ ou } JS}(\hat{\phi}_{MVC}) \leq \Delta_{KL \text{ ou } JS}(\hat{\phi}_{MQNC}).$$

Avaliação via Δ_{TKL} :

Forma Reduzida: Na Figura 4.4, encontram-se os valores da medida Tipo *Kullback-Leibler* na forma reduzida para alguns tamanhos de amostras. Em geral, verifica-se a seguinte relação:

$$\Delta_{TKL}(\hat{\phi}_{MVNC}) \leq \Delta_{TKL}(\hat{\phi}_{MVC}) \leq \Delta_{TKL}(\hat{\phi}_{MQNC}).$$

Independentemente da estimação, os valores das divergências aumentam quando $|\phi|$ se aproxima de um. Em particular para $n = 15$, esse crescimento é mais expressivo quando $\phi > 0.70$ e $\phi < -0.70$. Esse aumento é atenuado para tamanhos de amostras grandes

(conforme a Figura 4.4. A estimação por MVNC se mostrou mais robusta que os outros métodos de estimação na fronteira de estabilidade para todos os tamanhos amostrais estudados, visto que apresentou os menores valores da medida Δ_{TKL} .

Estamos em posição de concluir que:

A distância Tipo Kullback-Leibler indica que os estimadores MVNC como mais adequados, principalmente para processos sob correlação forte ($|\phi| \uparrow 1$), com poucas observações.

Forma Completa: Para a estrutura completa Tipo *Kullback-Leibler*, verifica-se que: Sejam os mapeamentos $\delta^-(n) : \mathbb{N} \mapsto (-1, 0)$ and $\delta^+(n) : \mathbb{N} \mapsto (0, 1)$,

$$\begin{cases} \Delta_{TKL}(\hat{\phi}_{MVNC}) \leq \Delta_{TKL}(\hat{\phi}_{MVC}) \leq \Delta_{TKL}(\hat{\phi}_{MQNC}), & \text{para } \phi < \delta^-(n) \text{ ou } \phi > \delta^+(n), \\ \Delta_{TKL}(\hat{\phi}_{MQNC}) \leq \Delta_{TKL}(\hat{\phi}_{MVNC}) \leq \Delta_{TKL}(\hat{\phi}_{MVC}), & \text{para } \phi > \delta^-(n) \text{ e } \phi < \delta^+(n), \end{cases}$$

em que $\delta^-(15) = -0.7$, $\delta^-(25) = -0.8$, $\delta^+(15) = 0.7$ e $\delta^+(25) = 0.8$. A igualdade entre os membros das desigualdades é alcançada para n suficientemente grande, indicado que não diferença entre os métodos de estimação assintoticamente.

Avaliação via Δ_B e Δ_H :

Forma Reduzida: As Figuras 4.6 e 4.7 apresentam os valores das medidas de *Bhattacharyya* e de *Hellinger* com formas reduzidas para alguns tamanhos de amostras. Em ambos os caso, vê-se que a estimação por MVNC obteve o melhor desempenho dentre os métodos de estimação, visto que minimiza a medida de divergência em comparação com os métodos MVC e MQNC. Assim, segue-se o seguinte comportamento:

$$\Delta_{B \text{ ou } H}(\hat{\phi}_{MVNC}) \leq \Delta_{B \text{ ou } H}(\hat{\phi}_{MVC}) \leq \Delta_{B \text{ ou } H}(\hat{\phi}_{MQNC}).$$

Em particular, para $n = 15$ e $\phi > 0.5$ ou $\phi < -0.5$, a diferença é mais expressiva. Para tamanhos de amostras grandes, os métodos tendem a apresentar desempenhos similares; mas, mesmo nestes casos, o impacto das estimativas MVC e MQNC—para $|\phi| \uparrow 1$ sobre o ajuste via Δ_B ou Δ_H —é bem expressivo se comparado àquele induzido da estimação por MVNC.

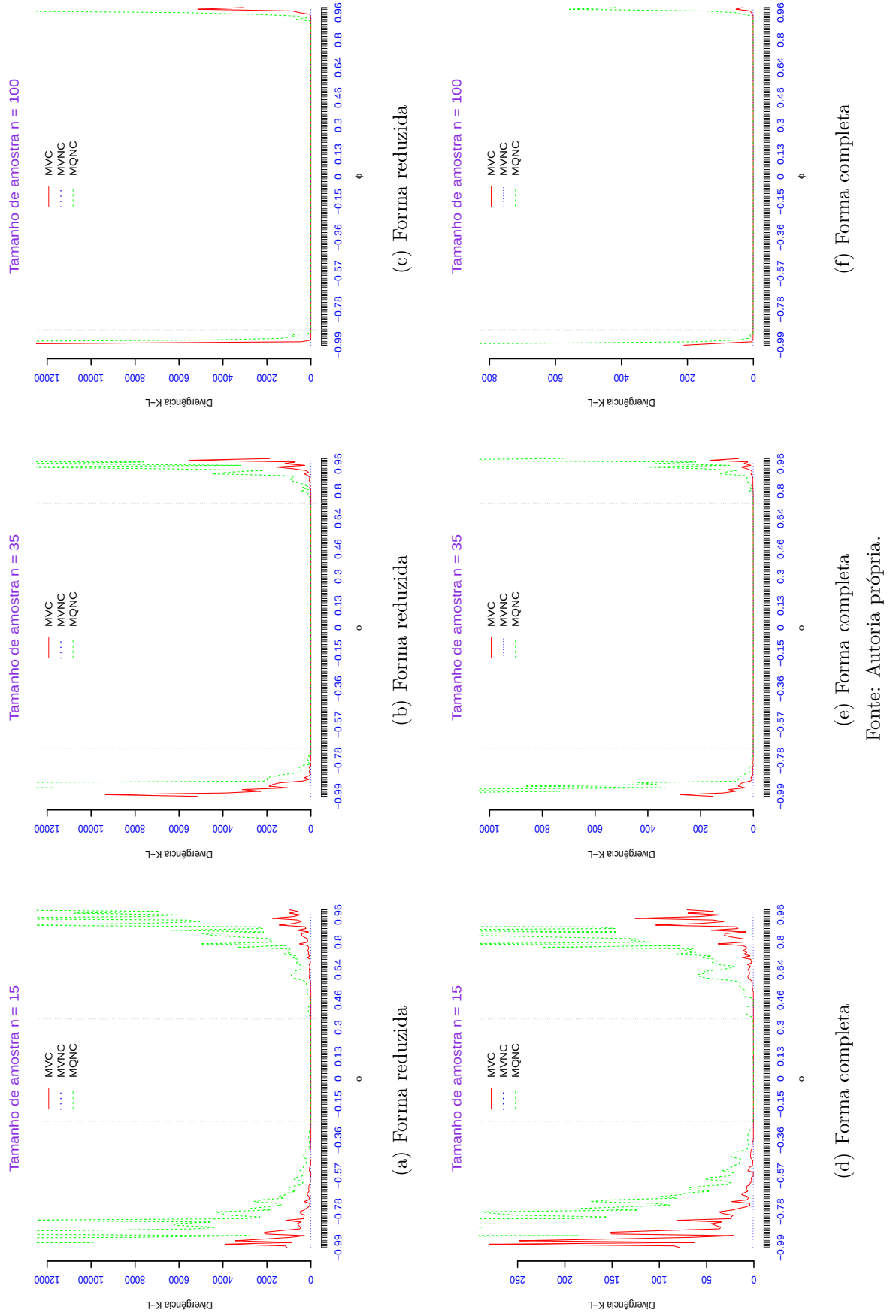
Desta maneira, conclui-se que

(1) as divergências Δ_B e Δ_H combinadas à estimação por MVNC podem ser ferramentas robustas para processos do tipo $X_t = X_{t-1} + Z_t$ e (2) o método de MVNC é o mais indicado entre termos de influência da estimação sobre a densidade conjunta de um processo autoregressivo gaussiano via Δ_B e Δ_H .

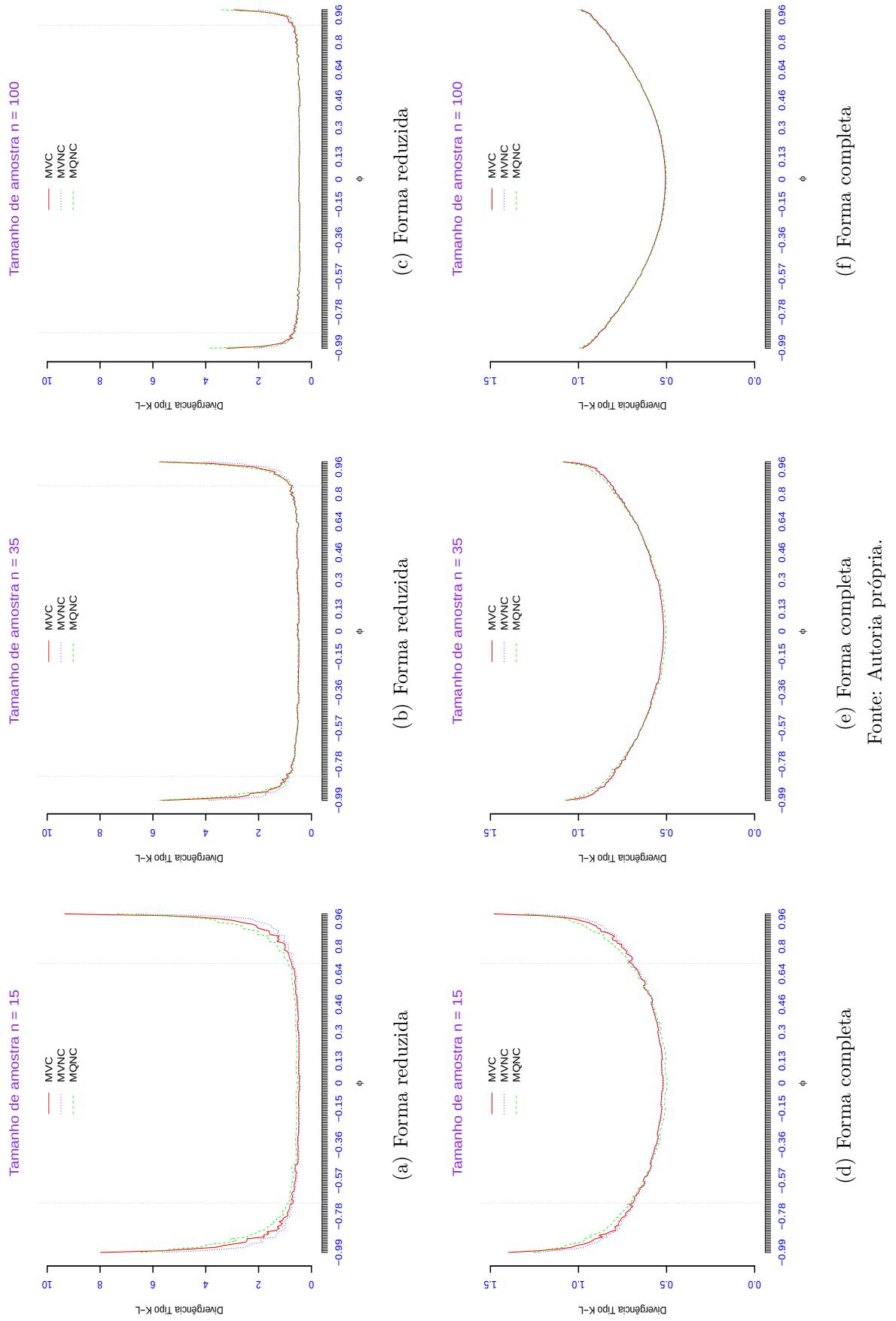
Forma Completa: Para a estrutura completa *Bhattacharyya* e *Hellinger*, verifica-se que: Sejam os mapeamentos $\delta^-(n) : \mathbb{N} \mapsto (-1, 0)$ and $\delta^+(n) : \mathbb{N} \mapsto (0, 1)$,

$$\begin{cases} \Delta_{\text{B ou H}}(\hat{\phi}_{MVNC}) \leq \Delta_{\text{B ou H}}(\hat{\phi}_{MVC}) \leq \Delta_{\text{B ou H}}(\hat{\phi}_{MQNC}), & \text{para } \phi < \delta^-(n) \text{ ou } \phi > \delta^+(n), \\ \Delta_{\text{B ou H}}(\hat{\phi}_{MQNC}) \leq \Delta_{\text{B ou H}}(\hat{\phi}_{MVNC}) \leq \Delta_{\text{B ou H}}(\hat{\phi}_{MVC}), & \text{para } \phi > \delta^-(n) \text{ e } \phi < \delta^+(n), \end{cases}$$

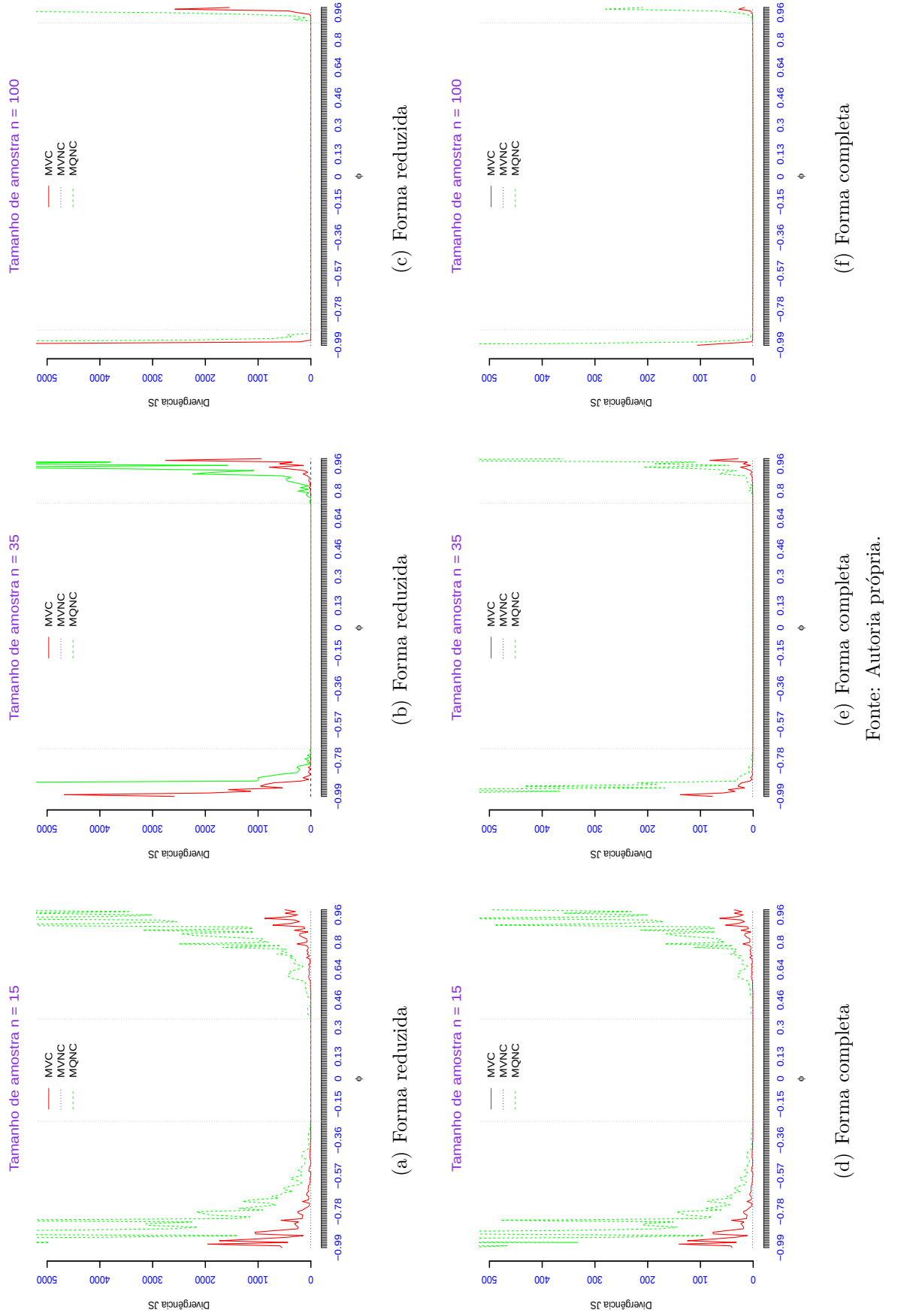
em que $\delta^-(15) = -0.7$, $\delta^-(25) = -0.8$, $\delta^+(15) = 0.7$ e $\delta^+(25) = 0.8$. A igualdade entre os membros das desigualdades é alcançada para n suficientemente grande, indicando que não há diferença entre os métodos de estimação assintoticamente.

Figura 4.3: Média da divergência *Kullback-Leibler* sob a região de estabilidade do AR(1)

Fonte: Autoria própria.

Figura 4.4: Média da divergência Tipo *Kullback-Leibler* sob a região de estabilidade do AR(1).

Fonte: Autoria própria.

Figura 4.5: Média da divergência *Jensen-Shannon* sob os pontos discretizados da região de estabilidade para o processo AR(1).

Fonte: Autoria própria.

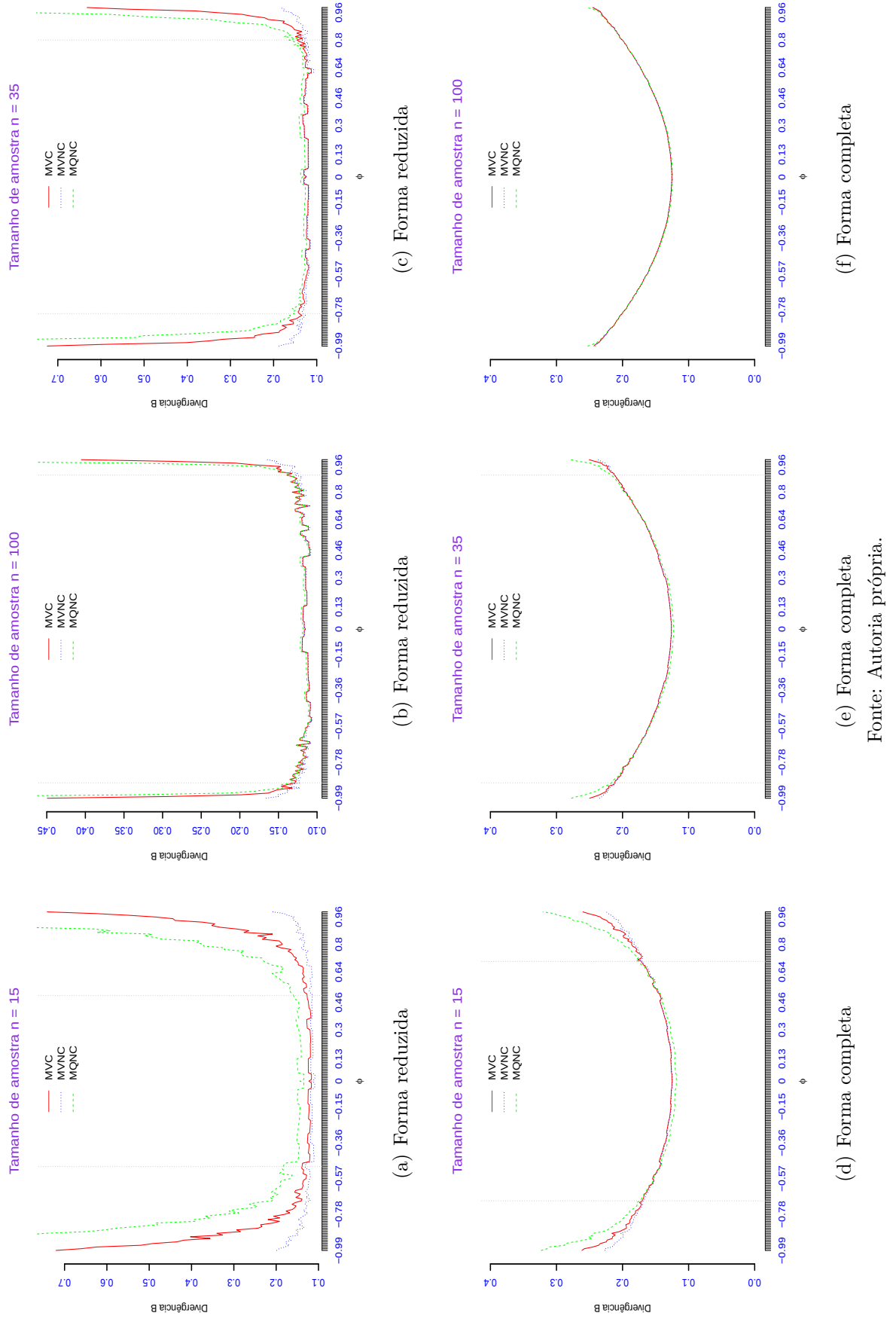
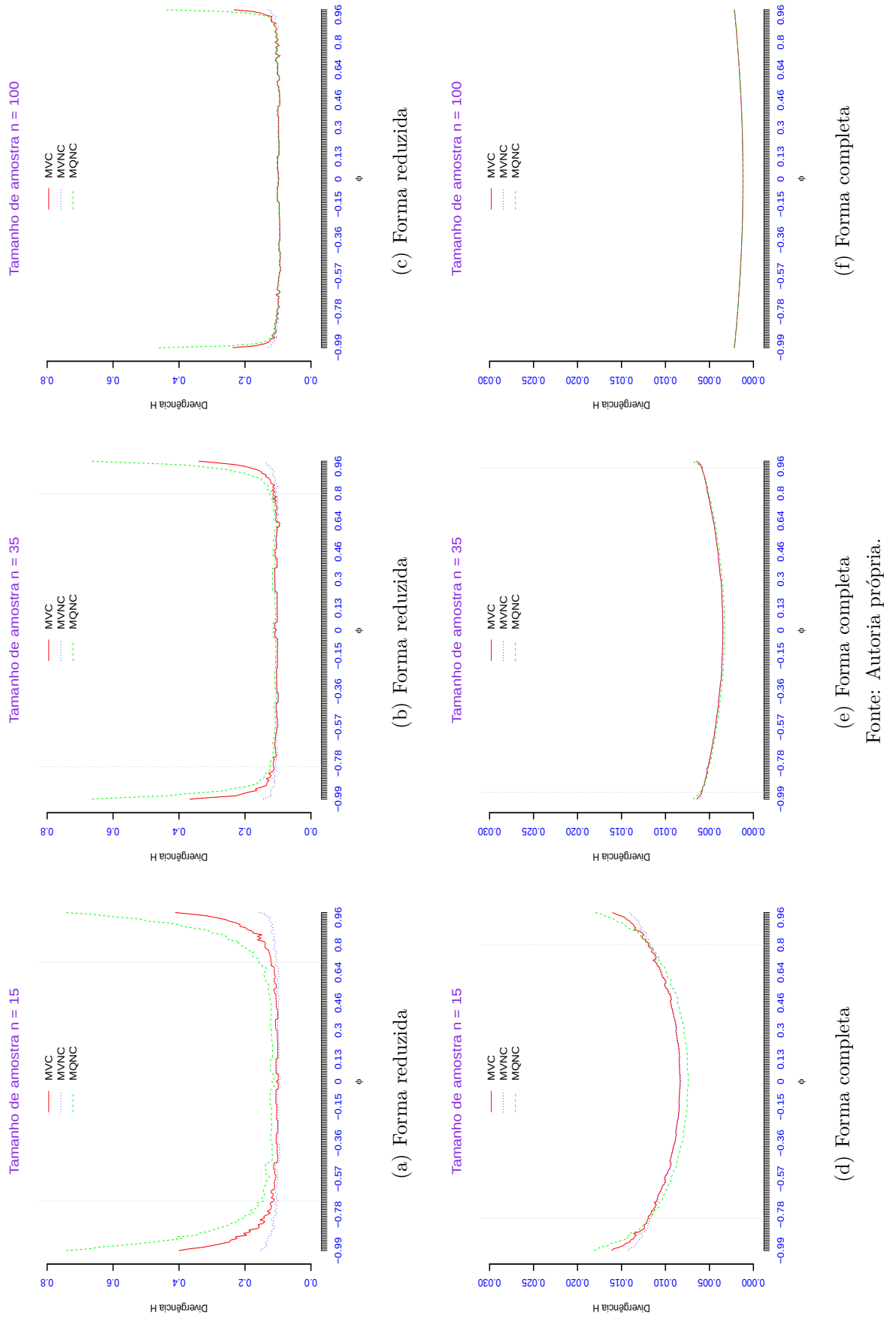
Figura 4.6: Média da divergência *Bhattacharyya* sob os pontos discretizados da região de estabilidade para o processo AR(1).

Figura 4.7: Média da divergência *Hellinger* sob os pontos discretizados da região de estabilidade para o processo AR(1).

Fonte: Autoria própria.

4.1.2 Estimação no modelo MA(1)

Nesta seção, avalia-se a estimação em processos MA(1). As médias das estimações do parâmetro θ de *médias móveis*, são apresentadas na Figura 4.8. Para tamanho de amostra $n = 15$, as estimações de MQNC é menos viesada a θ do que aquelas obtidas pelos métodos MVNC e MVC:

$$|B(\hat{\theta}_{MQNC})| \leq |B(\hat{\theta}_{MVNC})| \leq |B(\hat{\theta}_{MVC})|.$$

Para valores de θ próximos de zero, o desempenho dos métodos de estimação são similares. Um comportamento incomum pode ser identificado a partir da estimação do parâmetro θ , nota-se que para uma determinada região onde $0.7 < \theta < 0.9 \cap -0.9 < \theta < -0.7$, o viés da estimação dos métodos não diminui com o tamanho da amostra. Assim, como sugestão de trabalhos futuros, deve-se aplicar um estudo de investigação mais profundo nessa região de estabilidade.

A estimação do parâmetro σ^2 (que nesta dissertação será considerado como um parâmetro de *perturbação*, obedecendo à definição de Cordeiro (1999)) é apresentada na Figura 4.8. Note que, quando $|\theta| \uparrow 1$, o valor de σ^2 é superestimado, ultrapassando o valor quinze quando deveria ser um. Por outro lado, quando θ é próximo de zero, os métodos tem boa aproximação do valor verdadeiro de σ^2 . Em resumo, a estimação deste parâmetro parece funcionar como um bloqueio quando o $|\theta| \uparrow 1$. Neste caso, assumir a variância do ruído branco como um poderá melhorar os resultados de identificação (como será discutido adiante).

Na Figura 4.9, apresenta-se o EQM dos estimadores de θ . Para tamanho de amostra $n = 15$ e $\theta < -0.7$ e $\theta > 0.7$, (como apresentado na Figura 4.9), a seguinte relação é satisfeita:

$$\begin{cases} \text{EQM}(\hat{\theta}_{MQNC}) \leq \text{EQM}(\hat{\theta}_{MVNC}) \leq \text{EQM}(\hat{\theta}_{MVC}), & \text{para } \theta < -0.7 \text{ e } \theta > 0.7 \\ \text{EQM}(\hat{\theta}_{MVNC}) \leq \text{EQM}(\hat{\theta}_{MVC}) \leq \text{EQM}(\hat{\theta}_{MQNC}), & \text{para } |\theta| < 0.7 \end{cases}$$

Entretanto, com o aumento do tamanho amostral, os métodos tendem a apresentar o mesmo EQM.

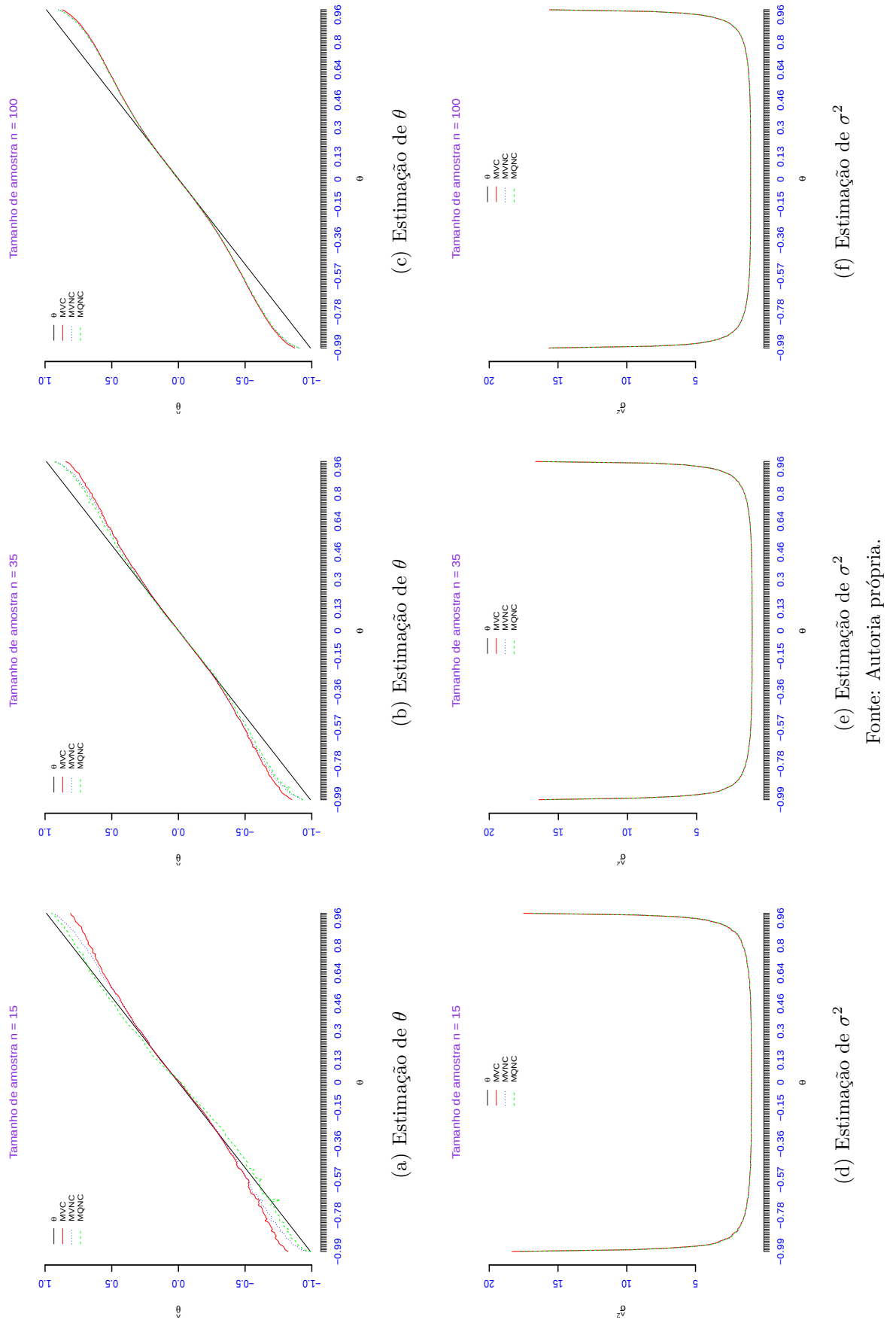
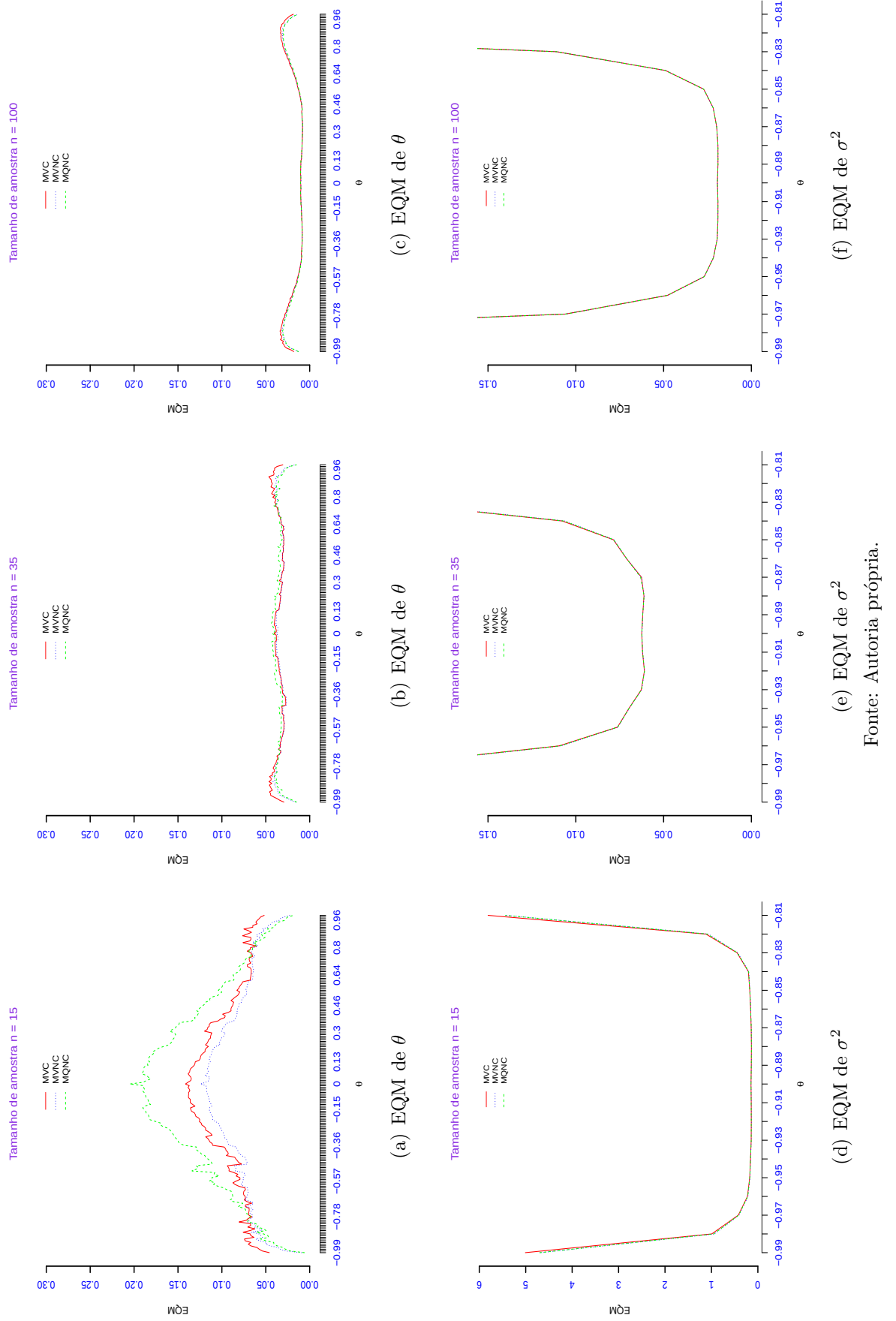
Figura 4.8: Média das estimativas dos parâmetros θ e σ^2 sob os pontos discretizados da região de estabilidade para o processo MA(1).

Figura 4.9: Erro Quadrático Médio das estimativas de θ e σ^2 sob os pontos discretizados da região de estabilidade para o processo MA(1).

Em relação ao EQM da estimativa de σ^2 nota-se a partir da Figura 4.9 não existe dissimilaridade significativa entre os métodos, apenas com tamanho $n = 15$, a estimação por *MVC* tem maior erro quadrático médio.

Avaliação via Δ_{KL} , Δ_{TKL} , Δ_{JS} , Δ_B e Δ_H :

A média das medidas de divergências consideradas em formas reduzidas e completas e são apresentadas nas Figuras 4.10, 4.11, 4.12, 4.13 e 4.14.

Forma Reduzida: Sob a forma reduzida das medidas de divergência, em geral, o comportamento entre as divergências foram similares. Sejam os mapeamentos $\delta^-(n) : \mathbb{N} \mapsto (-1, 0)$ e $\delta^+(n) : \mathbb{N} \mapsto (0, 1)$,

$$\begin{cases} \Delta_{\mathcal{M}}(\hat{\phi}_{MVNC}) \leq \Delta_{\mathcal{M}}(\hat{\phi}_{MVC}) \leq \Delta_{\mathcal{M}}(\hat{\phi}_{MQNC}), & \text{para } \phi < \delta^-(n) \text{ ou } \phi > \delta^+(n), \\ \Delta_{\mathcal{M}}(\hat{\phi}_{MQNC}) \leq \Delta_{\mathcal{M}}(\hat{\phi}_{MVNC}) \leq \Delta_{\mathcal{M}}(\hat{\phi}_{MVC}), & \text{para } \phi > \delta^-(n) \text{ e } \phi < \delta^+(n), \end{cases}$$

em que $\mathcal{M} = \{KL, TKL, JS, B, H\}$ e $(\delta^-(15), \delta^-(35), \delta^-(100)) = -0.7$ e $(\delta^+(15), \delta^+(35), \delta^+(100)) = 0.7$. Outro fato importante identificado nesse estudo de simulação é que as curvas das divergências dos métodos de estimação se aproximam com o aumento do tamanho das amostras; por exemplo, como pode ser visto na comparação entre os tamanhos amostrais $n = 15$ e 100 na Figura 4.10 para a medida de Δ_{KL} na forma reduzida. Assim, os métodos são assintoticamente equivalentes em termos da distorção da verossimilhança conjunta.

Forma Completa: Utilizando a forma completa das medidas de divergência, no geral, não existem disparidades significantes entre os métodos de estimação. No entanto, quando $|\theta|$ é próximo de 1, a estimação por *MVC* apresenta uma divergência menor em relação aos outros métodos, configurando-se como o melhor método de estimação nessa região. Esse comportamento está em discordância com a forma reduzida que aponta as estimações não condicionais de máxima verossimilhança e mínimos quadrados como melhores.

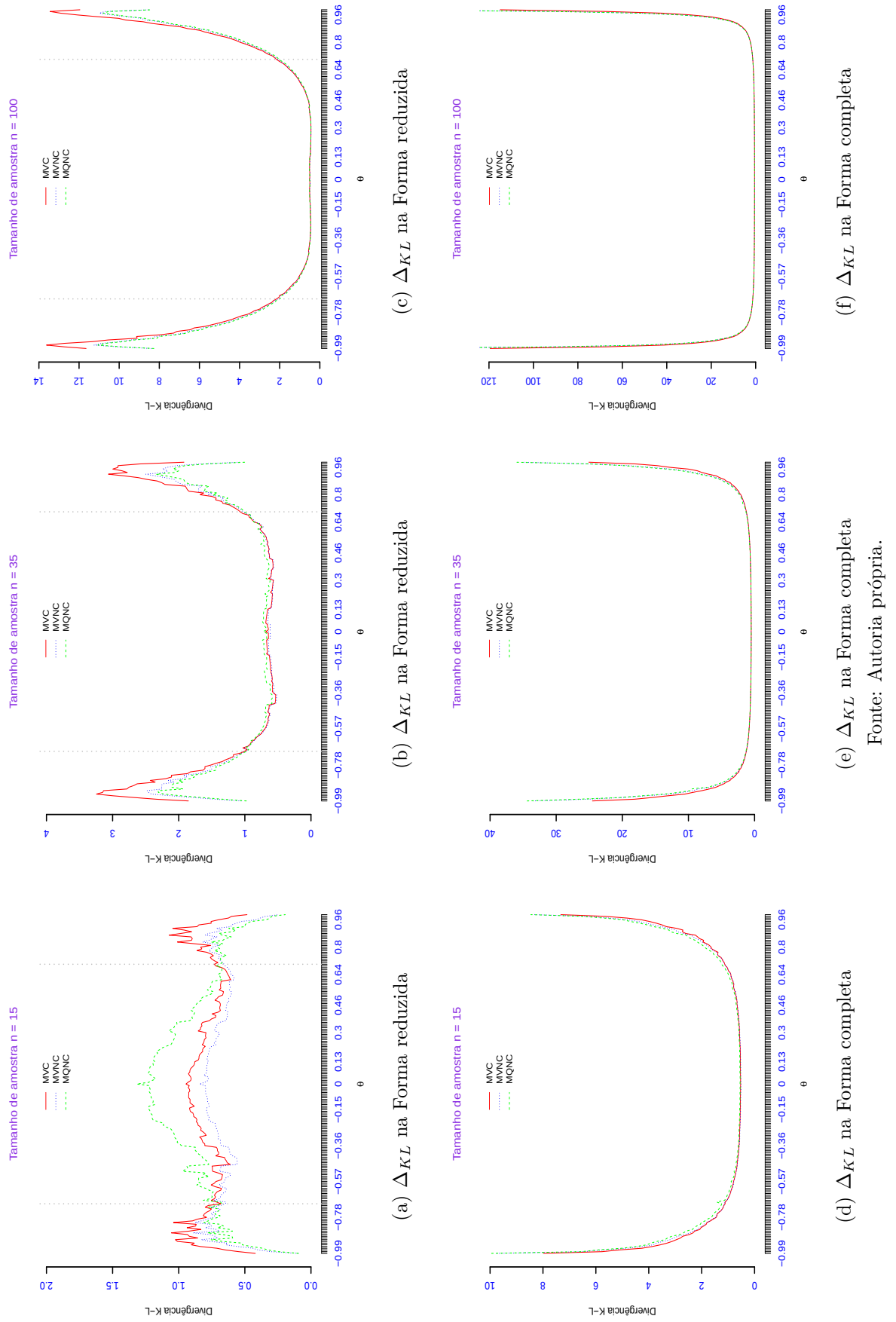
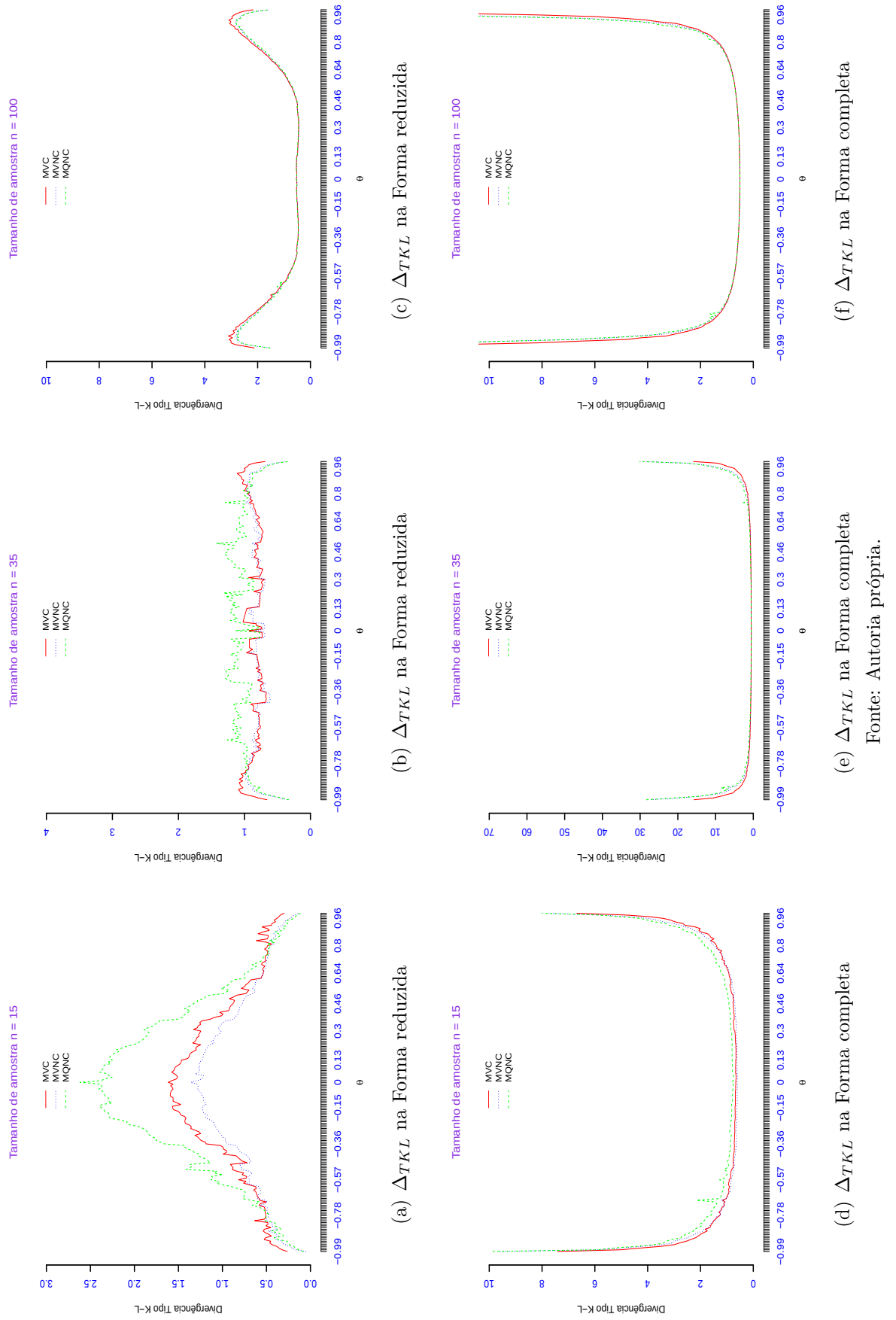
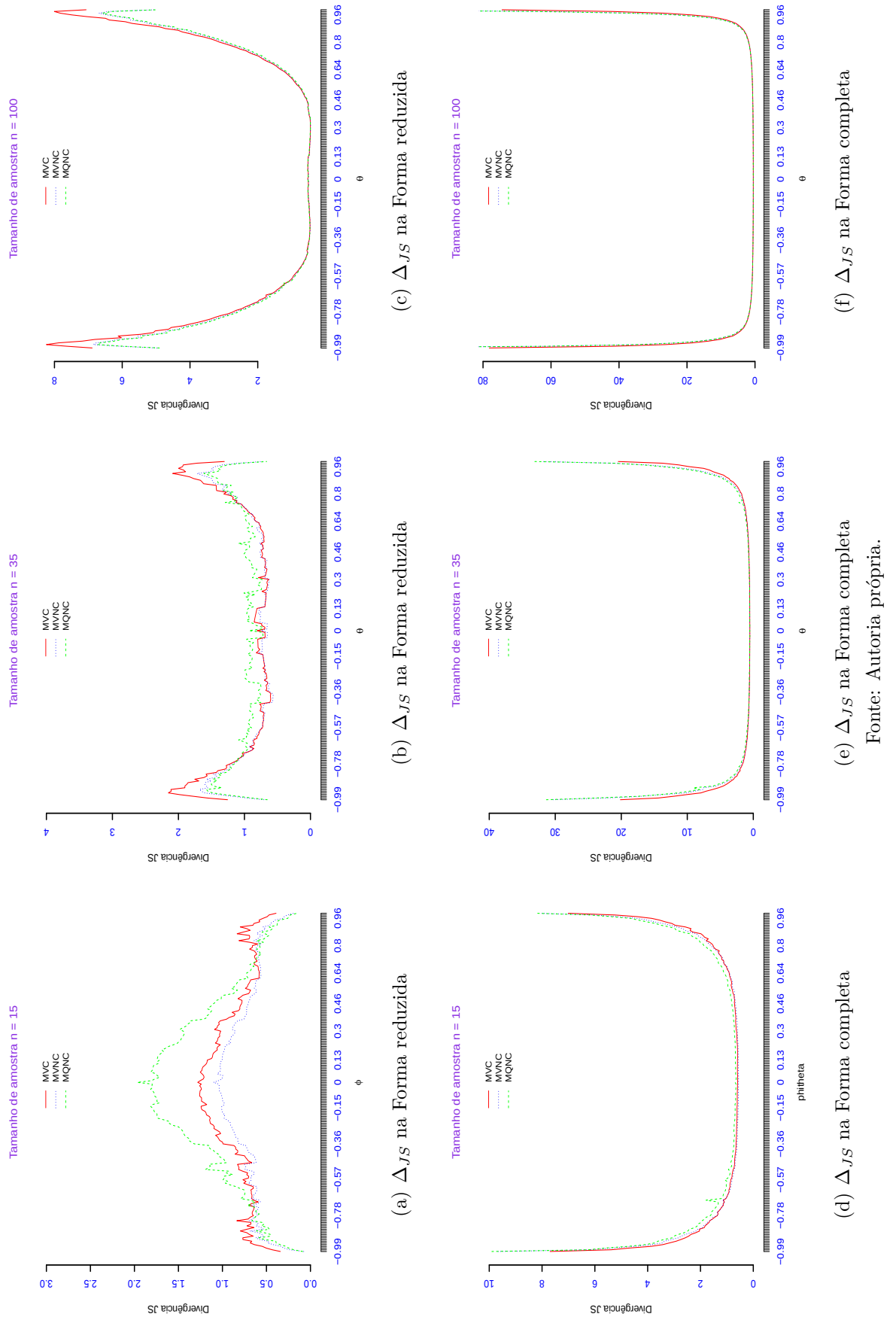
Figura 4.10: Média da divergência *Kullback-Leibler* sob os pontos discretizados da região de estabilidade para o processo MA(1)

Figura 4.11: Média da divergência Tipo *Kullback-Leibler* sob os pontos discretizados da região de estabilidade para o processo MA(1).

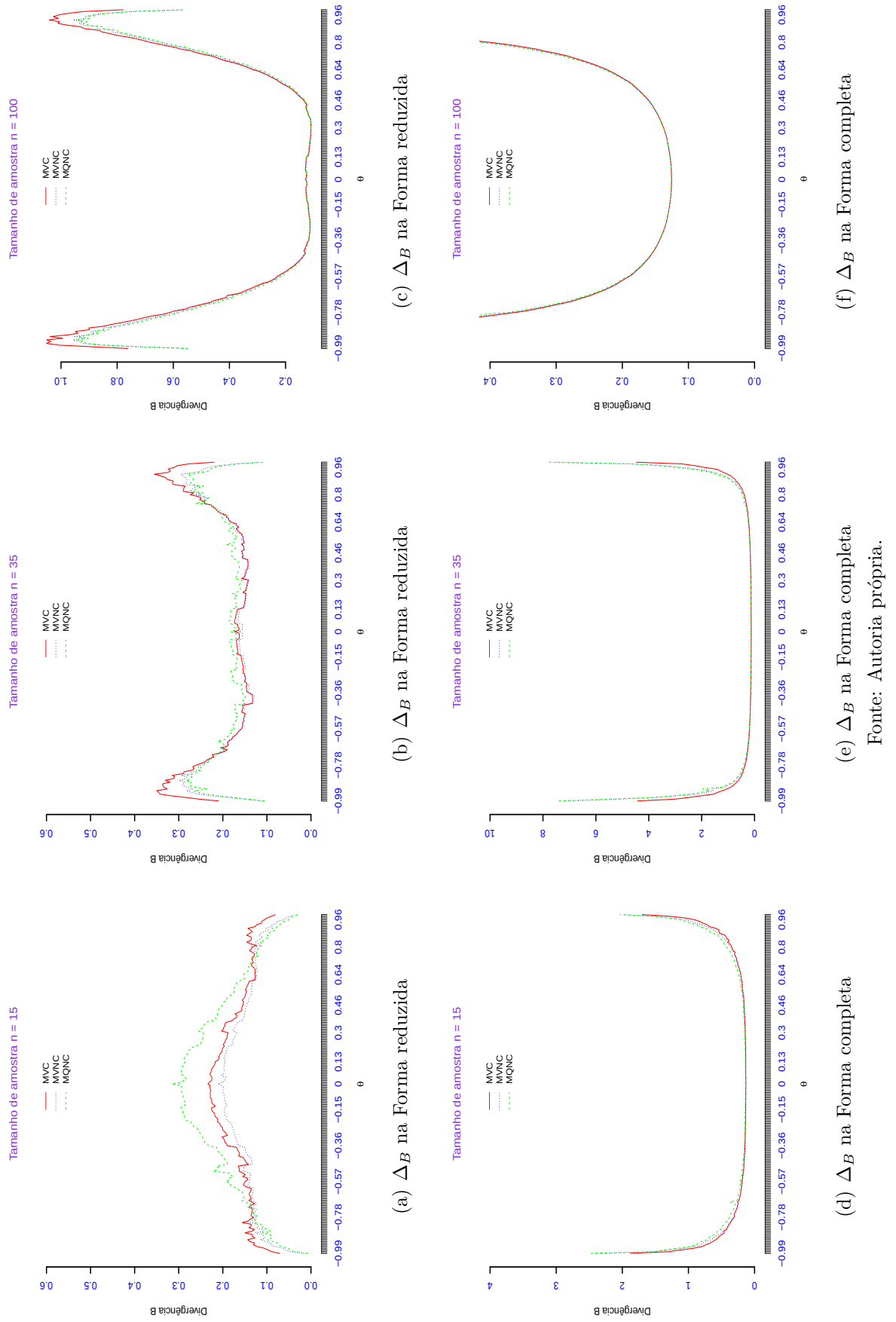


Fonte: Autoria própria.

Figura 4.12: Média da divergência *Jensen-Shannon* sob os pontos discretizados da região de estabilidade para o processo MA(1).

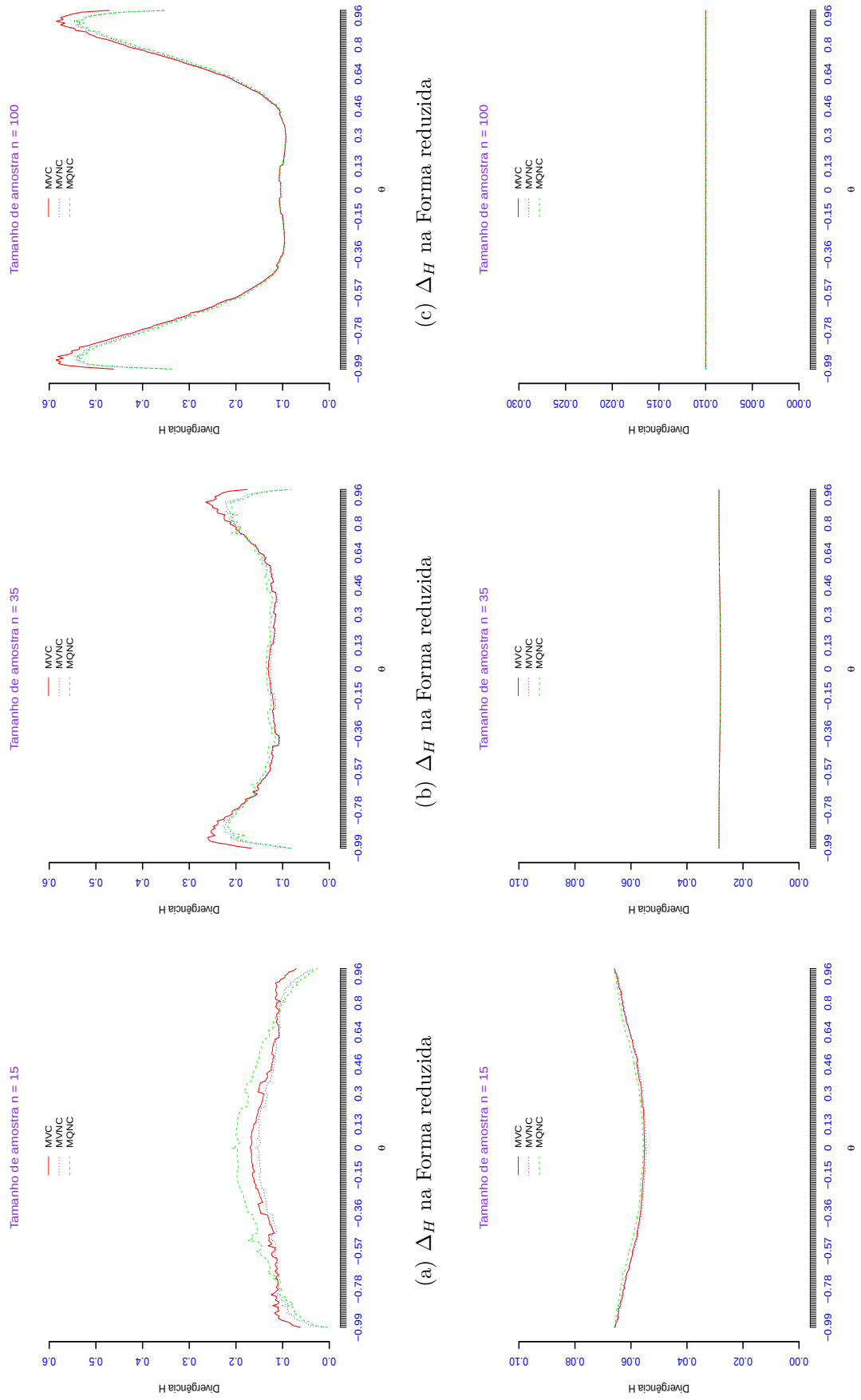


Fonte: Autoria própria.

Figura 4.13: Média da divergência *Battacharrya* sob os pontos discretizados da região de estabilidade para o processo MA(1).

Fonte: Autoria própria.

Figura 4.14: Média da divergência *Hellinger* sob os pontos discretizados da região de estabilidade para o processo MA(1).



4.1.3 Estimação no modelo ARMA(1,1)

Sob a região de estabilidade do modelo $ARMA(1,1)$, avaliamos as estimativas do vetor de parâmetros (ϕ, θ, σ^2) pelos métodos MVNC, MVC e MQNC. Inicialmente, esta avaliação é realizada via viés e EQM. Subsequentemente, apresenta-se a avaliação dos métodos da perspectiva de divergências. O método que apresenta o menor valor da distância é identificado como o melhor método de estimação segundo a distorção entre verossimilhanças conjuntas.

Nas Figuras 4.15 e 4.16, apresenta-se o método de estimação menos viesado segundo sub regiões de um processo ARMA(1,1) estável, das estimações de ϕ e θ , respectivamente. Vale salientar que exibi-se apenas o melhor método conforme a legenda das figuras, considerando a região de estabilidade do processo.

Na Figura 4.15, considerando $n = 25$ e $n = 35$, identifica-se que quando $|\phi|$ assume valores próximos a 1, MQNC obteve o desempenho melhor em comparação com os outros métodos. Quando $|\theta|$ é próximo de 1, MVNC e MQNC tem melhores aproximações em termo da media das estimativas de ϕ . A estimação condicional teve melhores resultados quando ϕ e θ são próximos da origem. Para tamanho de amostra grande $n = 100$, o comportamento das estimações é semelhante àqueles referentes as amostras pequenas; entretanto, nota-se um domínio do método de estimação MQNC sob a região de estabilidade.

A média das estimações do parâmetro θ é apresentada na Figura 4.16 para diferentes tamanhos amostrais. No contexto das amostras pequenas ($n = 25$ e $n = 35$), MVC teve melhor desempenho na maioria dos pontos considerados principalmente na região de estabilidade onde $|\theta| \approx 0.9$ e $-0.9 < \phi < 0.9$ (vê Figura 4.16), seguindo do desempenho de MQNC. Para o tamanho amostral $n = 100$, temos comportamento similares. Em relação ao parâmetro σ^2 dado na Figura 4.17, a estimação por MQNC é melhor em termos da média entre as estimações com tamanhos de amostras $n = 25$ e $n = 35$. Quando aumentamos o tamanho amostral para $n = 100$, a estimação condicional é consideravelmente melhor; entretanto, quando $|\phi| \approx 1$ as estimações por MQNC e MVNC são melhores na maioria dos pontos nessa região.

A avaliação do EQM das estimativas do parâmetro ϕ para diferentes tamanhos de amostra é apresentada na Figura 4.18. Verifica-se que no intervalo $-0.7 < \phi < 0.7$ e $-0.7 < \theta < 0.7$, o método de estimação MVC foi mais eficiente em grande parte dos pontos selecionados. Na região em que $\phi = 0.9$, $\phi = -0.9$, $\theta = -0.9$ e $\theta = 0.9$, MVNC e MQNC apresentam os menores erros quadráticos médios. Em relação ao parâmetro θ , a Figura 4.19 exibi o gráfico de melhor desempenho dos métodos de estimação para o parâmetro θ . Para tamanho de amostra $n = 25$, a estimação po MVNC apresentou menores EQM de θ . A melhor foi evidenciada na região mais próxima da fronteira de

estabilidade, como visto na Figura 4.19. A partir do tamanho $n = 50$, as estimações MVC e MVNC obtiveram melhores desempenhos em relação ao MQNC. Por fim, a Figura 4.20 apresenta a comparação da performance dos métodos segundo o EQM do parâmetro σ^2 . Na região próxima a fronteira de estabilidade, o método MVC tem melhor desempenho na grande maioria dos pontos considerados nessa região. Em contra partida, quando ϕ e θ estão próximos da origem, a estimação por MQNC se configura como o método que minimiza o erro quadrático médio. Estes resultados são apresentados para casos particulares nas Tabelas 4.1 e 4.2, para os tamanhos $n = 15$ e 100, respectivamente.

Figura 4.15: Média da estimativas do método menos viesado para o parâmetro ϕ sob a região de estabilidade de um $ARMA(1, 1)$.

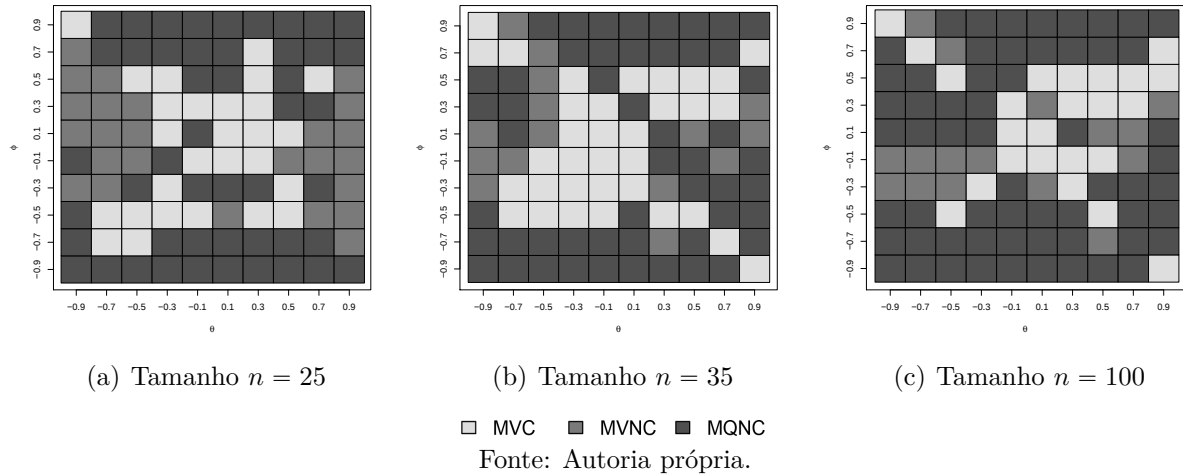


Tabela 4.1: Estimativas dos parâmetros para $ARMA(1, 1)$ com $n = 15$

(ϕ, θ, σ^2)	$\hat{\phi}$	$\hat{\theta}$	$\hat{\sigma}^2$	$EQM(\hat{X})$	$EQM(\hat{\phi})$	$EQM(\hat{\theta})$	$EQM(\hat{\sigma}^2)$	$CV(\hat{\phi})$	$CV(\hat{\theta})$	$CV(\hat{\sigma}^2)$
Condicional										
(-0.9,-0.9,1)	-0.848	-0.796	1.013	2.319	0.024	0.068	0.195	-0.171	-0.300	0.436
(0.1,-0.9,1)	-0.154	-0.667	0.984	1.100	0.166	0.257	0.171	-2.078	-0.675	0.420
(0.1,0.9,1)	0.226	0.722	1.018	1.154	0.093	0.161	0.176	1.231	0.498	0.413
(0.9,0.9,1)	0.834	0.794	1.006	2.307	0.032	0.074	0.198	0.199	0.315	0.443
(-0.9,0.1,1)	-0.785	0.085	0.824	1.666	0.072	0.248	0.147	-0.310	5.832	0.413
(0.1,0.1,1)	-0.102	0.333	0.842	0.984	0.319	0.511	0.151	-5.154	2.030	0.421
(0.9,0.1,1)	0.796	0.148	0.859	1.526	0.054	0.218	0.143	0.260	3.154	0.409
Não Condicional										
(-0.9,-0.9,1)	-0.837	-0.902	0.891	2.270	0.021	0.031	0.124	-0.157	-0.195	0.377
(0.1,-0.9,1)	-0.047	-0.807	0.898	1.049	0.127	0.143	0.134	-6.909	-0.455	0.392
(0.1,0.9,1)	0.129	0.861	0.911	1.094	0.081	0.059	0.130	2.200	0.279	0.384
(0.9,0.9,1)	0.829	0.881	0.910	2.281	0.028	0.044	0.137	0.184	0.238	0.395
(-0.9,0.1,1)	-0.788	0.012	0.851	1.109	0.068	0.198	0.134	-0.299	37.463	0.393
(0.1,0.1,1)	-0.229	0.491	0.866	0.911	0.402	0.483	0.147	-2.371	1.172	0.415
(0.9,0.1,1)	0.791	0.216	0.876	1.265	0.061	0.182	0.135	0.281	1.908	0.395
Mínimos Quadrados										
(-0.9,-0.9,1)	-0.897	-0.959	2.303	2.303	0.016	0.022	4.435	-0.143	-0.143	0.719
(0.1,-0.9,1)	-0.079	-0.868	1.057	1.057	0.133	0.123	0.181	-4.026	-0.403	0.399
(0.1,0.9,1)	0.162	0.934	1.103	1.103	0.080	0.042	0.205	1.709	0.216	0.401
(0.9,0.9,1)	0.890	0.938	2.314	2.314	0.021	0.039	5.126	0.164	0.206	0.798
(-0.9,0.1,1)	-0.842	0.043	1.132	1.132	0.063	0.255	0.285	-0.290	11.735	0.458
(0.1,0.1,1)	-0.224	0.528	0.935	0.935	0.397	0.576	0.156	-2.409	1.188	0.417
(0.9,0.1,1)	0.846	0.219	1.285	1.285	0.053	0.238	0.554	0.264	2.164	0.536

Figura 4.16: Média das estimativas do parâmetro θ sob a região de estabilidade para o processo $ARMA(1, 1)$.

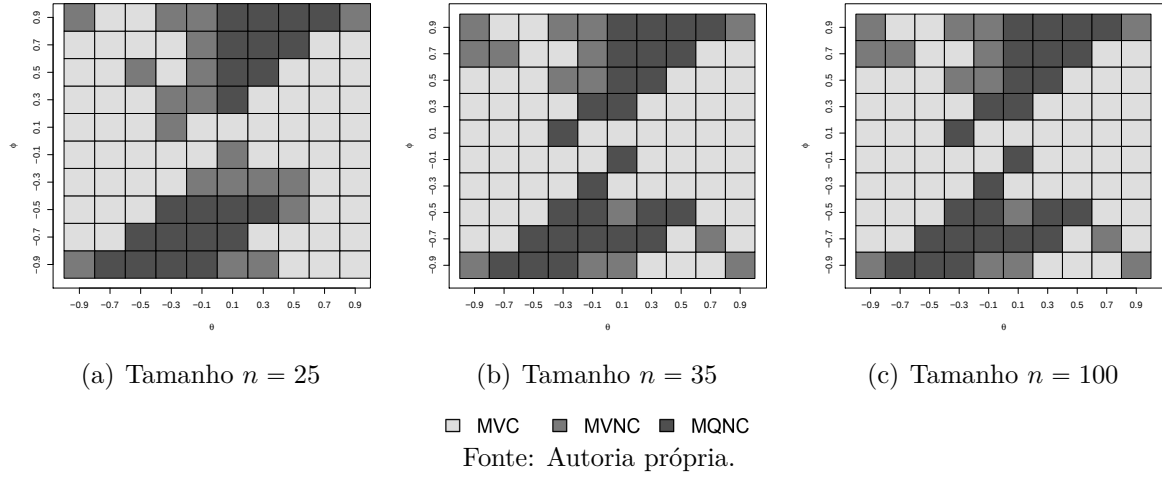
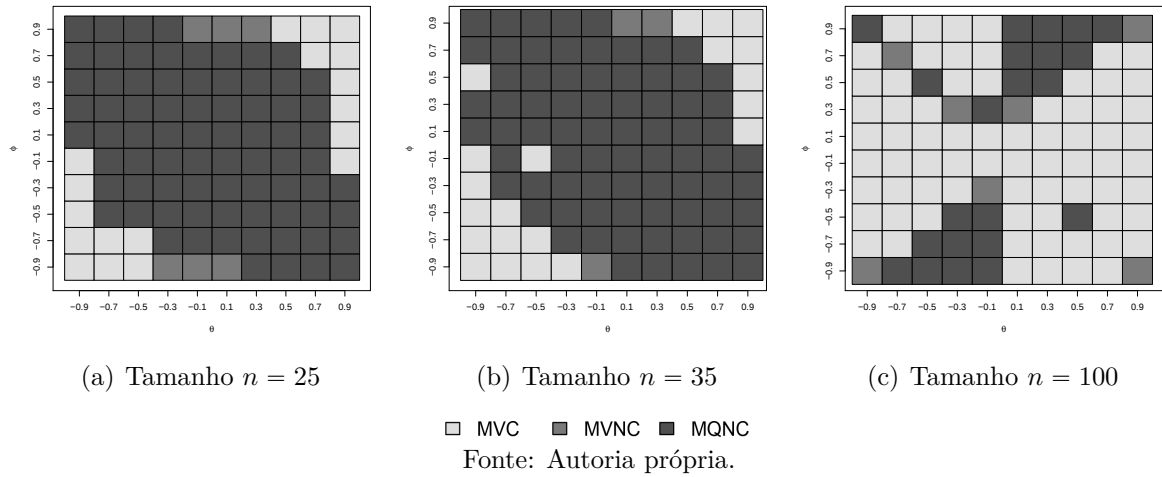


Figura 4.17: Média das estimativas do parâmetro σ^2 sob a região de estabilidade de um processo $ARMA(1, 1)$.



Avaliação via Δ_{KL} , Δ_{TKL} , Δ_{JS} , Δ_B e Δ_H :

Nas Figuras 4.21, 4.23, 4.25 e 4.27. apresentam-se os desempenhos dos métodos em relação as medidas de distâncias *Bhattacharyya*, Tipo *Kullback-Leibler*, *Kullback-Leibler* e *Jensen-Shannon* nas formas reduzidas e adicionalmente as Figuras 4.22, 4.24, 4.26 e 4.28 são apresentadas para a forma completa. Neste sentido, preenchemos a região de estabilidade com o método que apresentou a menor distância entre os métodos de estimação considerados.

Figura 4.18: Avaliação do erro quadrático médio das estimativas do parâmetro ϕ sob a região de estabilidade de um processo $ARMA(1, 1)$.

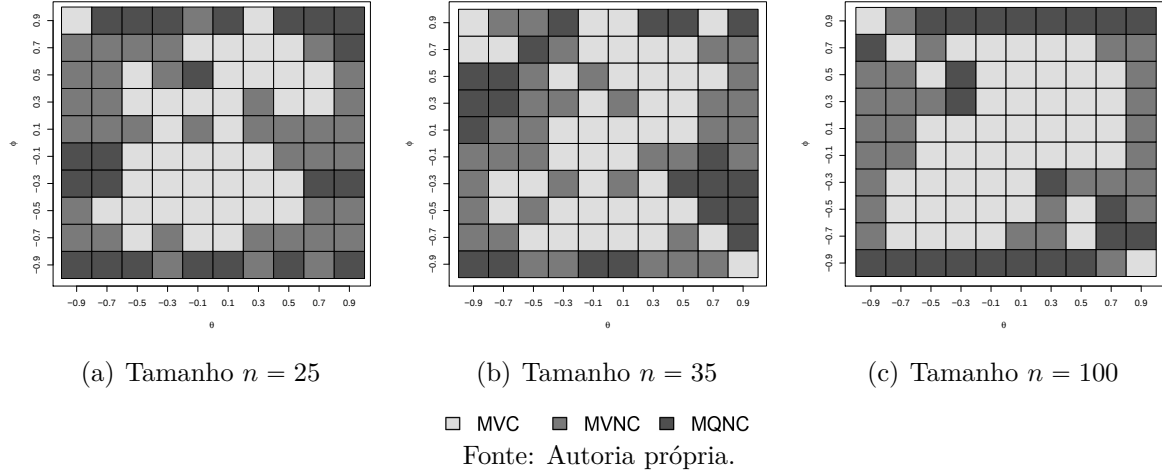
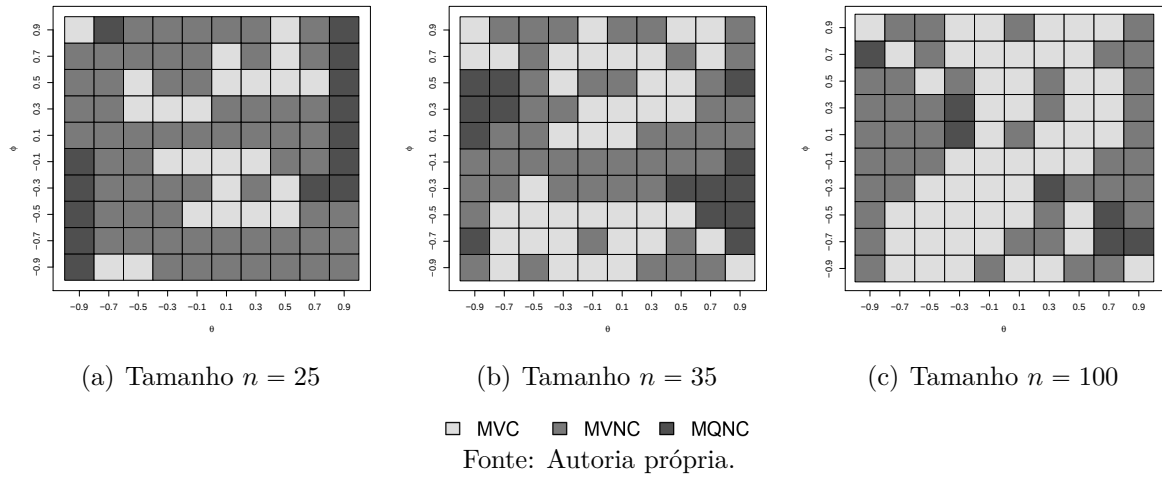
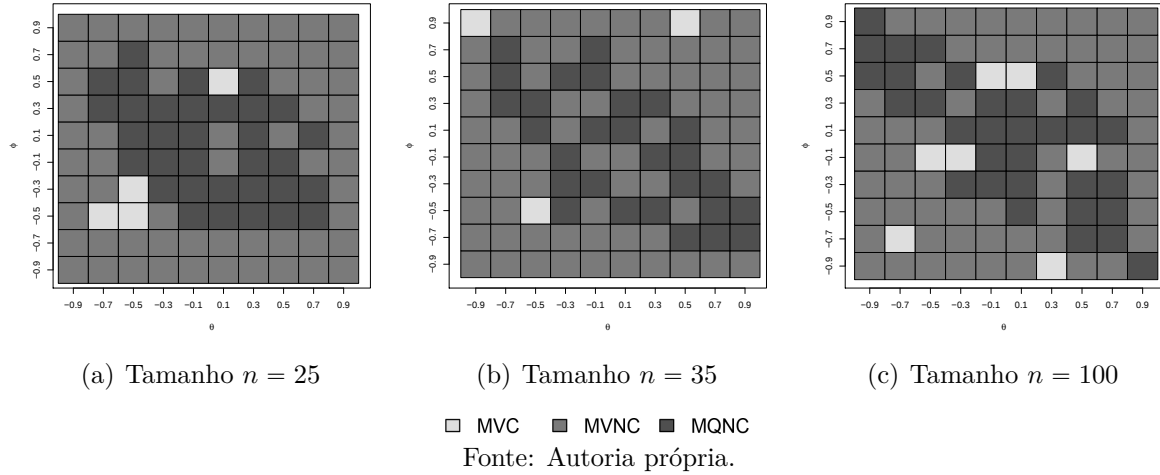


Figura 4.19: Avaliação do erro quadrático médio das estimativas do parâmetro θ sob a região de estabilidade de um processo $ARMA(1, 1)$.



Considerando a forma reduzida, o comportamento entre as medidas Δ_{KL} , Δ_{JS} , Δ_B e Δ_H são similares. A estimação MVNC teve superioridade na maioria dos pontos, para tamanhos de amostras $n = 25$ e 35 . Mas se é considerado $n = 100$, em direção a região $|\phi| < 0.9 \cap |\theta| < 0.9$ o estimador de MVC revelou-se como de melhor performance.

As medidas Δ_{KL} , Δ_{TKL} , Δ_{JS} , Δ_B e Δ_H em geral, apresentaram o mesmo comportamento para a forma completa. Com tamanho amostral $n = 25$, verifica-se que o método de estimação MVNC foi melhor na maior parte da região de estabilidade. para $n = 35$, segundo a região $\theta = -0.9 \cap (-0.7 \leq \phi \leq 0.3)$ e com tamanho de amostra grande representado por

Figura 4.20: Avaliação do erro quadrático médio das estimativas do parâmetro σ^2 sob a região de estabilidade de um processo $ARMA(1,1)$.**Tabela 4.2:** Estimativas dos parâmetros para $ARMA(1,1)$ com $n = 100$

(ϕ, θ, σ^2)	$\hat{\phi}$	$\hat{\theta}$	$\hat{\sigma}^2$	$EQM(\hat{X})$	$EQM(\hat{\phi})$	$EQM(\hat{\theta})$	$EQM(\hat{\sigma}^2)$	$CV(\hat{\phi})$	$CV(\hat{\theta})$	$CV(\hat{\sigma}^2)$
Condicional										
$(-0.9, -0.9, 1)$	-0.887	-0.882	1.011	1.194	0.003	0.004	0.022	-0.057	-0.069	0.148
$(0.1, -0.9, 1)$	0.053	-0.869	1.009	1.009	0.017	0.007	0.022	2.291	-0.089	0.148
$(0.1, 0.9, 1)$	0.132	0.877	1.004	1.009	0.015	0.005	0.021	0.892	0.078	0.144
$(0.9, 0.9, 1)$	0.887	0.884	1.007	1.180	0.002	0.004	0.020	0.053	0.074	0.139
$(-0.9, 0.1, 1)$	-0.878	0.084	0.986	1.016	0.004	0.015	0.021	-0.071	1.420	0.147
$(0.1, 0.1, 1)$	0.089	0.116	0.975	0.979	0.159	0.191	0.022	4.468	3.772	0.151
$(0.9, 0.1, 1)$	0.880	0.115	0.985	1.036	0.004	0.013	0.019	0.066	0.980	0.141
Não Condicional										
$(-0.9, -0.9, 1)$	-0.884	-0.910	0.978	1.186	0.002	0.003	0.020	-0.053	-0.059	0.142
$(0.1, -0.9, 1)$	0.093	-0.910	0.974	0.998	0.014	0.004	0.020	1.275	-0.070	0.143
$(0.1, 0.9, 1)$	0.106	0.910	0.972	1.000	0.012	0.003	0.020	1.041	0.063	0.141
$(0.9, 0.9, 1)$	0.883	0.915	0.973	1.172	0.002	0.003	0.018	0.052	0.063	0.135
$(-0.9, 0.1, 1)$	-0.878	0.085	0.986	1.016	0.004	0.015	0.021	-0.070	1.426	0.146
$(0.1, 0.1, 1)$	0.034	0.176	0.975	0.976	0.166	0.181	0.022	11.845	2.390	0.149
$(0.9, 0.1, 1)$	0.880	0.115	0.985	1.037	0.004	0.013	0.019	0.064	0.973	0.139
Mínimos Quadrados										
$(-0.9, -0.9, 1)$	-0.892	-0.929	1.189	1.189	0.002	0.004	0.125	-0.053	-0.063	0.252
$(0.1, -0.9, 1)$	0.099	-0.928	1.000	1.000	0.014	0.005	0.020	1.209	-0.073	0.143
$(0.1, 0.9, 1)$	0.103	0.927	1.001	1.001	0.013	0.004	0.020	1.099	0.065	0.142
$(0.9, 0.9, 1)$	0.891	0.935	1.175	1.175	0.002	0.005	0.111	0.053	0.065	0.242
$(-0.9, 0.1, 1)$	-0.887	0.086	1.016	1.016	0.004	0.015	0.024	-0.070	1.419	0.151
$(0.1, 0.1, 1)$	-0.016	0.231	0.977	0.977	0.194	0.191	0.021	-26.893	1.809	0.148
$(0.9, 0.1, 1)$	0.889	0.116	1.037	1.037	0.003	0.013	0.029	0.064	0.981	0.161

$n = 100$, quando $\theta \uparrow 0.9 \cap \theta \downarrow -0.9$ o método MVC tem melhor desempenho.

Com base na distância Tipo *Kullback-Leibler*, com tamanhos de amostras pequenos $n = 25$ e 35 , na grande maioria dos pontos o método MVNC apresenta as menores distâncias; no entanto, quando $|\theta| = 0.9$, MQNC obteve um melhor desempenho. Para 100 , tem-se destaque para o método MVC.

Adicionalmente, realizamos estudos de casos para alguns pontos na região de estabilidade nas Tabelas 4.1 e 4.2.

Desta maneira, conclui-se que

As medidas de divergências *Kullback-Leibler*, *Jensen-Shannon* e *Bhattacharyya* com forma reduzida, apontam o método de MVNC como o melhor método nos pontos considerados, dado que apresentam os menores valores

dessas medidas. MQNC foi melhor nos pontos $(\phi = 0.1; \theta = -0.9; \sigma^2 = 1)$ e $(\phi = 0.1; \theta = 0.9; \sigma^2 = 1)$ considerando a Tipo *Kullback-Leibler*.

Figura 4.21: Identificação do melhor método de estimação em termos da distância *Bhattacharyya* na forma reduzida.

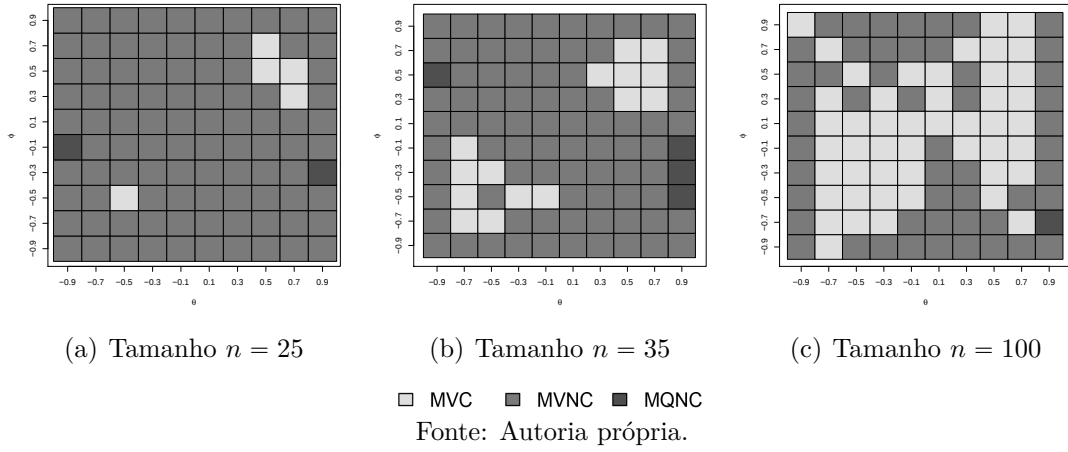


Figura 4.22: Identificação do melhor método de estimação em termos da distância *Bhattacharyya* na forma completa.

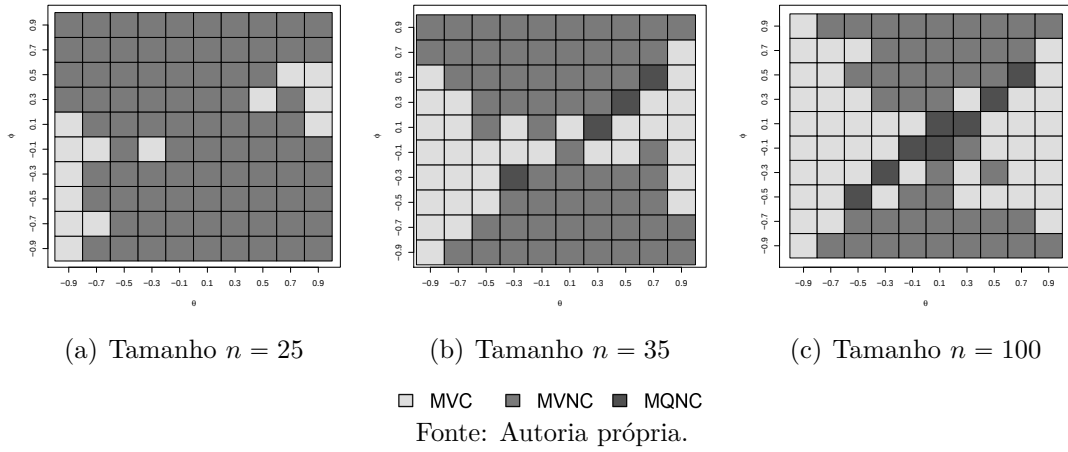


Figura 4.23: Identificação do melhor método de estimação em termos da distância Tipo *Kullback – Leibler* na forma *reduzida*.

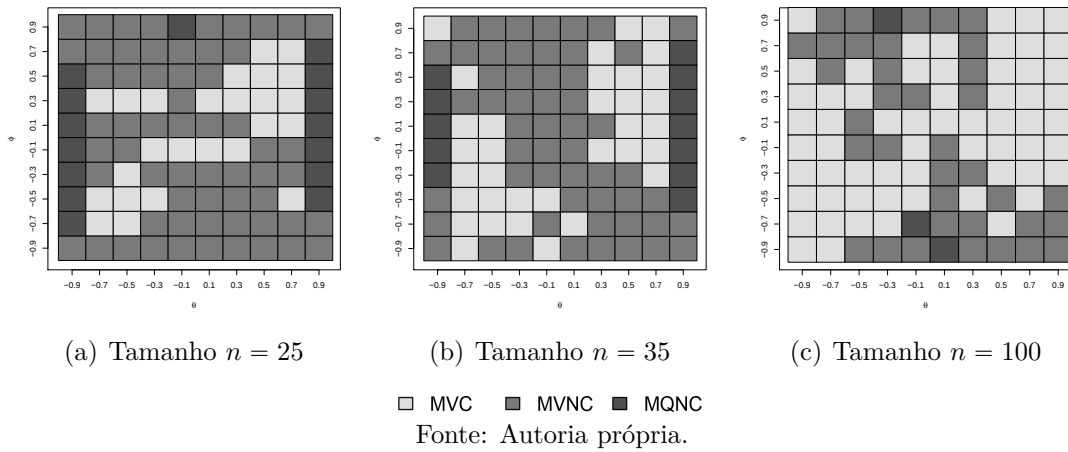


Figura 4.24: Identificação do melhor método de estimação em termos da distância Tipo *Kullback – Leibler* na forma *completa*.

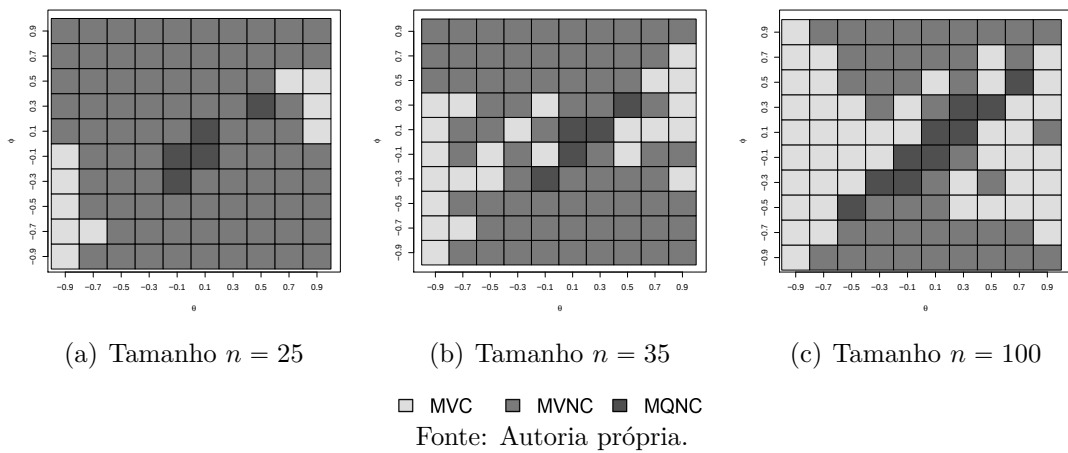


Figura 4.25: Identificação do melhor método de estimação em termos da distância *Kullback – Leibler* na forma *reduzida*.

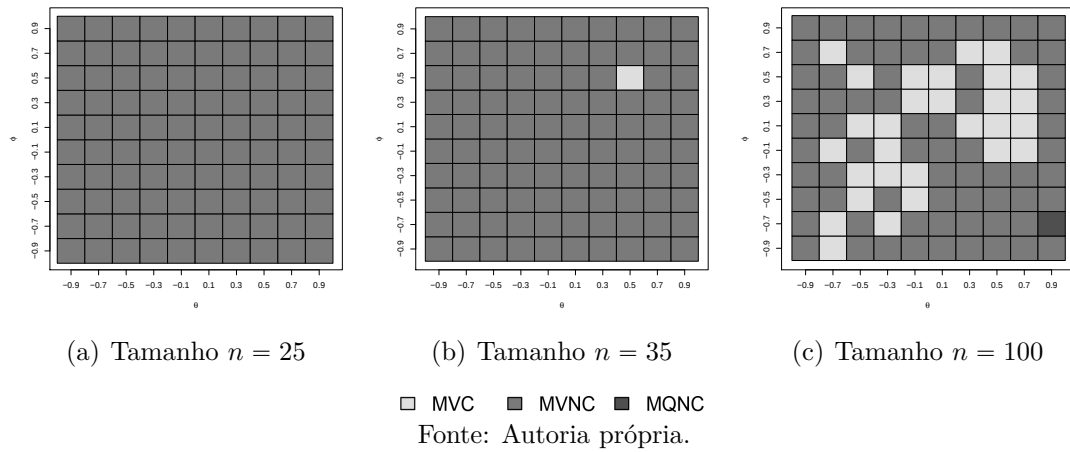


Figura 4.26: Identificação do melhor método de estimação em termos da distância *Kullback – Leibler* na forma *completa*.

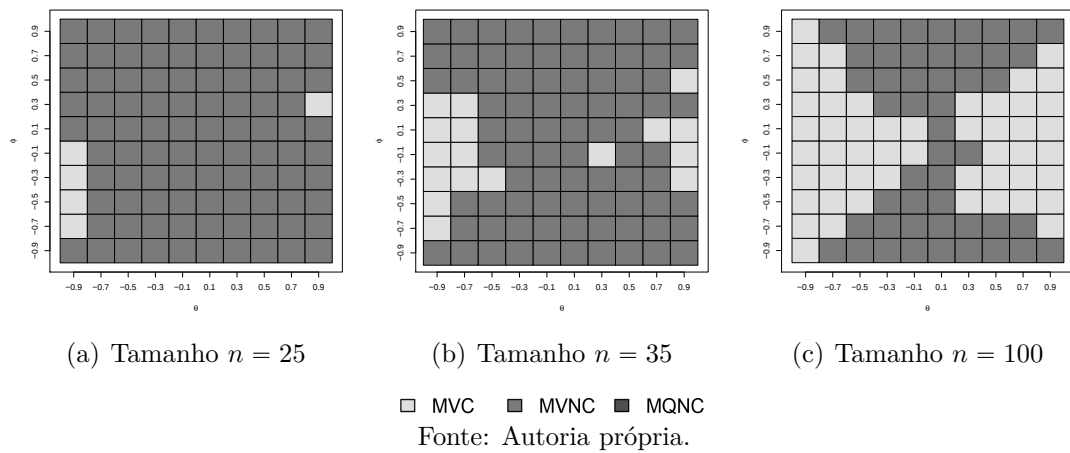


Figura 4.27: Identificação do melhor método de estimação em termos da distância *Jensen-Shannon* na forma *reduzida*.

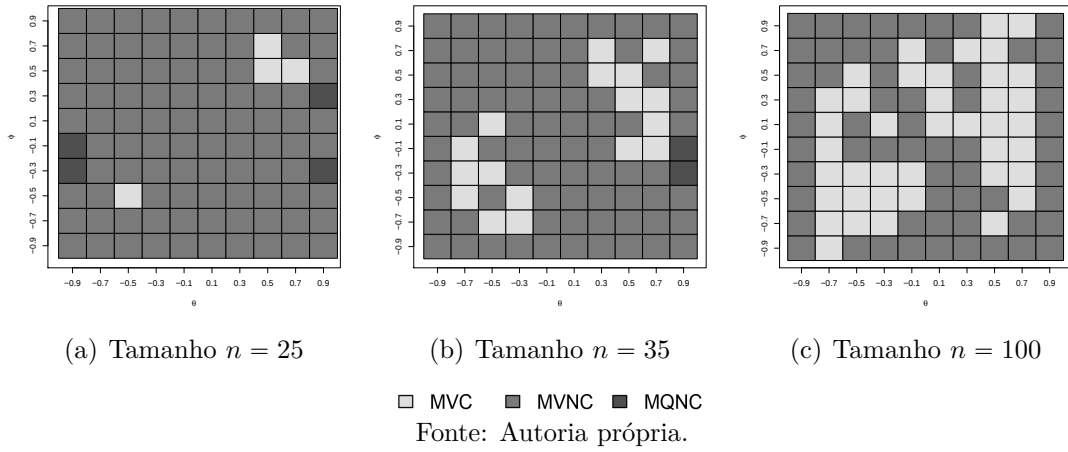
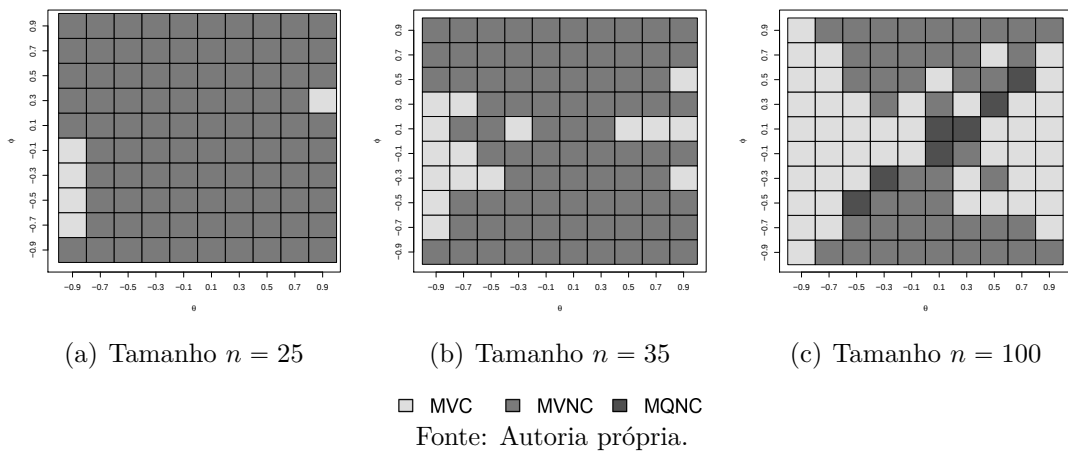


Figura 4.28: Identificação do melhor método de estimação em termos da distância *Jensen-Shannon* na forma *completa*.



4.2 Testes para identificação: Divergência vs. Razão entre verossimilhança

Nesta seção, apresenta-se um estudo comparativo do desempenho entre sete testes de hipóteses:

- Sob divergências: S_{KL} , S_{TKL} , S_{JS} , S_B e S_H .
- Sob razão entre verossimilhanças: S_{RVNC} e S_{RVC} .

Foi realizado um estudo de simulação Monte Carlo com 1000 réplicas e tamanhos variados de amostras. Adicionalmente, foram considerados três métodos de estimação clássicos: MVC, MVNC e MQNC. Avaliamos cinco propostas de testes de hipótese bilaterais (caso em que não existe teste *uniformemente mais poderosos*) para identificação em modelos ARMA e comparamos ela com o teste da razão entre verossimilhanças nas estruturas condicionais e não condicionais. Nas análises, assumiu-se que a hipótese $\sigma^2 = 1$, em que σ^2 foi entendida como um parâmetro de perturbação seguindo a abordagem de Cordeiro (1999). Finalmente os métodos, foram comparados utilizando como critérios a quantificação da função poder definida com se segue: Sejam Υ um teste de hipótese para $\mathcal{H}_0$ e Θ o espaço paramétrico associado ao teste, então sua função poder é definida como

$$\begin{aligned} \pi_{\Upsilon}(\theta) &= P(\{\text{Rejeitar } \mathcal{H}_0\} | \theta \in \Theta) \\ &= \begin{cases} \alpha \triangleq P(\{\text{Rejeitar } \mathcal{H}_0\} | \theta \in \Theta_0), & \text{para } \theta = \theta_0, \\ 1 - \beta \triangleq P(\{\text{Rejeitar } \mathcal{H}_0\} | \theta \in \Theta - \Theta_0), & \text{para } \theta = \theta_1 \neq \theta_0, \end{cases} \end{aligned}$$

em que α é o tamanho do erro do tipo I, β é o tamanho do erro do tipo II e $1 - \beta$ é o poder do teste. A fim de tornar o estudo do poder factível, assumiremos as seguintes parametrizações para o poder do teste:

- Para o caso uniparamétrico-AR(1) e MA(1): $\mathcal{H}_0 : \theta = \theta_0 \times \mathcal{H}_1 : \theta = \theta_0 + \Delta$ para $\Delta \in \{x \in \mathbb{R} - \{0\} : |\theta| < 1\}$.
- Para o caso biparamétrico-ARMA(1,1): $\mathcal{H}_0 : (\theta_1, \theta_2) = (\theta_{01}, \theta_{02}) \times \mathcal{H}_1 : (\theta_1, \theta_2) = (\theta_{01}, \theta_{02}) + (\Delta_1, \Delta_2)$ para $(\Delta_1, \Delta_2) \in \{(x_1, x_2) \in \mathbb{R}^2 - \{(0, 0)\} : |\theta_1| < 1, |\theta_2| < 1\}$.

4.2.1 Testes de identificação para classe de modelo AR(1)

Como um comportamento geral, o poder empírico dos testes cresce a medida que Δ se afasta de zero e/ou o tamanho de amostra n aumenta. As Tabelas 4.3, 4.4 e 4.5 apresentam a versão empírica da função poder dos testes considerando para $\phi = 0.1$, $\phi = 0.5$ e $\phi = 0.8$, respectivamente, para um modelo $AR(1)$ com tamanhos amostras $n = 35, 50$ e 100 . Os

poderes empíricos dos testes para EMVNC com os tamanhos amostrais $n = 35$ e 100 são apresentados nas Figuras 4.29 e 4.30.

Para a Tabela 4.3, os poderes empíricos dos testes da razão entre verossimilhanças para as versões condicional e não condicional foram equivalentes para todos os tamanhos de amostras considerados. Considerando o método de estimação MQNC, os poderes empíricos dos testes baseados nas distâncias *Kullback-Leibler*, *Jensen-Shannon* e *Bhattacharyya* são maiores quando comparados com o teste da razão entre verossimilhanças, obedecendo a seguinte relação: Seja $\pi_{D:n}(\theta)$ para $D \in \{KL, TKL, JS, B, H, RVC, RVNC\}$ a versão empírica a função poder obtida com base em um vetor ARMA de tamanho n ,

$$\pi_{KL:n=35,50}(\phi + \Delta) > \pi_{JS:n=35,50}(\phi + \Delta) > \pi_{B:n=35,50}(\phi + \Delta).$$

O teste induzido da distância de *Hellinger* apresentou pior desempenho. Contudo, os demais testes em geral apresentam tamanhos empíricos próximos do nível nominal de 5%.

As seguintes desigualdades de funções poder empíricas podem ser verificadas:

Quando $\phi = 0.1$ e $\Delta \downarrow -1$:

$$\begin{aligned} \pi_{KL:35}(\phi + \Delta) &> \pi_{JS:35}(\phi + \Delta) > \pi_{TKL:35}(\phi + \Delta) > \pi_{B:35}(\phi + \Delta) \\ &> \pi_{TRVC:35}(\phi + \Delta) \approx \pi_{TRVNC:35}(\phi + \Delta) > \pi_{H:35}(\phi + \Delta) \\ \pi_{KL:50}(\phi + \Delta) &> \pi_{JS:50}(\phi + \Delta) > \pi_{TKL:50}(\phi + \Delta) > \pi_{B:50}(\phi + \Delta) \\ &> \pi_{TRVC:50}(\phi + \Delta) \approx \pi_{TRVNC:50}(\phi + \Delta) > \pi_{H:50}(\phi + \Delta) \\ \pi_{KL:100}(\phi + \Delta) &\approx \pi_{JS:100}(\phi + \Delta) \approx \pi_{TKL:100}(\phi + \Delta) \approx \pi_{B:100}(\phi + \Delta) \\ &\approx \pi_{TRVC:100}(\phi + \Delta) \approx \pi_{TRVNC:100}(\phi + \Delta) > \pi_{H:100}(\phi + \Delta) \end{aligned}$$

Quando $\phi = 0.1$ e $\Delta \uparrow 0.8$:

$$\begin{aligned} \pi_{KL:35}(\phi + \Delta) &> \pi_{JS:35}(\phi + \Delta) > \pi_{RVNC:35}(\phi + \Delta) > \pi_{RVC:35}(\phi + \Delta) \\ &> \pi_{B:35}(\phi + \Delta) \approx \pi_{TKL:35}(\phi + \Delta) > \pi_{H:35}(\phi + \Delta) \\ \pi_{KL:50}(\phi + \Delta) &> \pi_{JS:50}(\phi + \Delta) > \pi_{RVNC:50}(\phi + \Delta) > \pi_{RVC:50}(\phi + \Delta) \\ &> \pi_{B:50}(\phi + \Delta) \approx \pi_{TKL:50}(\phi + \Delta) > \pi_{H:50}(\phi + \Delta) \\ \pi_{KL:100}(\phi + \Delta) &\approx \pi_{RVNC:100}(\phi + \Delta) \approx \pi_{RVC:100}(\phi + \Delta) \approx \pi_{JS:100}(\phi + \Delta) \\ &\approx \pi_{B:100}(\phi + \Delta) \approx \pi_{TKL:100}(\phi + \Delta) > \pi_{H:100}(\phi + \Delta) \end{aligned}$$

De acordo com a Tabela 4.4, podemos extrair os resultados: O poder empírico dos testes da razão de verossimilhanças com versões condicional e não condicional, são aproximadas.

Considerando o método de estimação MQNC, os poderes empíricos dos testes baseados nas distâncias *Kullback-Leibler*, *Jensen-Shannon* e *Bhattacharyya* são maiores quando comparados com o teste da razão de verossimilhança, obedecendo a seguinte relação

$$\pi_{\text{KL}:n=35,50}(\phi) > \pi_{\text{JS}:n=35,50}(\phi) > \pi_{\text{B}:n=35,50}(\phi).$$

Os testes em geral apresentam tamanhos empíricos próximos do nível nominal de 5%, com exceção da distância *Hellinger*. As seguintes desigualdades em relação aos respectivos poderes são satisfeitas:

Quando $\phi = 0.5$ e $\Delta \downarrow -1.4$:

$$\begin{aligned} \pi_{\text{TKL}:35;50}(\phi + \Delta) &> \pi_{\text{JS}:35;50}(\phi + \Delta) > \pi_{\text{B}:35;50}(\phi + \Delta) > \pi_{\text{KL}:35;50}(\phi + \Delta) \\ &> \pi_{\text{RVNC}:35;50}(\phi + \Delta) \approx \pi_{\text{RVC}:35;50}(\phi + \Delta) > \pi_{\text{H}:35;50}(\phi + \Delta) \\ \pi_{\text{TKL}:100}(\phi + \Delta) &> \pi_{\text{JS}:100}(\phi + \Delta) \approx \pi_{\text{B}:100}(\phi + \Delta) \approx \pi_{\text{KL}:100}(\phi + \Delta) \\ &\approx \pi_{\text{RVNC}:100}(\phi + \Delta) \approx \pi_{\text{RVC}:100}(\phi + \Delta) > \pi_{\text{H}:100}(\phi + \Delta) \end{aligned}$$

Quando $\phi = 0.5$ e $\Delta \uparrow 0.4$

$$\begin{aligned} \pi_{\text{KL}:35;50}(\phi + \Delta) &> \pi_{\text{JS}:35;50}(\phi + \Delta) > \pi_{\text{RVNC}:35;50}(\phi + \Delta) > \pi_{\text{RVC}:35;50}(\phi + \Delta) \\ &> \pi_{\text{B}:35;50}(\phi + \Delta) > \pi_{\text{TKL}:35;50}(\phi + \Delta) > \pi_{\text{H}:35;50}(\phi + \Delta) \\ \pi_{\text{KL}:100}(\phi + \Delta) &\approx \pi_{\text{JS}:100}(\phi + \Delta) \approx \pi_{\text{RVNC}:100}(\phi + \Delta) \approx \pi_{\text{RVC}:100}(\phi + \Delta) \\ &\approx \pi_{\text{B}:100}(\phi + \Delta) > \pi_{\text{TKL}:100}(\phi + \Delta) > \pi_{\text{H}:100}(\phi + \Delta) \end{aligned}$$

De acordo com a Tabela 4.5, podemos extrair os resultados: Como base na função poder empírica, tem-se que o teste da RVNC tem melhor desempenho do que aquele da RVC para tamanho de amostra pequeno ou moderado. Para $n = 100$, ambos os testes RVNC e RVC são assintoticamente equivalentes, como esperado. Considerando o método de estimação MVNC, os poderes empíricos dos testes baseados nas distâncias *Kullback-Leibler* e *Jensen-Shannon* são maiores quando comparados com o teste da razão de verossimilhança, obedecendo a seguinte relação

$$\pi_{\text{KL}:n=35,50}(\phi) > \pi_{\text{JS}:n=35,50}(\phi).$$

O tamanho empírico dos testes Tipo *Kullback-Leibler* e *Jensen-Shannon* apresentam taxas superiores ao nível nominal de 5%; entretanto tem poderes superior aos demais testes quando $\Delta \downarrow -1.7$. O poder do teste baseado na distância *Hellinger* para $\phi = 0.8$ é

competitivo em relação aos outros testes; por exemplo, para tamanho $n = 35$, tal teste teve melhor desempenho do que os testes da razão de verossimilhança. As seguintes desigualdades em relação aos respectivos poderes são satisfeitas:

Quando $\phi = 0.8$ e $\Delta \downarrow -1.7$:

$$\begin{aligned} \pi_{\text{TKL}:35}(\phi + \Delta) &> \pi_{\text{JS}:35}(\phi + \Delta) > \pi_{\text{B}:35}(\phi + \Delta) > \pi_{\text{KL}:35}(\phi + \Delta) \\ &> \pi_{\text{H}:35}(\phi + \Delta) \approx \pi_{\text{RVNC}:35}(\phi + \Delta) > \pi_{\text{RVC}:35}(\phi + \Delta) \\ \pi_{\text{TKL}:50}(\phi + \Delta) &> \pi_{\text{JS}:50}(\phi + \Delta) > \pi_{\text{B}:50}(\phi + \Delta) > \pi_{\text{KL}:50}(\phi + \Delta) \\ &> \pi_{\text{RVNC}:50}(\phi + \Delta) \approx \pi_{\text{RVC}:50}(\phi + \Delta) > \pi_{\text{H}:50}(\phi + \Delta) \\ \pi_{\text{TKL}:100}(\phi + \Delta) &> \pi_{\text{JS}:100}(\phi + \Delta) \approx \pi_{\text{B}:100}(\phi + \Delta) \approx \pi_{\text{KL}:100}(\phi + \Delta) \\ &\approx \pi_{\text{RVNC}:100}(\phi + \Delta) \approx \pi_{\text{RVC}:100}(\phi + \Delta) > \pi_{\text{H}:100}(\phi + \Delta) \end{aligned}$$

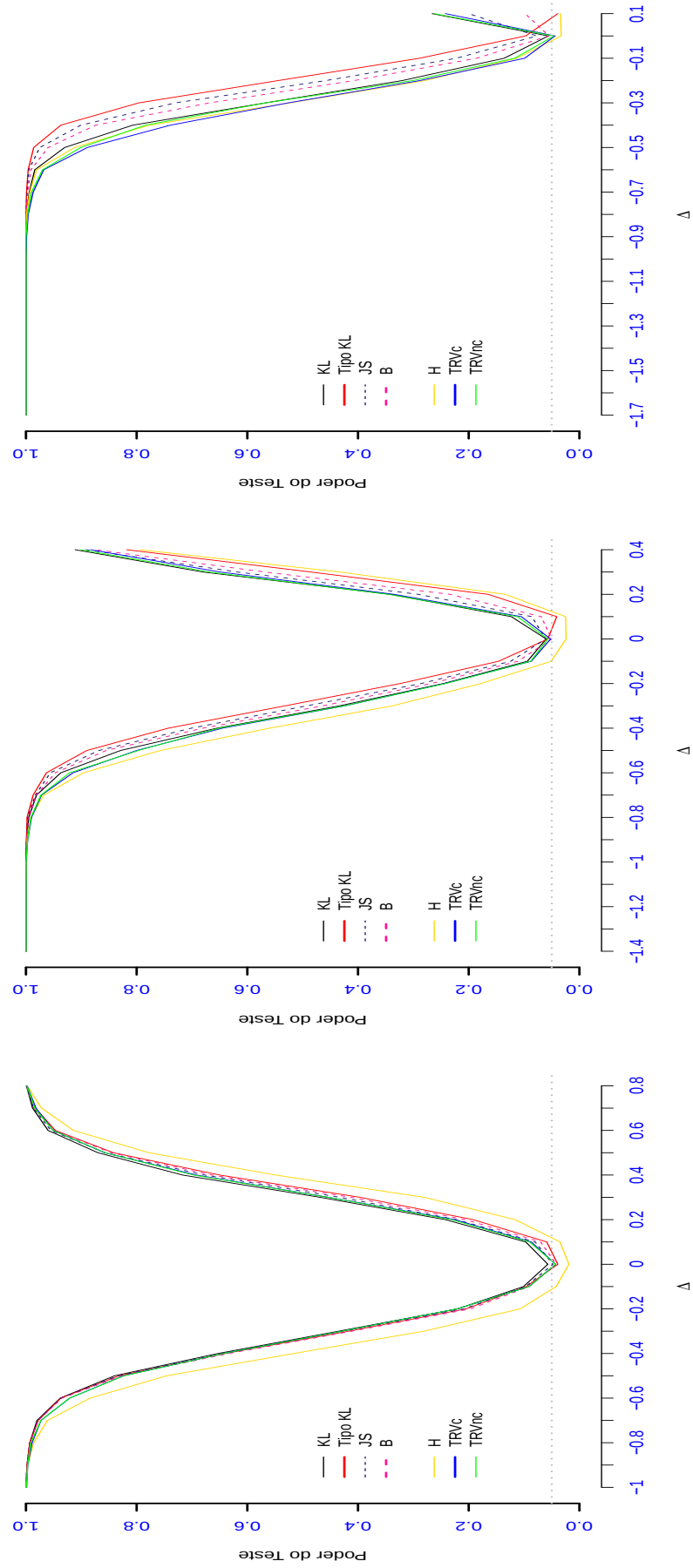
Quando $\phi = 0.8$ e $\Delta = 0.1$:

$$\begin{aligned} \pi_{\text{KL}:35}(\phi + \Delta) &> \pi_{\text{RVNC}:35}(\phi + \Delta) > \pi_{\text{RVC}:35}(\phi + \Delta) > \pi_{\text{JS}:35}(\phi + \Delta) \\ &> \pi_{\text{B}:35}(\phi + \Delta) \approx \pi_{\text{TKL}:35}(\phi + \Delta) > \pi_{\text{H}:35}(\phi + \Delta) \\ \pi_{\text{KL}:50}(\phi + \Delta) &> \pi_{\text{RVNC}:50}(\phi + \Delta) > \pi_{\text{RVC}:50}(\phi + \Delta) > \pi_{\text{JS}:50}(\phi + \Delta) \\ &> \pi_{\text{B}:50}(\phi + \Delta) \approx \pi_{\text{TKL}:50}(\phi + \Delta) > \pi_{\text{H}:50}(\phi + \Delta) \\ \pi_{\text{KL}:100}(\phi + \Delta) &> \pi_{\text{RVNC}:100}(\phi + \Delta) > \pi_{\text{RVC}:100}(\phi + \Delta) > \pi_{\text{JS}:100}(\phi + \Delta) \\ &> \pi_{\text{B}:100}(\phi + \Delta) > \pi_{\text{TKL}:100}(\phi + \Delta) > \pi_{\text{H}:100}(\phi + \Delta) \end{aligned}$$

Conclusivamente,

para processos $AR(1)$, os testes com estatísticas S_{KL} , S_{TKL} , S_{JS} e S_{B} combinados com a estimação MVNC se mostraram bastante competitivos e em alguns casos mais robustos do que o teste com base em S_{RVC} e S_{RVNC} , que são altamente utilizado na literatura.

Figura 4.29: Poder do Teste considerando $\phi = 0.1$, $\phi = 0.5$ e $\phi = 0.8$ para o processo Autoregressivo de ordem 1 com tamanho $n = 35$.



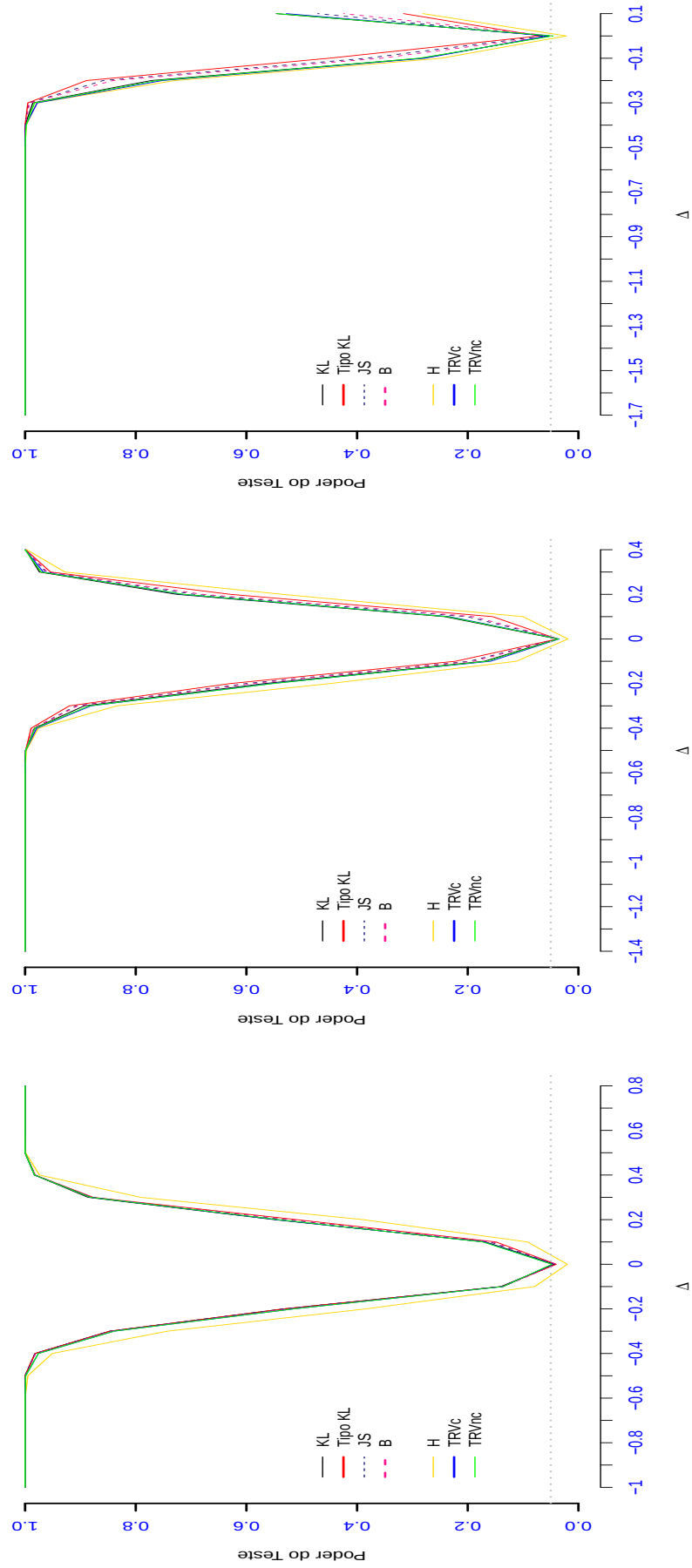
(a) $\phi = 0.1$ (EMVNC)

(b) $\phi = 0.5$ (EMVNC)

(c) $\phi = 0.8$ (EMVNC)

Fonte: Autoria própria.

Figura 4.30: Poder do Teste considerando $\phi = 0.1$, $\phi = 0.5$ e $\phi = 0.8$ para o processo Autoregressivo de ordem 1 com tamanho $n = 100$.



(a) $\phi = 0.1$ (EMVNC)

(b) $\phi = 0.5$ (EMVNC)

(c) $\phi = 0.8$ (EMVNC)

Fonte: Autoria própria.

[illegible]

Tabela 4.4: Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $AR(1)$ com $\phi = 0,5$

η	Δ	Condicional						Não Condicional						Mínimos Quadrados						TRVNC	TRVC	TRVNC							
		KL	TipoKL	JS	B	H	TRVNC	TRVC	KL	TipoKL	JS	B	H	TRVNC	TRVC	KL	TipoKL	JS	B				H						
35	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	0.999	0.999	0.999	0.999	0.999	0.997	0.997	0.997	0.999	0.999	0.999	0.999	0.998	0.997	0.997	0.997	0.999	0.999	0.999	0.998	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997
	-0.8	0.995	0.998	0.995	0.995	0.990	0.990	0.990	0.991	0.995	0.997	0.995	0.991	0.990	0.990	0.991	0.995	0.998	0.995	0.995	0.990	0.990	0.990	0.990	0.990	0.991	0.991	0.991	0.991
	-0.7	0.980	0.986	0.983	0.983	0.968	0.983	0.973	0.973	0.983	0.987	0.983	0.972	0.968	0.972	0.973	0.995	0.986	0.985	0.985	0.965	0.965	0.965	0.965	0.965	0.965	0.965	0.965	0.965
	-0.6	0.937	0.958	0.953	0.944	0.896	0.915	0.921	0.921	0.937	0.963	0.956	0.947	0.896	0.915	0.921	0.934	0.958	0.951	0.941	0.888	0.915	0.921	0.921	0.921	0.921	0.921	0.921	0.921
50	-0.5	0.832	0.889	0.867	0.854	0.755	0.800	0.797	0.829	0.890	0.868	0.855	0.757	0.800	0.797	0.829	0.890	0.868	0.855	0.757	0.800	0.797	0.829	0.890	0.868	0.855	0.757	0.800	0.797
	-0.4	0.656	0.745	0.698	0.682	0.557	0.645	0.650	0.655	0.743	0.701	0.680	0.560	0.645	0.651	0.646	0.742	0.725	0.684	0.672	0.548	0.651	0.646	0.742	0.725	0.684	0.672	0.548	0.651
	-0.3	0.435	0.531	0.493	0.465	0.339	0.424	0.425	0.434	0.532	0.486	0.464	0.338	0.424	0.425	0.434	0.532	0.486	0.464	0.338	0.424	0.425	0.434	0.532	0.486	0.464	0.338	0.424	0.425
	-0.2	0.246	0.328	0.287	0.269	0.180	0.244	0.242	0.246	0.324	0.286	0.268	0.178	0.244	0.242	0.237	0.307	0.274	0.273	0.259	0.170	0.244	0.242	0.237	0.307	0.274	0.273	0.259	0.170
	-0.1	0.099	0.145	0.125	0.111	0.055	0.087	0.088	0.094	0.146	0.124	0.111	0.051	0.087	0.088	0.094	0.135	0.121	0.102	0.050	0.087	0.088	0.094	0.135	0.121	0.102	0.050	0.087	0.088
	0	0.071	0.055	0.057	0.051	0.024	0.052	0.056	0.060	0.057	0.061	0.056	0.025	0.052	0.057	0.078	0.057	0.068	0.058	0.027	0.052	0.057	0.078	0.057	0.068	0.058	0.027	0.052	0.057
	0.1	0.126	0.045	0.096	0.073	0.028	0.108	0.110	0.123	0.041	0.086	0.069	0.025	0.105	0.112	0.169	0.073	0.135	0.105	0.047	0.106	0.110	0.169	0.073	0.135	0.105	0.047	0.106	0.110
	0.2	0.354	0.181	0.286	0.243	0.138	0.339	0.343	0.342	0.165	0.291	0.234	0.134	0.335	0.345	0.423	0.237	0.335	0.320	0.130	0.333	0.344	0.423	0.237	0.335	0.320	0.130	0.333	0.344
	0.3	0.672	0.470	0.610	0.568	0.424	0.654	0.667	0.680	0.485	0.618	0.568	0.433	0.649	0.669	0.741	0.569	0.690	0.647	0.530	0.643	0.667	0.741	0.569	0.690	0.647	0.530	0.643	0.667
0.4	0.899	0.812	0.882	0.850	0.779	0.891	0.899	0.911	0.818	0.892	0.875	0.794	0.883	0.891	0.931	0.873	0.915	0.906	0.853	0.880	0.899	0.931	0.873	0.915	0.906	0.853	0.880	0.899	
100	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
150	-0.5	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
	-0.4	0.983	0.987	0.985	0.985	0.977	0.979	0.982	0.983	0.989	0.984	0.984	0.976	0.981	0.982	0.982	0.986	0.984	0.984	0.976	0.981	0.982	0.982	0.986	0.984	0.984	0.976	0.981	0.982
	-0.3	0.889	0.914	0.895	0.900	0.835	0.881	0.884	0.893	0.920	0.908	0.905	0.834	0.881	0.884	0.884	0.911	0.903	0.899	0.882	0.911	0.884	0.884	0.911	0.903	0.899	0.882	0.911	0.884
	-0.2	0.571	0.626	0.602	0.592	0.447	0.555	0.551	0.567	0.629	0.600	0.589	0.449	0.555	0.551	0.559	0.616	0.584	0.574	0.440	0.555	0.551	0.559	0.616	0.584	0.574	0.440	0.555	0.551
	-0.1	0.163	0.217	0.192	0.187	0.110	0.159	0.165	0.170	0.222	0.197	0.189	0.112	0.159	0.165	0.155	0.209	0.189	0.182	0.107	0.159	0.165	0.155	0.209	0.189	0.182	0.107	0.159	0.165
	0	0.038	0.039	0.036	0.036	0.020	0.035	0.036	0.038	0.038	0.037	0.036	0.019	0.035	0.036	0.038	0.038	0.038	0.038	0.021	0.035	0.036	0.038	0.038	0.038	0.021	0.035	0.036	0.036
	0.1	0.243	0.153	0.203	0.196	0.107	0.234	0.237	0.242	0.155	0.201	0.190	0.100	0.233	0.237	0.246	0.175	0.234	0.227	0.126	0.233	0.237	0.246	0.175	0.234	0.227	0.126	0.233	0.237
	0.2	0.721	0.627	0.694	0.683	0.540	0.716	0.713	0.726	0.629	0.694	0.679	0.535	0.716	0.713	0.746	0.661	0.713	0.711	0.568	0.716	0.713	0.746	0.661	0.713	0.711	0.568	0.716	0.713
	0.3	0.969	0.950	0.958	0.956	0.918	0.969	0.971	0.974	0.954	0.962	0.960	0.928	0.967	0.971	0.977	0.960	0.972	0.970	0.941	0.967	0.971	0.977	0.960	0.972	0.970	0.941	0.967	0.971
0.4	1.000	0.999	0.999	0.999	0.996	0.996	1.000	0.999	0.999	0.999	0.999	0.999	0.997	1.000	0.999	0.999	0.999	0.999	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.998	0.999	0.999	

Tabela 4.5: Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $AR(1)$ com $\phi = 0, 8$ e $\sigma^2 = 1$ conhecido

n	Δ	Condicional					Não Condicional					Mínimos Quadrados									
		KL	TipoKL	JS	H	TRVC	TRVNC	KL	TipoKL	JS	B	H	TRVC	TRVNC	KL	TipoKL	JS	B	H	TRVC	TRVNC
35	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	0.999	1.000	1.000	1.000	1.000	0.997	1.000	0.998	1.000	1.000	1.000	0.999	1.000	0.996	1.000	1.000	1.000	1.000	0.999	1.000
50	-0.7	0.995	0.999	0.998	0.998	0.995	0.987	0.990	0.995	0.999	0.997	0.993	0.987	0.990	0.993	0.998	0.997	0.996	0.992	0.987	0.990
	-0.6	0.984	0.999	0.994	0.992	0.979	0.968	0.970	0.984	0.996	0.991	0.981	0.968	0.970	0.980	0.995	0.991	0.988	0.974	0.968	0.969
	-0.5	0.931	0.987	0.974	0.958	0.913	0.889	0.901	0.930	0.986	0.975	0.960	0.912	0.889	0.901	0.910	0.978	0.963	0.946	0.893	0.901
	-0.4	0.802	0.931	0.904	0.871	0.771	0.743	0.783	0.807	0.937	0.901	0.874	0.777	0.739	0.784	0.780	0.916	0.876	0.839	0.742	0.781
	-0.3	0.566	0.793	0.732	0.668	0.520	0.527	0.565	0.567	0.796	0.730	0.668	0.523	0.526	0.567	0.527	0.749	0.672	0.622	0.479	0.526
	-0.2	0.319	0.552	0.470	0.414	0.290	0.291	0.301	0.320	0.544	0.458	0.407	0.279	0.288	0.301	0.282	0.478	0.415	0.364	0.244	0.300
	-0.1	0.138	0.288	0.226	0.192	0.114	0.100	0.116	0.136	0.222	0.186	0.113	0.099	0.117	0.118	0.242	0.188	0.159	0.095	0.096	0.116
	0	0.068	0.098	0.090	0.072	0.037	0.045	0.044	0.056	0.097	0.077	0.055	0.033	0.044	0.046	0.104	0.087	0.072	0.037	0.045	0.045
	0.1	0.287	0.071	0.230	0.142	0.062	0.248	0.249	0.266	0.039	0.200	0.098	0.034	0.242	0.263	0.436	0.158	0.364	0.172	0.238	0.233
	0.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.4	0.999	1.000	1.000	1.000	1.000	0.999	0.998	1.000	0.999	1.000	1.000	0.998	0.999	0.998	0.999	0.999	1.000	1.000	0.997	0.999
	-0.3	0.986	0.995	0.991	0.990	0.977	0.978	0.983	0.986	0.995	0.993	0.992	0.978	0.983	0.983	0.983	0.993	0.988	0.975	0.978	0.983
	-0.2	0.776	0.892	0.852	0.834	0.727	0.754	0.760	0.772	0.889	0.859	0.841	0.729	0.752	0.761	0.750	0.873	0.810	0.702	0.750	0.761
	-0.1	0.289	0.452	0.384	0.351	0.251	0.278	0.283	0.284	0.451	0.383	0.357	0.248	0.278	0.288	0.265	0.416	0.350	0.216	0.278	0.287
	0	0.061	0.080	0.070	0.061	0.027	0.057	0.046	0.050	0.070	0.063	0.048	0.022	0.056	0.046	0.061	0.058	0.062	0.026	0.054	0.046
	0.1	0.541	0.308	0.471	0.423	0.286	0.531	0.542	0.546	0.316	0.471	0.424	0.281	0.528	0.545	0.618	0.400	0.561	0.379	0.524	0.544
	0.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

4.2.2 Testes de identificação para classe de modelo MA(1)

No que tange ao modelo MA(1), as versões empíricas da função poder dos testes foram avaliadas para $\theta = 0.1, 0.5$ e 0.8 , representadas nas Tabelas 4.6, 4.7 e 4.8, respectivamente. Os poderes empíricos dos testes para EMVNC com os tamanhos amostrais $n = 35$ e 100 são apresentados nas Figuras 4.31 e 4.32.

Para o primeiro caso com $\theta = 0.1$ apresentado na Tabela 4.6, segue-se discussão. Para $n = 35$ e 50 , identifica-se uma diferença expressiva entre os poderes dos testes da razão entre verossimilhanças condicional e não condicional avaliadas nos métodos de estimação MVNC e MQNC. Em particular, o teste da RVC apresentou menor poder em relação ao teste da RVNC. Assim, existe um ganho evidente em utilizar uma verossimilhança com estrutura não condicional quando o tamanho de amostra é pequeno. Por outro lado, os testes a luz das medidas de distância (com exceção daquele baseado na distância de *Hellinger*) para tamanhos de amostras $n = 35$ e 50 têm poder empírico superior aos testes da razão de verossimilhanças. Entretanto, em termos do tamanho empírico dos testes, os testes das medidas Tipo *Kullback-Leibler*, *Jensen-Shannon* e *Bhattacharyya* apresentam tamanhos acima do nível especificado. Isto é, eles tendem a errar mais na captura de similaridade (ou equivalentemente, têm maiores erro do tipo I empírico ou maiores taxas de alarme falso). Vale salientar que, para $n = 100$, os testes se aproximam do nível nominal e os poderes dos testes são assintoticamente similares. As seguintes desigualdades podem ser verificadas:

Quando $\theta = 0.1$ e $\Delta \downarrow -1$ e $\Delta \uparrow 0.8$:

$$\begin{aligned} \pi_{\text{TKL:35}}(\theta + \Delta) &> \pi_{\text{JS:35}}(\theta + \Delta) > \pi_{\text{B:35}}(\theta + \Delta) > \pi_{\text{KL:35}}(\theta + \Delta) \\ &> \pi_{\text{RVNC:35}}(\theta + \Delta) \approx \pi_{\text{RVC:35}}(\theta + \Delta) > \pi_{\text{H:35}}(\theta + \Delta) \\ \pi_{\text{TKL:50}}(\theta + \Delta) &> \pi_{\text{JS:50}}(\theta + \Delta) > \pi_{\text{B:50}}(\theta + \Delta) > \pi_{\text{KL:50}}(\theta + \Delta) \\ &> \pi_{\text{RVNC:50}}(\theta + \Delta) \approx \pi_{\text{RVC:50}}(\theta + \Delta) > \pi_{\text{H:50}}(\theta + \Delta) \\ \pi_{\text{TKL:100}}(\theta + \Delta) &> \pi_{\text{JS:100}}(\theta + \Delta) \approx \pi_{\text{B:100}}(\theta + \Delta) \approx \pi_{\text{KL:100}}(\theta + \Delta) \\ &\approx \pi_{\text{RVNC:100}}(\theta + \Delta) \approx \pi_{\text{RVC:100}}(\theta + \Delta) > \pi_{\text{H:100}}(\theta + \Delta) \end{aligned}$$

O segundo caso, com $\theta = 0.5$, é apresentado na Tabela 4.7. Para $n = 35$, os testes em geral apresentaram tamanhos empíricos acima do nível de significância adotado de 5%, com exceção do teste da RVNC. Para $n = 50$, todos os testes tiveram tamanhos empíricos elevados; entretanto, os tamanhos para os testes induzidos das medidas são maiores em relação aos testes da razão de verossimilhanças. Apenas para $n = 100$, os tamanhos empíricos se aproximam do nível adotado. Os testes induzidos das distâncias *Kullback-*

Leibler, Tipo *Kullback-Leibler*, *Jensen-Shannon* e *Bhattacharyya* apresentam maior poder empírico em relação aos testes da razão entre verossimilhanças quando $\Delta \uparrow 0.4$. Os testes são aproximados quando $\Delta \downarrow -1.4$, com exceção da distância *Hellinger*. As seguintes desigualdades podem ser verificadas neste caso:

Quando $\theta = 0.5$ e $\Delta \downarrow -1.4$:

$$\begin{aligned} \pi_{KL:35}(\theta + \Delta) &> \pi_{JS:35}(\theta + \Delta) > \pi_{B:35}(\theta + \Delta) > \pi_{TKL:35}(\theta + \Delta) \\ &> \pi_{RVNC:35}(\theta + \Delta) \approx \pi_{RVC:35}(\theta + \Delta) > \pi_{H:35}(\theta + \Delta) \\ \pi_{KL:50}(\theta + \Delta) &> \pi_{JS:50}(\theta + \Delta) > \pi_{B:50}(\theta + \Delta) > \pi_{TKL:50}(\theta + \Delta) \\ &> \pi_{RVNC:50}(\theta + \Delta) \approx \pi_{RVC:50}(\theta + \Delta) > \pi_{H:50}(\theta + \Delta) \\ \pi_{KL:100}(\theta + \Delta) &> \pi_{JS:100}(\theta + \Delta) \approx \pi_{B:100}(\theta + \Delta) \approx \pi_{TKL:100}(\theta + \Delta) \\ &\approx \pi_{RVNC:100}(\theta + \Delta) \approx \pi_{RVC:100}(\theta + \Delta) > \pi_{H:100}(\theta + \Delta) \end{aligned}$$

Quando $\theta = 0.5$ e $\Delta \uparrow 0.4$ e Métodos MVNC e MQNC:

$$\begin{aligned} \pi_{TKL:35}(\theta + \Delta) &> \pi_{JS:35}(\theta + \Delta) > \pi_{B:35}(\theta + \Delta) > \pi_{KL:35}(\theta + \Delta) \\ &> \pi_{RVNC:35}(\theta + \Delta) \approx \pi_{RVC:35}(\theta + \Delta) > \pi_{H:35}(\theta + \Delta) \\ \pi_{TKL:50}(\theta + \Delta) &> \pi_{JS:50}(\theta + \Delta) > \pi_{B:50}(\theta + \Delta) > \pi_{KL:50}(\theta + \Delta) \\ &> \pi_{RVNC:50}(\theta + \Delta) \approx \pi_{RVC:50}(\theta + \Delta) > \pi_{H:50}(\theta + \Delta) \\ \pi_{TKL:100}(\theta + \Delta) &> \pi_{JS:100}(\theta + \Delta) \approx \pi_{B:100}(\theta + \Delta) \approx \pi_{KL:100}(\theta + \Delta) \\ &\approx \pi_{RVNC:100}(\theta + \Delta) \approx \pi_{RVC:100}(\theta + \Delta) > \pi_{H:100}(\theta + \Delta) \end{aligned}$$

Para $\theta = 0.8$, os poderes dos testes são apresentadas na Tabela 4.8. Identifica-se uma notória diferença entre os testes da razão de verossimilhanças com estruturas condicional e não condicional quando $n = 35$ e 50 . O teste da RVC apresentou um tamanho empírico do teste (quando $\Delta = 0$) bastante elevado, enquanto o teste da RVNC apresentou uma taxa de rejeição próxima do nível nominal. Em relação ao tamanho empírico dos testes induzidos da divergência de *Kullback-Leibler*, o tamanho tem melhor aproximação do nível nominal quando o método de estimação é MVNC. Deste modo, o teste da divergência Δ_{KL} tem seu desempenho otimizado com EMVNC.

O teste RVNC teve melhor desempenho em termos do tamanho do teste, entretanto, o poder empírico do teste em função da divergência de *Kullback-Leibler* teve melhor performance.

Conclusivamente,

os testes em geral apresentaram um tamanho empírico acima do nível nominal para identificação em processos MA(1). Dentre eles, o teste da RVNC combinado com a EMVC apresentou o melhor resultado.

Tabela 4.6: Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $MA(1)$ com $\theta = 0.1$

n	Δ	Condicional					Não Condicional					Mínimos Quadrados										
		KL	TKL	JS	B	H	TRVC	TRVNC	KL	TKL	JS	B	H	TRVC	TRVNC	KL	TKL	JS	B	H	TRVC	TRVNC
35	-1.	0.997	0.998	0.997	0.997	0.997	0.994	0.997	0.997	0.999	0.999	0.999	0.999	0.999	0.998	0.999	0.999	0.999	0.999	0.999	0.958	0.998
	-0.9	0.995	0.996	0.995	0.995	0.995	0.993	0.994	0.994	0.998	0.998	0.998	0.998	0.998	0.995	0.998	0.998	0.998	0.998	0.997	0.969	0.995
	-0.8	0.994	0.994	0.994	0.994	0.990	0.979	0.980	0.980	0.995	0.995	0.995	0.991	0.991	0.965	0.995	0.995	0.995	0.995	0.994	0.948	0.980
	-0.7	0.980	0.981	0.981	0.980	0.970	0.947	0.951	0.985	0.987	0.987	0.985	0.975	0.975	0.940	0.989	0.989	0.989	0.987	0.978	0.934	0.951
	-0.6	0.947	0.950	0.950	0.941	0.892	0.899	0.903	0.951	0.953	0.953	0.949	0.900	0.900	0.895	0.904	0.957	0.960	0.959	0.955	0.912	0.892
	-0.5	0.835	0.843	0.838	0.829	0.739	0.785	0.799	0.843	0.848	0.844	0.840	0.741	0.741	0.783	0.802	0.851	0.858	0.858	0.847	0.760	0.781
	-0.4	0.649	0.658	0.654	0.643	0.508	0.614	0.612	0.653	0.669	0.658	0.642	0.518	0.518	0.610	0.616	0.671	0.685	0.679	0.664	0.534	0.608
	-0.3	0.424	0.441	0.430	0.417	0.308	0.413	0.414	0.412	0.428	0.422	0.410	0.303	0.303	0.410	0.421	0.430	0.453	0.437	0.422	0.325	0.412
	-0.2	0.259	0.271	0.262	0.254	0.167	0.231	0.231	0.249	0.258	0.252	0.239	0.156	0.156	0.228	0.233	0.274	0.293	0.283	0.272	0.188	0.236
	-0.1	0.125	0.137	0.131	0.120	0.076	0.096	0.095	0.118	0.130	0.123	0.117	0.069	0.093	0.098	0.133	0.142	0.139	0.129	0.080	0.093	0.099
50	0	0.085	0.103	0.094	0.086	0.056	0.053	0.053	0.083	0.094	0.091	0.086	0.052	0.051	0.055	0.095	0.108	0.101	0.095	0.061	0.052	0.055
	0.1	0.132	0.161	0.149	0.138	0.085	0.107	0.095	0.130	0.154	0.142	0.132	0.080	0.102	0.095	0.142	0.169	0.159	0.147	0.098	0.104	0.094
	0.2	0.236	0.276	0.253	0.241	0.179	0.223	0.220	0.236	0.277	0.262	0.248	0.172	0.216	0.223	0.266	0.303	0.285	0.269	0.196	0.217	0.222
	0.3	0.415	0.477	0.452	0.429	0.320	0.423	0.419	0.421	0.485	0.456	0.435	0.319	0.419	0.430	0.458	0.521	0.492	0.471	0.350	0.409	0.427
	0.4	0.670	0.740	0.707	0.690	0.541	0.625	0.630	0.686	0.748	0.712	0.695	0.566	0.613	0.638	0.712	0.775	0.752	0.732	0.605	0.611	0.634
	0.5	0.867	0.896	0.884	0.878	0.802	0.817	0.824	0.878	0.904	0.892	0.882	0.808	0.808	0.829	0.893	0.919	0.906	0.903	0.844	0.787	0.821
	0.6	0.959	0.965	0.964	0.961	0.931	0.909	0.919	0.969	0.973	0.970	0.969	0.945	0.945	0.925	0.972	0.978	0.974	0.973	0.957	0.890	0.921
	0.7	0.987	0.990	0.989	0.988	0.976	0.963	0.964	0.992	0.994	0.993	0.992	0.982	0.928	0.967	0.995	0.996	0.996	0.995	0.987	0.924	0.962
	0.8	0.995	0.996	0.996	0.995	0.991	0.982	0.987	0.998	0.998	0.998	0.998	0.997	0.950	0.991	0.998	0.998	0.998	0.998	0.998	0.934	0.990
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.986	1.000
100	-0.9	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.993	1.000	1.000	1.000	1.000	0.991	1.000	
	-0.8	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	1.000	1.000	1.000	1.000	0.994	1.000	
	-0.7	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.992	1.000	1.000	1.000	1.000	0.999	1.000	
	-0.6	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.990	1.000	1.000	1.000	1.000	0.997	1.000	
	-0.5	0.999	0.999	0.999	0.999	0.999	0.996	0.996	0.997	0.997	0.997	0.997	0.994	0.996	0.996	0.996	0.997	0.997	0.997	0.994	0.996	
	-0.4	0.981	0.981	0.981	0.981	0.955	0.967	0.968	0.979	0.979	0.979	0.979	0.957	0.966	0.968	0.979	0.979	0.979	0.979	0.958	0.966	
	-0.3	0.864	0.864	0.864	0.859	0.762	0.844	0.847	0.865	0.865	0.865	0.860	0.772	0.844	0.848	0.868	0.868	0.868	0.862	0.776	0.844	
	-0.2	0.525	0.525	0.525	0.517	0.367	0.517	0.508	0.520	0.518	0.518	0.512	0.367	0.517	0.509	0.525	0.525	0.525	0.517	0.373	0.517	
	-0.1	0.148	0.149	0.149	0.143	0.085	0.145	0.140	0.139	0.140	0.139	0.138	0.083	0.145	0.140	0.140	0.142	0.162	0.158	0.149	0.089	0.107
	0	0.061	0.064	0.061	0.061	0.032	0.052	0.051	0.057	0.062	0.060	0.058	0.033	0.057	0.051	0.060	0.067	0.064	0.062	0.034	0.052	0.051

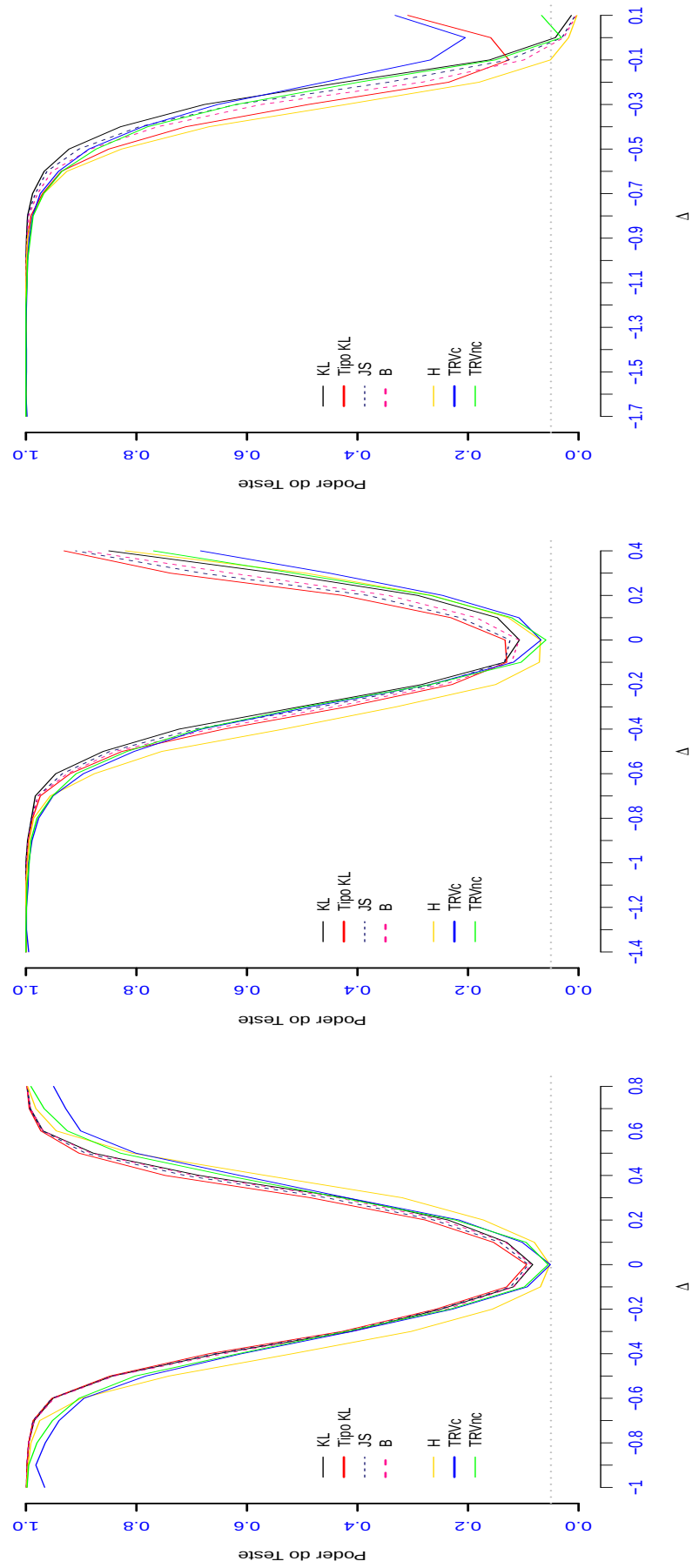
Tabela 4.7: Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $MA(1)$ com $\theta = 0.5$

n	Δ	Condicional					Não Condicional					Mínimos Quadrados											
		KL	TKL	JS	B	H	$TRVC$	$TRVNC$	KL	TKL	JS	B	H	$TRVC$	$TRVNC$	KL	TKL	JS	B	H	$TRVC$	$TRVNC$	
35	-1.4	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	1.000	1.000	1.000	1.000	1.000	0.983	1.000	
	-1.3	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	0.997	1.000	
	-1.2	1.000	1.000	1.000	1.000	1.000	0.998	0.998	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	0.997	0.999	
	-1.1	1.000	1.000	1.000	1.000	1.000	0.999	0.996	1.000	1.000	1.000	1.000	1.000	1.000	0.996	1.000	1.000	1.000	1.000	1.000	0.997	0.997	
	-1.0	0.999	0.997	0.998	0.998	0.996	0.995	0.994	1.000	0.998	0.999	0.997	0.997	0.997	0.995	0.994	1.000	0.998	0.999	0.998	0.997	0.993	
	-0.9	0.996	0.995	0.996	0.995	0.993	0.989	0.990	0.997	0.995	0.996	0.996	0.994	0.989	0.989	0.991	0.996	0.995	0.995	0.995	0.993	0.994	
	-0.8	0.991	0.990	0.991	0.990	0.986	0.978	0.980	0.990	0.988	0.990	0.988	0.986	0.986	0.977	0.980	0.990	0.988	0.988	0.986	0.977	0.980	
	-0.7	0.982	0.971	0.980	0.974	0.950	0.953	0.952	0.983	0.974	0.978	0.956	0.951	0.951	0.952	0.980	0.979	0.970	0.977	0.974	0.954	0.952	
	-0.6	0.938	0.914	0.927	0.921	0.870	0.895	0.905	0.946	0.920	0.934	0.925	0.876	0.896	0.896	0.909	0.941	0.915	0.926	0.921	0.873	0.894	0.903
	-0.5	0.854	0.814	0.841	0.829	0.746	0.807	0.815	0.860	0.826	0.848	0.838	0.754	0.804	0.804	0.817	0.855	0.818	0.841	0.833	0.747	0.805	0.818
50	-0.4	0.722	0.645	0.684	0.664	0.537	0.683	0.686	0.723	0.642	0.694	0.666	0.534	0.682	0.687	0.716	0.635	0.684	0.660	0.531	0.684	0.688	
	-0.3	0.507	0.419	0.457	0.440	0.332	0.486	0.493	0.505	0.419	0.468	0.442	0.330	0.485	0.496	0.501	0.413	0.460	0.437	0.327	0.490	0.497	
	-0.2	0.294	0.240	0.269	0.254	0.160	0.270	0.265	0.284	0.229	0.259	0.244	0.150	0.269	0.278	0.242	0.232	0.257	0.234	0.154	0.276	0.268	
	-0.1	0.141	0.134	0.137	0.125	0.078	0.118	0.102	0.134	0.130	0.134	0.121	0.071	0.118	0.104	0.153	0.149	0.152	0.134	0.094	0.127	0.104	
	0	0.098	0.124	0.114	0.103	0.059	0.070	0.056	0.107	0.133	0.124	0.108	0.069	0.068	0.059	0.144	0.175	0.158	0.144	0.111	0.097	0.054	
	0.1	0.136	0.211	0.184	0.160	0.114	0.103	0.096	0.147	0.231	0.214	0.185	0.127	0.108	0.121	0.214	0.290	0.262	0.247	0.183	0.142	0.107	
	0.2	0.228	0.360	0.316	0.279	0.197	0.246	0.237	0.291	0.427	0.381	0.344	0.265	0.246	0.268	0.375	0.509	0.460	0.429	0.345	0.294	0.250	
	0.3	0.410	0.627	0.560	0.504	0.364	0.442	0.463	0.547	0.739	0.682	0.629	0.494	0.451	0.516	0.655	0.820	0.762	0.725	0.614	0.510	0.484	
	0.4	0.646	0.816	0.776	0.743	0.591	0.671	0.703	0.850	0.931	0.910	0.894	0.820	0.684	0.769	0.916	0.967	0.953	0.941	0.893	0.736	0.733	
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	0.993	1.000	
100	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	1.000	
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	
	-1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.997	
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.997	
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.997	
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.997	
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.997	
	-0.5	0.998	0.997	0.998	0.998	0.994	0.992	0.994	0.998	0.998	0.998	0.998	0.993	0.992	0.994	0.998	0.996	0.998	0.998	0.993	0.992	0.994	
	-0.4	0.981	0.970	0.978	0.976	0.957	0.976	0.974	0.978	0.971	0.976	0.975	0.956	0.976	0.974	0.978	0.969	0.973	0.973	0.955	0.976	0.974	

Tabela 4.8: Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $MA(1)$ com $\theta = 0.8$

n	Δ	Conditional					Não Condicional					Mínimos Quadrados										
		KL	TKL	JS	B	H	$TRVC$	$TRVNC$	KL	TKL	JS	B	H	$TRVC$	$TRVNC$	KL	TKL	JS	B	H	$TRVC$	$TRVNC$
35	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	0.998	0.998	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000
	-1	1.000	0.999	1.000	1.000	0.998	0.997	0.997	1.000	0.999	1.000	1.000	1.000	0.999	0.997	0.997	0.998	0.999	0.998	0.998	0.996	0.996
	-0.9	0.998	0.997	0.998	0.996	0.997	0.994	0.992	0.997	0.999	0.997	0.998	0.998	0.997	0.994	0.992	0.996	0.997	0.996	0.994	0.994	0.992
	-0.8	0.995	0.992	0.995	0.994	0.988	0.988	0.987	0.997	0.991	0.997	0.995	0.988	0.987	0.988	0.987	0.995	0.987	0.993	0.991	0.984	0.987
	-0.7	0.990	0.974	0.987	0.982	0.966	0.974	0.966	0.988	0.970	0.983	0.979	0.968	0.974	0.968	0.983	0.970	0.974	0.969	0.957	0.974	0.967
50	-0.6	0.971	0.940	0.963	0.952	0.919	0.941	0.933	0.967	0.936	0.962	0.953	0.926	0.941	0.935	0.957	0.929	0.944	0.933	0.908	0.943	0.934
	-0.5	0.926	0.849	0.899	0.874	0.817	0.886	0.873	0.922	0.850	0.904	0.883	0.827	0.886	0.873	0.904	0.835	0.881	0.864	0.801	0.888	0.873
	-0.4	0.828	0.719	0.793	0.766	0.672	0.789	0.779	0.829	0.711	0.795	0.762	0.669	0.789	0.780	0.800	0.704	0.763	0.732	0.636	0.800	0.780
	-0.3	0.683	0.517	0.639	0.586	0.445	0.650	0.618	0.677	0.489	0.619	0.574	0.421	0.653	0.620	0.635	0.497	0.575	0.521	0.377	0.683	0.618
	-0.2	0.453	0.254	0.367	0.306	0.192	0.439	0.395	0.425	0.235	0.344	0.283	0.180	0.460	0.398	0.369	0.315	0.294	0.240	0.146	0.509	0.400
	-0.1	0.186	0.129	0.153	0.121	0.061	0.227	0.153	0.161	0.126	0.127	0.099	0.051	0.268	0.151	0.125	0.291	0.096	0.079	0.041	0.361	0.151
	0	0.061	0.118	0.041	0.035	0.023	0.086	0.029	0.042	0.159	0.030	0.027	0.018	0.205	0.031	0.030	0.448	0.025	0.022	0.012	0.349	0.034
	0.1	0.021	0.170	0.018	0.016	0.009	0.096	0.027	0.013	0.309	0.006	0.004	0.003	0.332	0.067	0.004	0.738	0.004	0.003	0.002	0.541	0.067
	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
100	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	0.999	0.999	0.999	0.999	0.997	0.997	0.997	0.999	0.999	0.999	0.999	0.996	0.997	0.997	0.999	0.996	0.996	0.999	0.996	0.997	0.997
	-0.4	0.994	0.991	0.993	0.993	0.988	0.992	0.991	0.993	0.990	0.993	0.993	0.987	0.992	0.993	0.993	0.988	0.993	0.992	0.985	0.991	0.992
	-0.3	0.963	0.944	0.960	0.951	0.932	0.954	0.950	0.963	0.940	0.956	0.952	0.932	0.954	0.951	0.960	0.934	0.948	0.943	0.923	0.954	0.950
	-0.2	0.833	0.718	0.799	0.780	0.664	0.816	0.792	0.827	0.713	0.789	0.770	0.652	0.813	0.797	0.807	0.690	0.765	0.742	0.618	0.815	0.796
	-0.1	0.392	0.255	0.323	0.297	0.198	0.394	0.346	0.386	0.269	0.340	0.304	0.218	0.414	0.379	0.595	0.504	0.545	0.472	0.638	0.605	0.638
	0	0.076	0.098	0.095	0.081	0.045	0.070	0.046	0.077	0.114	0.105	0.094	0.058	0.087	0.056	0.110	0.151	0.134	0.119	0.092	0.110	0.053
	0.1	0.208	0.417	0.332	0.292	0.188	0.235	0.276	0.392	0.612	0.544	0.494	0.380	0.340	0.382	0.492	0.708	0.645	0.594	0.477	0.420	0.353

Figura 4.31: Poder do Teste considerando $\phi = 0.1$, $\phi = 0.5$ e $\phi = 0.8$ para o processo de Médias Móveis de ordem 1 com tamanho $n = 35$.



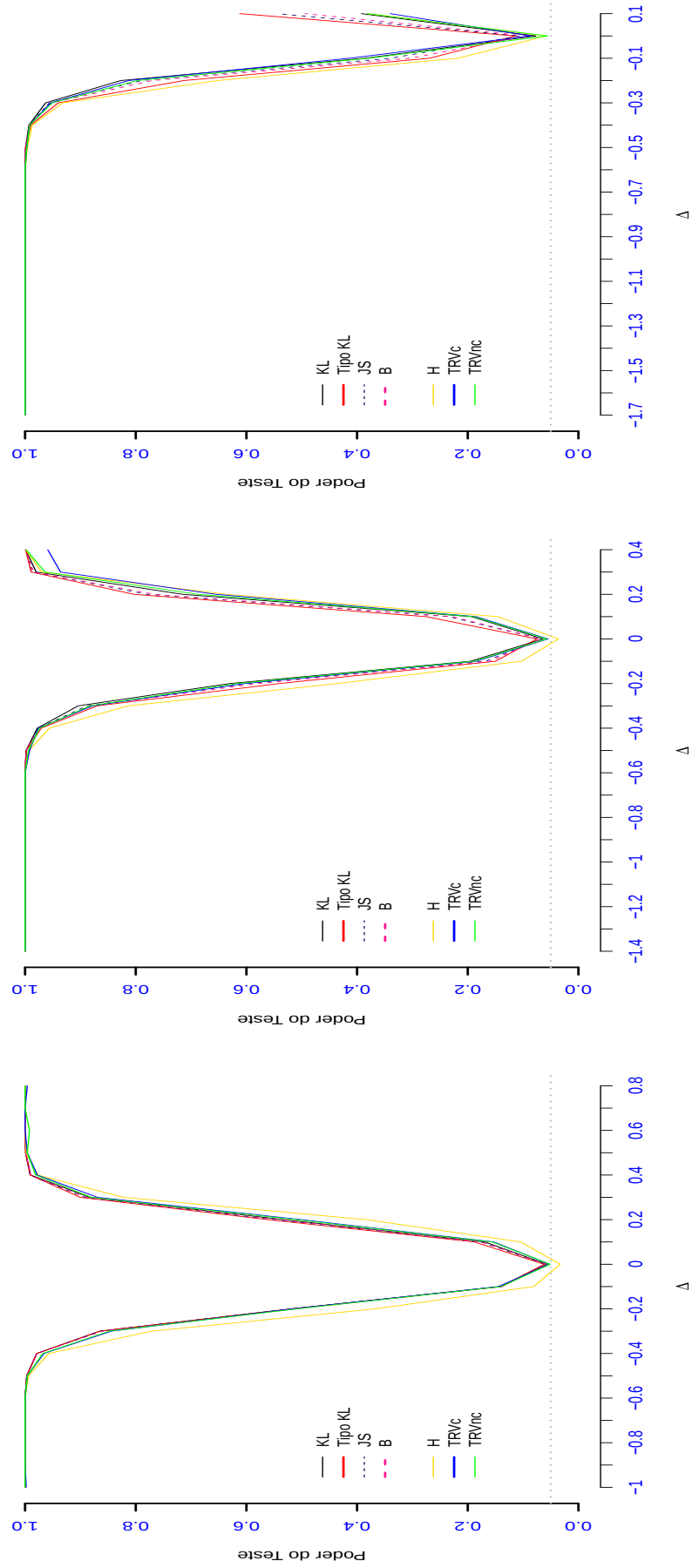
(a) $\phi = 0.1$ (EMVNC)

(b) $\phi = 0.5$ (EMVNC)

(c) $\phi = 0.8$ (EMVNC)

Fonte: Autoria própria.

Figura 4.32: Poder do Teste considerando $\phi = 0.1$, $\phi = 0.5$ e $\phi = 0.8$ para o processo de Médias Móveis de ordem 1 com tamanho $n = 100$.



(a) $\phi = 0.1$ (EMVNC)

(b) $\phi = 0.5$ (EMVNC)

(c) $\phi = 0.8$ (EMVNC)

Fonte: Autoria própria.

4.2.3 Identificação nos modelos ARMA(1,1)

Para o espaço biparamétrico do modelo ARMA(1,1), objetiva-se testar as hipóteses:

$$\mathcal{H}_0 : (\phi, \theta) = (\phi_0, \theta_0) \times \mathcal{H}_1 : (\phi, \theta) = (\phi_0, \theta_0) + (\Delta_1, \Delta_2),$$

em que $(\Delta_1, \Delta_2) \in \{(x_1, x_2) \in \mathbb{R}^2 / \{(0, 0)\} : |\phi| < 1, |\theta| < 1\}$. Especificamente, considerar-se-á os seguintes casos particulares de $\mathcal{H}_1$: Para $\Delta \in \mathbb{R}$ tal que $|\phi| < 1$ e $|\theta| < 1$,

(a) $\phi = \theta$, ou $\mathcal{H}_1 : (\phi, \theta) = (\phi_0, \theta_0) + (\Delta, \Delta)$,

(b) $\phi \neq \theta$ ou $\mathcal{H}_1 : (\phi, \theta) = (\phi_0, \theta_0) + (\Delta, 0)$ e

(c) $\phi \neq \theta$ ou $\mathcal{H}_1 : (\phi, \theta) = (\phi_0, \theta_0) + (0, \Delta)$.

Cenário (a): $\phi = \theta$

Pela Tabela 4.9, os tamanhos empíricos são bem competitivos entre divergências e RV.

Tabela 4.9: Tamanhos empíricos para medidas de divergências e razão de verossimilhanças para $\phi = \theta = 0.1, 0.5, 0.8$

ϕ	N	KL	TKL	JS	B	H	RVC	$RVNC$
0.1	35	0.049	0.131	0.099	0.059	0.011	0.101	0.041
	50	0.073	0.133	0.111	0.079	0.015	0.071	0.041
	150	0.057	0.064	0.059	0.054	0.010	0.050	0.046
0.5	35	0.111	0.152	0.139	0.108	0.011	0.079	0.066
	50	0.076	0.113	0.096	0.079	0.021	0.051	0.055
	150	0.050	0.054	0.052	0.048	0.000	0.050	0.050
0.8	35	0.059	0.089	0.075	0.051	0.003	0.140	0.043
	50	0.055	0.199	0.163	0.049	0.003	0.126	0.049
	150	0.056	0.108	0.082	0.054	0.018	0.050	0.034

Em geral, o poder empírico dos testes cresce a medida que o Δ se afasta de zero e/ou tamanho de amostra n aumenta.

Especificamente, o teste RVNC apresentou taxas mais próximas ao nível nominal especificado de 5%. Considerando o método de estimação MQNC, os testes em geral com exceção do teste RVNC, apresentam tamanho empírico bem acima do nível para tamanhos amostrais $n = 35$ e 50 . Entretanto, com tamanho superior de amostra o tamanho empírico se aproxima do nível. Em relação ao método de estimação MVNC, os testes induzidos das medidas de distância apresentam tamanhos empíricos mais próximo do nível.

Com base na Tabela 4.10, o poder empírico do teste RVC apresentou o melhor desempenho quando o método de estimação MVC é considerado. Entretanto, para os métodos de estimação MVNC e MQNC, o teste RVNC teve poder superior ao teste

RVC. O poder é avaliado assumindo os valores $\phi = \theta = 0.1, 0.6$ e 0.8 . Nas Tabelas 4.10, 4.11 e 4.12, apresentam-se os poderes empíricos dos testes para os valores de $\phi = \theta = 0.1, 0.6$ e 0.8 , respectivamente. Para EMVNC com os tamanhos amostrais $n = 35$ e 100 são apresentados nas Figuras 4.33 e 4.34. As seguintes desigualdades de funções poder empíricas podem ser verificadas:

- (i) Para $\phi = \theta = 0.1$, se (i.1) $\Delta \downarrow -1$ ou $\Delta \uparrow 0.8$ (Estimação MVC) ou se (i.2) $\Delta \downarrow -1$ (Estimação MVNC ou MQNC):

$$\begin{aligned} \pi_{JS:35}(\phi + \Delta, \theta + \Delta) &> \pi_{KL:35}(\phi + \Delta, \theta + \Delta) > \pi_{TKL:35}(\phi + \Delta, \theta + \Delta) > \pi_{B:35}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{RVC:35}(\phi + \Delta, \theta + \Delta) > \pi_{RVNC:35}(\phi + \Delta, \theta + \Delta) > \pi_{H:35}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

$$\begin{aligned} \pi_{JS:50}(\phi + \Delta, \theta + \Delta) &> \pi_{KL:50}(\phi + \Delta, \theta + \Delta) > \pi_{TKL:50}(\phi + \Delta, \theta + \Delta) > \pi_{B:50}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{RVC:50}(\phi + \Delta, \theta + \Delta) > \pi_{RVNC:50}(\phi + \Delta, \theta + \Delta) > \pi_{H:50}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

$$\begin{aligned} \pi_{JS:150}(\phi + \Delta, \theta + \Delta) &> \pi_{KL:150}(\phi + \Delta, \theta + \Delta) > \pi_{TKL:150}(\phi + \Delta, \theta + \Delta) > \pi_{B:150}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{TRVC:150}(\phi + \Delta, \theta + \Delta) > \pi_{TRVNC:150}(\phi + \Delta, \theta + \Delta) > \pi_{H:150}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

- (ii) Para $\phi = \theta = 0.1$, se $\Delta \uparrow 0.8$ (Estimação MVNC ou MQNC):

$$\begin{aligned} \pi_{TKL:35}(\phi + \Delta, \theta + \Delta) &> \pi_{JS:35}(\phi + \Delta, \theta + \Delta) > \pi_{B:35}(\phi + \Delta, \theta + \Delta) > \pi_{KL:35}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{RVC:35}(\phi + \Delta, \theta + \Delta) > \pi_{RVNC:35}(\phi + \Delta, \theta + \Delta) > \pi_{H:35}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

$$\begin{aligned} \pi_{JS:50}(\phi + \Delta, \theta + \Delta) &> \pi_{KL:50}(\phi + \Delta, \theta + \Delta) > \pi_{TKL:50}(\phi + \Delta, \theta + \Delta) > \pi_{B:50}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{RVNC:50}(\phi + \Delta, \theta + \Delta) > \pi_{RVC:50}(\phi + \Delta, \theta + \Delta) > \pi_{H:50}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

$$\begin{aligned} \pi_{JS:150}(\phi + \Delta, \theta + \Delta) &> \pi_{KL:150}(\phi + \Delta, \theta + \Delta) > \pi_{TKL:150}(\phi + \Delta, \theta + \Delta) > \pi_{B:150}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{TRVNC:150}(\phi + \Delta, \theta + \Delta) > \pi_{TRVC:150}(\phi + \Delta, \theta + \Delta) > \pi_{H:150}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

Na Tabela 4.11, assim como no caso anterior, o poder empírico do teste RVC apresentou maior poder quando o método de estimação MVC é considerado. Contudo, o teste RVNC teve poder superior ao teste RVC, com os métodos de estimação MVNC e MQNC. Em relação ao tamanho empírico, o teste RVNC apresentou taxas mais próximas ao nível nominal especificado de 5%. Com tamanho $n = 150$ os testes têm performance aproximada.

Considerando o método de estimação MQNC, os testes induzidos das medidas em geral, apresentam tamanho empírico bem acima do nível principalmente para os testes baseado nas distâncias *Tipo Kullback-Leibler* e *Jensen-Shannon* para tamanhos amostrais $n = 35$ e 50 . Entretanto, com tamanho superior de amostra, o tamanho empírico se aproxima do nível. As seguintes desigualdades de funções poder empíricas podem ser verificadas:

(i) Para $\phi = \theta = 0.5$, se $\Delta \downarrow -1.4$ (Estimação MVC):

$$\pi_{\text{TKL}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{JS}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:35}(\phi + \Delta, \theta + \Delta) \\ > \pi_{\text{RVC}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVNC}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:35}(\phi + \Delta, \theta + \Delta)$$

$$\pi_{\text{TKL}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{JS}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:50}(\phi + \Delta, \theta + \Delta) \\ > \pi_{\text{RVC}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVNC}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:50}(\phi + \Delta, \theta + \Delta)$$

$$\pi_{\text{TKL}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{JS}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:150}(\phi + \Delta, \theta + \Delta) \\ > \pi_{\text{RVC}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVNC}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:150}(\phi + \Delta, \theta + \Delta)$$

(ii) Para $\phi = \theta = 0.5$, se $\Delta \downarrow -1.4$ (Estimação MVNC ou MQNC): A relação de desigualdades acima assumindo para os termos 5º e 6º que $\pi_{\text{RVNC}:n}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVC}:n}(\phi + \Delta, \theta + \Delta)$ para $n = 35, 50, 150$.

(iii) Para $\phi = \theta = 0.5$, se $\Delta \uparrow 0.4$ (Estimação MVC): A mesma relação do item (i) assumindo para 1º, 2º e 3º que $\pi_{\text{JS}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{TKL}:35}(\phi + \Delta, \theta + \Delta)$.

(iv) Quando $\phi = \theta = 0.5$ e $\Delta \uparrow 0.4$ (Estimação MVNC ou MQNC):

$$\pi_{\text{JS}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{TKL}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:35}(\phi + \Delta, \theta + \Delta) \\ > \pi_{\text{RVNC}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVC}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:35}(\phi + \Delta, \theta + \Delta)$$

$$\pi_{\text{JS}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{TKL}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:50}(\phi + \Delta, \theta + \Delta) \\ > \pi_{\text{RVNC}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVC}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:50}(\phi + \Delta, \theta + \Delta)$$

$$\pi_{\text{JS}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{TKL}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:150}(\phi + \Delta, \theta + \Delta) \\ > \pi_{\text{RVNC}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVC}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:150}(\phi + \Delta, \theta + \Delta)$$

De acordo com a Tabela 4.12, podemos extrair os resultados: O teste RVNC teve poder superior ao teste RVC, via estimação MVNC e MQNC. Contudo, quando o método de estimação é MVC, o teste RVC tem maior poder empírico. Em relação ao tamanho empírico, o teste RVNC apresentou taxas mais próximas ao nível nominal especificado de 5%. Além disso, via estimação não condicional, o teste RVC apresentou um tamanho empírico muito elevado para tamanhos de amostras $n = 35$ e 50 . Entretanto, com tamanho amostral $n = 150$, a taxa de rejeição quando $\Delta = 0$ se aproxima do nível de significância; em contra partida, o teste RVNC apresentou tamanho empírico inferior ao nível.

Os testes à luz das medidas de distâncias *Kullback-Leibler* e *Bhattacharyya* obtiveram tamanhos empíricos (para $\Delta = 0$) mais próximo ao nível de significância em relação a outras medidas. As medidas *Tipo Kullback-Leibler* e *Jensen-Shannon* se destacaram para o tamanho amostral $n = 50$.

O poder empírico, das distâncias de *Kullback-Leibler* e *Bhattacharyya* são maiores em relação aos testes RVC e RVNC. Mas, o tamanho (para $\Delta = 0$) apresentado pelos testes induzidos das medidas de distâncias são maiores que o nível especificado. As seguintes desigualdades podem ser verificadas:

- (i) Para $\phi = \theta = 0.8$, se $\Delta \downarrow -1.7$ (Estimação MVC) ou se $\Delta \downarrow -1.7$ (Estimação MVNC ou MQNC):

$$\begin{aligned} \pi_{\text{TKL}:35}(\phi + \Delta, \theta + \Delta) &> \pi_{\text{JS}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:35}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{\text{RVC}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVNC}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:35}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

$$\begin{aligned} \pi_{\text{JS}:50}(\phi + \Delta, \theta + \Delta) &> \pi_{\text{TKL}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:50}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{\text{RVC}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVNC}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:50}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

$$\begin{aligned} \pi_{\text{JS}:150}(\phi + \Delta, \theta + \Delta) &> \pi_{\text{KL}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{TKL}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:150}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{\text{RVC}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVNC}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:150}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

- (ii) Quando $\phi = \theta = 0.8$ e $\Delta = 0.1$ (Estimação MVC, MVNC ou MQNC):

$$\begin{aligned} \pi_{\text{TKL}:35}(\phi + \Delta, \theta + \Delta) &> \pi_{\text{JS}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{KL}:35}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{\text{RVC}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVNC}:35}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:35}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

$$\begin{aligned} \pi_{\text{JS}:50}(\phi + \Delta, \theta + \Delta) &> \pi_{\text{KL}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{TKL}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:50}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{\text{RVNC}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVC}:50}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:50}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

$$\begin{aligned} \pi_{\text{JS}:150}(\phi + \Delta, \theta + \Delta) &> \pi_{\text{KL}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{TKL}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{B}:150}(\phi + \Delta, \theta + \Delta) \\ &> \pi_{\text{RVNC}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{RVC}:150}(\phi + \Delta, \theta + \Delta) > \pi_{\text{H}:150}(\phi + \Delta, \theta + \Delta) \end{aligned}$$

Tabela 4.10: Cenário(a): Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $ARMA(1, 1)$ com $\phi = \theta = 0.1$ e $\sigma^2 = 1$ conhecido

n	Δ	Condicional						Não Condicional						Mínimos Quadrados								
		KL	TKL	JS	B	H	$RVNC$	RVC	KL	TKL	JS	B	H	$RVNC$	RVC	KL	TKL	JS	B	H	$RVNC$	RVC
35	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	0.998	0.997	0.998	0.998	0.984	0.994	0.994	0.999	0.998	1.000	1.000	0.998	0.986	0.996	0.999	1.000	1.000	1.000	1.000	0.992	0.990
	-0.4	0.980	0.981	0.982	0.979	0.893	0.975	0.967	0.987	0.988	0.990	0.985	0.902	0.963	0.974	0.988	0.993	0.993	0.991	0.925	0.951	0.970
	-0.3	0.872	0.862	0.876	0.851	0.597	0.858	0.828	0.901	0.866	0.904	0.860	0.593	0.822	0.846	0.915	0.907	0.925	0.904	0.656	0.823	0.837
	-0.2	0.565	0.603	0.600	0.552	0.222	0.558	0.487	0.594	0.603	0.621	0.557	0.209	0.521	0.521	0.642	0.667	0.678	0.622	0.291	0.558	0.509
	-0.1	0.199	0.289	0.265	0.220	0.057	0.212	0.136	0.205	0.282	0.267	0.207	0.046	0.227	0.152	0.299	0.405	0.375	0.317	0.129	0.280	0.167
50	0	0.072	0.158	0.129	0.091	0.018	0.075	0.049	0.049	0.131	0.099	0.059	0.011	0.101	0.041	0.126	0.226	0.198	0.134	0.051	0.148	0.050
	0.1	0.193	0.294	0.284	0.199	0.040	0.173	0.156	0.192	0.317	0.301	0.216	0.033	0.315	0.148	0.217	0.537	0.462	0.301	0.091	0.506	0.105
	0.2	0.581	0.555	0.584	0.523	0.205	0.545	0.518	0.580	0.579	0.594	0.533	0.197	0.522	0.529	0.634	0.626	0.656	0.603	0.270	0.534	0.523
	0.3	0.910	0.887	0.908	0.887	0.597	0.870	0.869	0.907	0.886	0.909	0.881	0.601	0.855	0.874	0.925	0.918	0.931	0.916	0.668	0.856	0.869
	0.4	0.979	0.978	0.980	0.977	0.905	0.968	0.953	0.940	0.977	0.971	0.959	0.845	0.932	0.871	0.903	0.983	0.971	0.940	0.825	0.886	0.777
	0.5	0.999	0.999	0.999	0.999	0.990	0.999	1.000	1.000	1.000	1.000	1.000	0.994	0.996	1.000	1.000	1.000	1.000	1.000	0.994	0.987	1.000
	0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	0.991	1.000
	0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	1.000
	0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
150	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.2	0.978	0.980	0.980	0.978	0.928	0.988	0.984	0.990	0.990	0.990	0.988	0.962	0.978	0.988	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.1	0.570	0.578	0.574	0.554	0.192	0.626	0.570	0.624	0.644	0.636	0.614	0.128	0.588	0.618	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0	0.063	0.065	0.065	0.058	0.009	0.054	0.045	0.057	0.064	0.059	0.054	0.010	0.050	0.046	0.063	0.067	0.065	0.056	0.012	0.055	0.045
	0.1	0.604	0.612	0.616	0.598	0.254	0.580	0.580	0.606	0.606	0.606	0.600	0.256	0.578	0.586	0.616	0.622	0.626	0.608	0.268	0.576	0.586
	0.2	0.992	0.996	0.994	0.994	0.954	0.994	0.994	0.992	0.994	0.994	0.994	0.956	0.994	0.994	0.994	0.994	0.994	0.994	0.958	0.994	0.994
	0.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.2	0.978	0.980	0.980	0.978	0.928	0.988	0.984	0.990	0.990	0.990	0.988	0.962	0.978	0.988	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.1	0.570	0.578	0.574	0.554	0.192	0.626	0.570	0.624	0.644	0.636	0.614	0.128	0.588	0.618	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0	0.063	0.065	0.065	0.058	0.009	0.054	0.045	0.057	0.064	0.059	0.054	0.010	0.050	0.046	0.063	0.067	0.065	0.056	0.012	0.055	0.045
	0.1	0.604	0.612	0.616	0.598	0.254	0.580	0.580	0.606	0.606	0.606	0.600	0.256	0.578	0.586	0.616	0.622	0.626	0.608	0.268	0.576	0.586
	0.2	0.992	0.996	0.994	0.994	0.954	0.994															

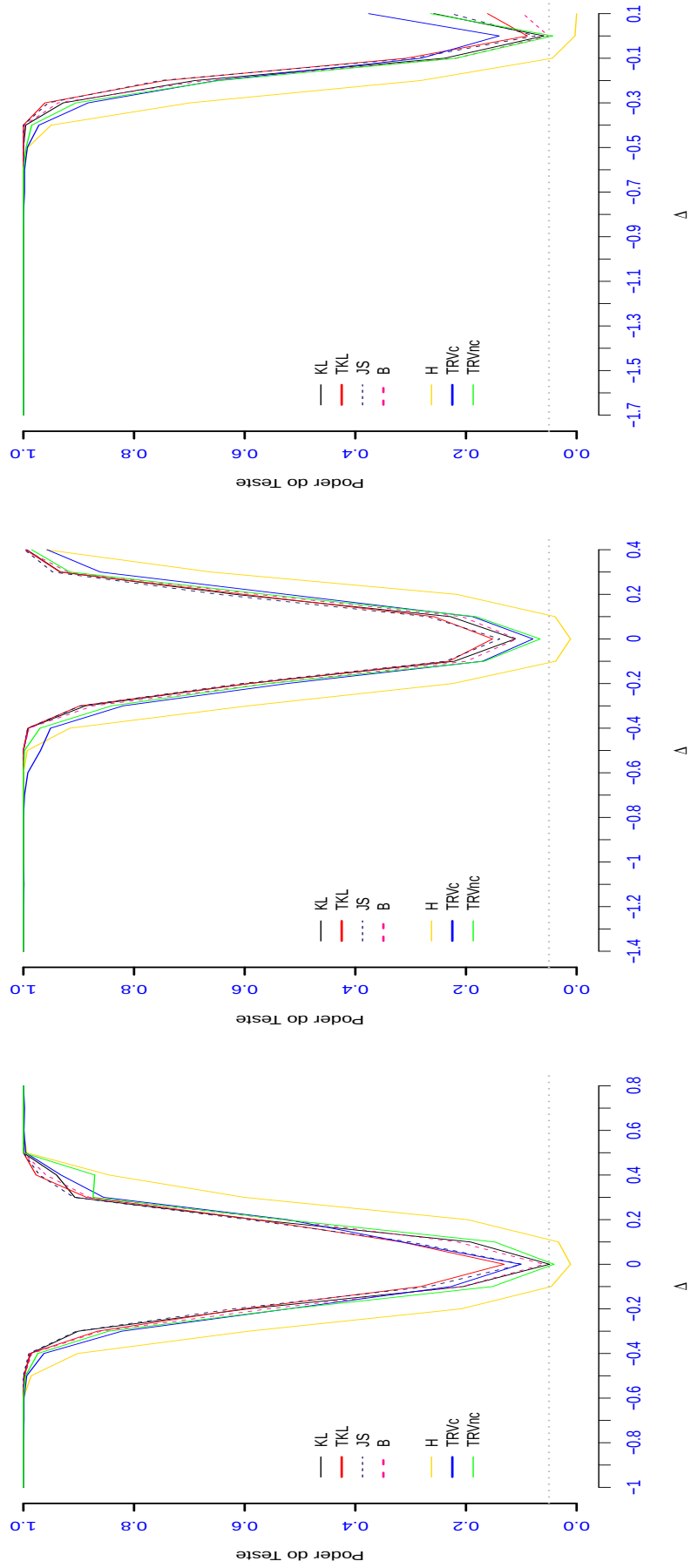
Tabela 4.11: Cenário(a): Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $ARMA(1, 1)$ com $\phi = \theta = 0.5$ e $\sigma^2 = 1$ conhecido

n	Δ	Condicional					Não Condicional					Mínimos Quadrados					RVC	RVNC				
		KL	TKL	JS	B	H	RVC	RVNC	KL	TKL	JS	B	H	RVC	RVNC	KL			TKL	JS	B	H
35	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	-0.5	0.990	0.991	0.990	0.987	0.910	0.968	0.954	0.991	0.992	0.992	0.990	0.916	0.951	0.970	0.988	0.990	0.990	0.987	0.907	0.947	0.962
	-0.4	0.879	0.887	0.885	0.866	0.587	0.836	0.843	0.888	0.898	0.897	0.874	0.594	0.873	0.841	0.874	0.882	0.870	0.857	0.740	0.821	0.838
	-0.3	0.590	0.585	0.586	0.551	0.223	0.543	0.532	0.592	0.594	0.594	0.606	0.558	0.524	0.557	0.577	0.574	0.578	0.536	0.240	0.528	0.551
	-0.2	0.218	0.223	0.223	0.192	0.037	0.169	0.159	0.221	0.231	0.237	0.199	0.038	0.170	0.167	0.232	0.220	0.236	0.214	0.048	0.192	0.162
	0	0.131	0.147	0.142	0.111	0.021	0.073	0.056	0.111	0.152	0.139	0.107	0.011	0.079	0.066	0.197	0.222	0.211	0.183	0.031	0.127	0.063
	0.1	0.243	0.251	0.268	0.202	0.036	0.188	0.152	0.228	0.258	0.269	0.207	0.039	0.188	0.182	0.383	0.376	0.398	0.337	0.092	0.239	0.164
	0.2	0.605	0.605	0.613	0.544	0.201	0.558	0.536	0.618	0.602	0.648	0.577	0.217	0.525	0.562	0.742	0.728	0.756	0.756	0.316	0.858	0.548
	0.3	0.913	0.907	0.922	0.900	0.621	0.905	0.903	0.933	0.934	0.947	0.920	0.661	0.861	0.916	0.969	0.967	0.974	0.961	0.812	0.859	0.909
	0.4	0.989	0.985	0.990	0.989	0.930	0.983	0.980	0.994	0.996	0.998	0.997	0.959	0.957	0.985	0.998	0.999	0.999	0.998	0.983	0.922	0.985
150	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.2	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.996	0.996	0.994	0.994	0.960	0.994	0.994	0.994	0.992	0.996	0.992	0.958	0.994	0.994
	-0.1	0.590	0.606	0.608	0.598	0.224	0.594	0.590	0.612	0.622	0.616	0.602	0.236	0.590	0.590	0.574	0.600	0.594	0.574	0.210	0.590	0.590
	0	0.054	0.058	0.056	0.044	0.002	0.052	0.048	0.050	0.054	0.052	0.048	0.000	0.050	0.050	0.060	0.060	0.060	0.054	0.002	0.050	0.050
	0.1	0.694	0.624	0.630	0.620	0.274	0.618	0.622	0.630	0.642	0.646	0.622	0.272	0.614	0.626	0.668	0.680	0.682	0.670	0.316	0.612	0.624
	0.2	0.994	0.996	0.994	0.994	0.994	0.994	0.992	0.994	0.992	0.994	0.994	0.978	0.994	0.994	0.994	0.996	0.996	0.996	0.996	0.994	0.992
	0.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

Tabela 4.12: Cenário(a): Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $ARMA(1, 1)$ com $\phi = \theta = 0.8$ e $\sigma^2 = 1$ conhecido

n	Δ	Condicional						Não Condicional						Mínimos Quadrados						RVC	$RVNC$				
		KL	TKL	JS	B	H	$RVNC$	RVC	$RVNC$	KL	TKL	JS	B	H	$RVNC$	RVC	$RVNC$	KL	TKL			JS	B	H	
35	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	0.999	1.000	1.000	1.000	1.000	1.000	0.991	0.995	1.000	0.991	0.995	1.000	0.992	0.993	0.995	0.995	0.998	1.000	0.989	0.986	0.994	0.982	0.982	0.982
	-0.4	0.996	1.000	1.000	1.000	0.997	0.939	0.978	0.985	0.996	1.000	1.000	0.998	0.950	0.998	0.972	0.985	0.995	1.000	0.997	0.924	0.971	0.982	0.982	0.982
	-0.3	0.927	0.955	0.950	0.928	0.681	0.898	0.892	0.928	0.961	0.955	0.955	0.937	0.699	0.883	0.905	0.965	0.911	0.948	0.944	0.916	0.647	0.894	0.886	0.886
	-0.2	0.700	0.730	0.744	0.669	0.287	0.676	0.624	0.690	0.745	0.753	0.673	0.282	0.605	0.234	0.633	0.631	0.692	0.686	0.602	0.634	0.696	0.633	0.633	0.633
	-0.1	0.257	0.299	0.293	0.223	0.045	0.263	0.219	0.239	0.308	0.295	0.218	0.044	0.281	0.219	0.210	0.276	0.210	0.276	0.242	0.192	0.034	0.388	0.206	0.206
	0	0.093	0.101	0.109	0.067	0.009	0.077	0.030	0.059	0.089	0.075	0.051	0.003	0.043	0.043	0.095	0.115	0.095	0.115	0.110	0.061	0.005	0.320	0.035	0.035
100	0.1	0.289	0.142	0.253	0.141	0.020	0.267	0.204	0.258	0.161	0.227	0.097	0.000	0.376	0.264	0.439	0.422	0.431	0.294	0.063	0.544	0.226	0.544	0.226	0.226
	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.2	0.998	1.000	1.000	1.000	1.000	0.986	0.998	0.998	0.998	1.000	1.000	0.988	0.988	0.998	0.998	0.998	0.998	1.000	0.998	0.978	0.998	0.998	0.998	0.998
	-0.1	0.774	0.768	0.794	0.774	0.440	0.768	0.754	0.758	0.780	0.782	0.760	0.432	0.760	0.432	0.762	0.756	0.722	0.740	0.748	0.734	0.734	0.734	0.734	0.734
0	0.062	0.084	0.070	0.056	0.008	0.046	0.026	0.056	0.108	0.082	0.054	0.018	0.050	0.034	0.074	0.122	0.074	0.122	0.102	0.074	0.028	0.048	0.032	0.032	
100	0.1	0.810	0.766	0.810	0.768	0.478	0.812	0.830	0.872	0.894	0.878	0.578	0.754	0.864	0.864	0.914	0.918	0.914	0.918	0.926	0.898	0.862	0.862	0.862	0.862

Figura 4.33: Poder do Teste considerando $\phi = 0.1$, $\phi = 0.5$ e $\phi = 0.8$ para o processo Autoregressivo de Médias Móveis de ordem 1 com tamanho $n = 35$.



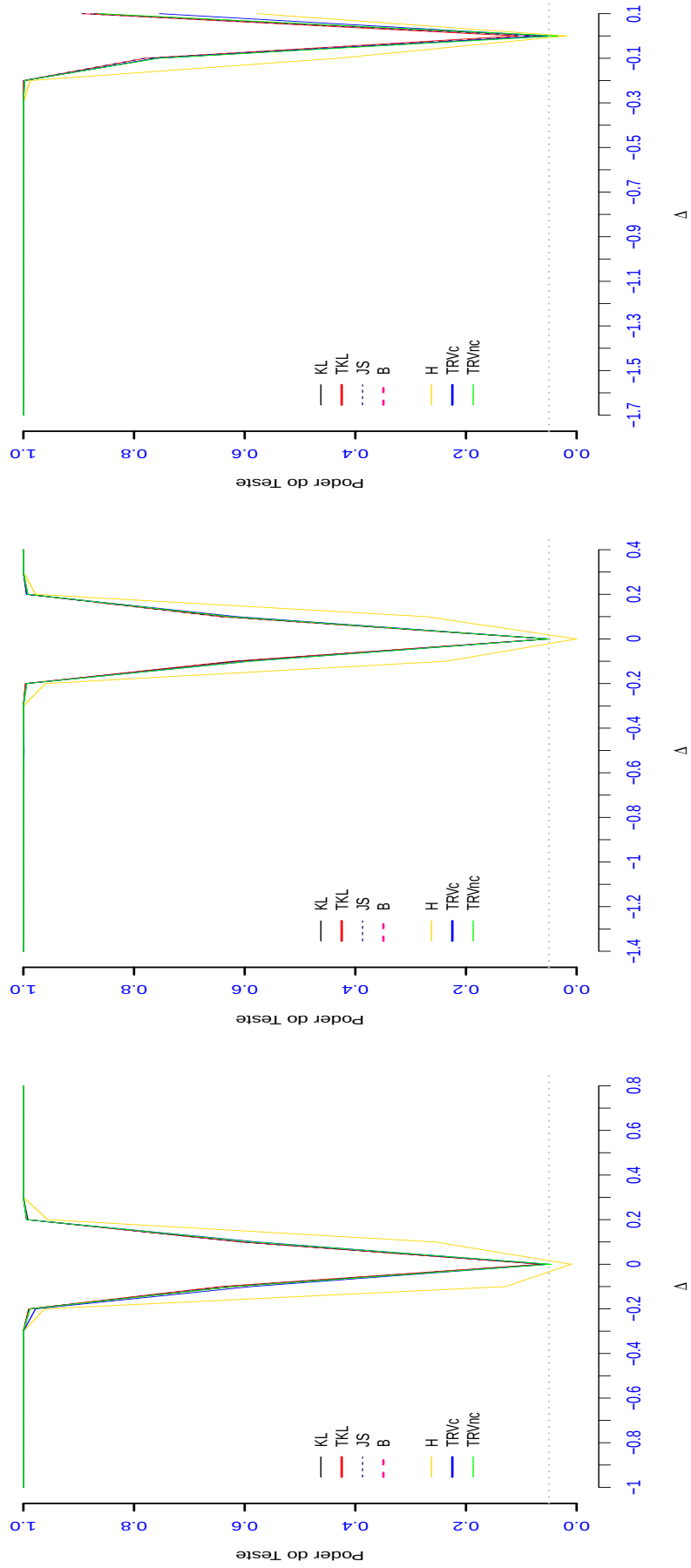
(c) $\phi = 0.8$ (EMVNC)

(b) $\phi = 0.5$ (EMVNC)

(a) $\phi = 0.1$ (EMVNC)

Fonte: Autoria própria.

Figura 4.34: Poder do Teste considerando $\phi = 0.1$, $\phi = 0.5$ e $\phi = 0.8$ para o processo Autoregressivo de Médias Móveis de ordem 1 com tamanho $n = 150$.



(a) $\phi = 0.1$ (EMVNC)

(b) $\phi = 0.5$ (EMVNC)

(c) $\phi = 0.8$ (EMVNC)

Fonte: Autoria própria.

Cenário (b): $\phi \neq \theta$ ($\phi + \Delta, \theta$)

Neste cenário avaliamos o poder empírico dos testes assumindo os valores verdadeiros dos parâmetros $\phi = 0.1 + \Delta, 0.5 + \Delta, 0.8 + \Delta$ e $\theta = 0.1, 0.6$ e 0.8 , apresentados respectivamente, nas Tabelas 4.13, 4.14 e 4.15.

No geral relacionado aos testes de razão entre verossimilhanças, via estimação MVC, o poder empírico do teste RVC é maior. Entretanto, segundo os métodos MVNC e MQNC, o teste RVNC teve poder superior ao teste RVC. Para tamanhos grandes de amostras os testes são aproximados.

Com base na Tabela 4.13, segue os resultados:

Considerando os métodos de estimação MVNC e MQNC, as distâncias *Kullback-Leibler* e *Bhattacharyya* apresenta desempenho superior do poder empírico dos testes, comparado com o método MVC.

As seguintes desigualdades podem ser verificadas:

(i) Quando $\phi = 0.1 + \Delta, \theta = 0.1$ e $\Delta \downarrow -1$ (Estimação MVC):

$$\begin{aligned} \pi_{\text{RVC};35;50}(\phi + \Delta, \theta) &> \pi_{\text{KL};35;50}(\phi + \Delta, \theta) > \pi_{\text{JS};35;50}(\phi + \Delta, \theta) > \pi_{\text{TKL};35;50}(\phi + \Delta, \theta) \\ &> \pi_{\text{RVNC};35;50}(\phi + \Delta, \theta) > \pi_{\text{B};35;50}(\phi + \Delta, \theta) > \pi_{\text{H};35;50}(\phi + \Delta, \theta) \\ \pi_{\text{RVC};150}(\phi + \Delta, \theta) &> \pi_{\text{KL};150}(\phi + \Delta, \theta) \approx \pi_{\text{JS};150}(\phi + \Delta, \theta) \approx \pi_{\text{TKL};150}(\phi + \Delta, \theta) \\ &\approx \pi_{\text{RVNC};150}(\phi + \Delta, \theta) \approx \pi_{\text{B};150}(\phi + \Delta, \theta) > \pi_{\text{H};150}(\phi + \Delta, \theta) \end{aligned}$$

(ii) Quando $\phi = 0.1 + \Delta, \theta = 0.1$ e $\Delta \downarrow -1$ (Estimação MVNC ou MQNC):

$$\begin{aligned} \pi_{\text{KL};35}(\phi + \Delta, \theta) &> \pi_{\text{JS};35}(\phi + \Delta, \theta) > \pi_{\text{RVNC};35}(\phi + \Delta, \theta) > \pi_{\text{B};35}(\phi + \Delta, \theta) \\ &> \pi_{\text{TKL};35}(\phi + \Delta, \theta) > \pi_{\text{RVC};35}(\phi + \Delta, \theta) > \pi_{\text{H};35}(\phi + \Delta, \theta) \\ \pi_{\text{KL};50}(\phi + \Delta, \theta) &> \pi_{\text{JS};50}(\phi + \Delta, \theta) > \pi_{\text{B};50}(\phi + \Delta, \theta) > \pi_{\text{RVNC};50}(\phi + \Delta, \theta) \\ &\approx \pi_{\text{RVC};50}(\phi + \Delta, \theta) \approx \pi_{\text{TKL};50}(\phi + \Delta, \theta) > \pi_{\text{H};50}(\phi + \Delta, \theta) \\ \pi_{\text{RVNC};150}(\phi + \Delta, \theta) &> \pi_{\text{KL};150}(\phi + \Delta, \theta) \approx \pi_{\text{TKL};150}(\phi + \Delta, \theta) \approx \pi_{\text{B};150}(\phi + \Delta, \theta) \\ &\approx \pi_{\text{RVC};150}(\phi + \Delta, \theta) \approx \pi_{\text{JS};150}(\phi + \Delta, \theta) > \pi_{\text{H};150}(\phi + \Delta, \theta) \end{aligned}$$

(iii) Quando $\phi = 0.1 + \Delta, \theta = 0.1$ e $\Delta \uparrow 0.8$ (Estimação MVC):

$$\begin{aligned} \pi_{\text{KL};35;50}(\phi + \Delta, \theta) &> \pi_{\text{RVC};35;50}(\phi + \Delta, \theta) > \pi_{\text{RVNC};35;50}(\phi + \Delta, \theta) > \pi_{\text{JS};35;50}(\phi + \Delta, \theta) \\ &> \pi_{\text{B};35;50}(\phi + \Delta, \theta) > \pi_{\text{TKL};35;50}(\phi + \Delta, \theta) > \pi_{\text{H};35;50}(\phi + \Delta, \theta) \\ \pi_{\text{KL};150}(\phi + \Delta, \theta) &> \pi_{\text{RVC};150}(\phi + \Delta, \theta) \approx \pi_{\text{RVNC};150}(\phi + \Delta, \theta) \approx \pi_{\text{JS};150}(\phi + \Delta, \theta) \\ &\approx \pi_{\text{B};150}(\phi + \Delta, \theta) \approx \pi_{\text{TKL};150}(\phi + \Delta, \theta) > \pi_{\text{H};150}(\phi + \Delta, \theta) \end{aligned}$$

(iv) Quando $\phi = 0.1 + \Delta$, $\theta = 0.1$ e $\Delta \uparrow 0.8$ (Estimação MVNC ou MQNC):

$$\begin{aligned} \pi_{\text{KL}:35}(\phi + \Delta, \theta) &> \pi_{\text{JS}:35}(\phi + \Delta, \theta) > \pi_{\text{RVNC}:35}(\phi + \Delta, \theta) > \pi_{\text{RVC}:35}(\phi + \Delta, \theta) \\ &> \pi_{\text{B}:35}(\phi + \Delta, \theta) > \pi_{\text{TKL}:35}(\phi + \Delta, \theta) > \pi_{\text{H}:35}(\phi + \Delta, \theta) \\ \pi_{\text{KL}:50}(\phi + \Delta, \theta) &> \pi_{\text{JS}:50}(\phi + \Delta, \theta) > \pi_{\text{B}:50}(\phi + \Delta, \theta) > \pi_{\text{TKL}:50}(\phi + \Delta, \theta) \\ &> \pi_{\text{RVC}:50}(\phi + \Delta, \theta) > \pi_{\text{RVNC}:50}(\phi + \Delta, \theta) > \pi_{\text{H}:50}(\phi + \Delta, \theta) \\ \pi_{\text{KL}:150}(\phi + \Delta, \theta) &> \pi_{\text{JS}:150}(\phi + \Delta, \theta) \approx \pi_{\text{B}:150}(\phi + \Delta, \theta) \approx \pi_{\text{RVNC}:150}(\phi + \Delta, \theta) \\ &\approx \pi_{\text{RVC}:150}(\phi + \Delta, \theta) \approx \pi_{\text{TKL}:150}(\phi + \Delta, \theta) > \pi_{\text{H}:150}(\phi + \Delta, \theta) \end{aligned}$$

De acordo com a Tabela 4.14, quando $\Delta \downarrow -1.4$, o poder do teste baseado nas distâncias *Tipo Kullback-Leibler* e *Jensen-Shannon* tiveram maior poder (onde $\pi_{\text{TKL}}(\phi + \Delta, \theta) > \pi_{\text{JS}}(\phi + \Delta, \theta)$) e quando $\Delta \uparrow 0.4$, os testes *Kullback-Leibler* e *Jensen-Shannon* tem melhores performances (onde $\pi_{\text{KL}}(\phi + \Delta, \theta) > \pi_{\text{JS}}(\phi + \Delta, \theta)$).

As seguintes desigualdades de funções poder empíricas podem ser verificadas:

(i) Quando $\phi = 0.5 + \Delta$, $\theta = 0.5$ e $\Delta \downarrow -1.4$ (Estimação MVC):

$$\begin{aligned} \pi_{\text{TKL}:35;50}(\phi + \Delta, \theta) &> \pi_{\text{JS}:35;50}(\phi + \Delta, \theta) > \pi_{\text{B}:35;50}(\phi + \Delta, \theta) > \pi_{\text{KL}:35;50}(\phi + \Delta, \theta) \\ &> \pi_{\text{RVC}:35;50}(\phi + \Delta, \theta) > \pi_{\text{RVNC}:35;50}(\phi + \Delta, \theta) > \pi_{\text{H}:35;50}(\phi + \Delta, \theta) \\ \pi_{\text{TKL}:150}(\phi + \Delta, \theta) &> \pi_{\text{JS}:150}(\phi + \Delta, \theta) \approx \pi_{\text{B}:150}(\phi + \Delta, \theta) \approx \pi_{\text{KL}:150}(\phi + \Delta, \theta) \\ &\approx \pi_{\text{RVC}:150}(\phi + \Delta, \theta) \approx \pi_{\text{RVNC}:150}(\phi + \Delta, \theta) > \pi_{\text{H}:150}(\phi + \Delta, \theta) \end{aligned}$$

(ii) Quando $\phi = 0.5 + \Delta$, $\theta = 0.5$ e $\Delta \downarrow -1.4$ (Estimação MVNC ou MQNC): A relação de desigualdades acima assumindo para os termos 5º e 6º que $\pi_{\text{RVNC}:n}(\phi + \Delta, \theta) > \pi_{\text{RVC}:n}(\phi + \Delta, \theta)$ para $n = 35, 50, 150$.

(iii) Quando $\phi = 0.5 + \Delta$, $\theta = 0.5$ e $\Delta \uparrow 0.4$ (Estimação MVC):

$$\begin{aligned} \pi_{\text{KL}:35;50}(\phi + \Delta, \theta) &> \pi_{\text{JS}:35;50}(\phi + \Delta, \theta) > \pi_{\text{RVC}:35;50}(\phi + \Delta, \theta) \approx \pi_{\text{RVNC}:35;50}(\phi + \Delta, \theta) \\ &> \pi_{\text{B}:35;50}(\phi + \Delta, \theta) > \pi_{\text{TKL}:35;50}(\phi + \Delta, \theta) > \pi_{\text{H}:35;50}(\phi + \Delta, \theta) \\ \pi_{\text{KL}:150}(\phi + \Delta, \theta) &> \pi_{\text{RVC}:150}(\phi + \Delta, \theta) \approx \pi_{\text{RVNC}:150}(\phi + \Delta, \theta) \approx \pi_{\text{JS}:150}(\phi + \Delta, \theta) \\ &\approx \pi_{\text{B}:150}(\phi + \Delta, \theta) \approx \pi_{\text{TKL}:150}(\phi + \Delta, \theta) > \pi_{\text{H}:150}(\phi + \Delta, \theta) \end{aligned}$$

- (iv) Quando $\phi = 0.5 + \Delta$, $\theta = 0.5$ e $\Delta \uparrow 0.4$ (Estimação MVNC ou MQNC): A relação de desigualdades acima assumindo para os termos 3º e 4º que $\pi_{\text{RVNC}:n}(\phi + \Delta, \theta) > \pi_{\text{RVC}:n}(\phi + \Delta, \theta)$ para $n = 35, 50, 150$.

Analisando a Tabela 4.15, pode-se chegar as seguintes conclusões:

Quando $\Delta \downarrow -1.4$, o poder do teste baseado nas distâncias *Tipo Kullback-Leibler* e *Jensen-Shannon* tiveram maior poder (onde $\pi_{\text{TKL}}(\phi + \Delta, \theta) > \pi_{\text{JS}}(\phi + \Delta, \theta)$) e quando $\Delta \uparrow 0.4$, os testes *Kullback-Leibler* e *Jensen-Shannon* tem melhores performances (onde $\pi_{\text{KL}}(\phi + \Delta, \theta) > \pi_{\text{JS}}(\phi + \Delta, \theta)$).

As seguintes desigualdades de funções poder empíricas podem ser verificadas:

- (i) Quando $\phi = 0.8 + \Delta$, $\theta = 0.8$ e $\Delta \downarrow -1.7$ (Estimação MVC):

$$\begin{aligned} \pi_{\text{TKL}:35;50}(\phi + \Delta, \theta) &> \pi_{\text{JS}:35;50}(\phi + \Delta, \theta) > \pi_{\text{B}:35;50}(\phi + \Delta, \theta) > \pi_{\text{KL}:35;50}(\phi + \Delta, \theta) \\ &> \pi_{\text{RVC}:35;50}(\phi + \Delta, \theta) > \pi_{\text{RVNC}:35;50}(\phi + \Delta, \theta) > \pi_{\text{H}:35;50}(\phi + \Delta, \theta) \\ \pi_{\text{TKL}:150}(\phi + \Delta, \theta) &> \pi_{\text{JS}:150}(\phi + \Delta, \theta) \approx \pi_{\text{B}:150}(\phi + \Delta, \theta) \approx \pi_{\text{KL}:150}(\phi + \Delta, \theta) \\ &\approx \pi_{\text{RVC}:150}(\phi + \Delta, \theta) \approx \pi_{\text{RVNC}:150}(\phi + \Delta, \theta) > \pi_{\text{H}:150}(\phi + \Delta, \theta) \end{aligned}$$

- (ii) Quando $\phi = 0.8 + \Delta$, $\theta = 0.8$ e $\Delta \downarrow -1.7$ (Estimação MVNC ou MQNC): A relação de desigualdades acima assumindo para os termos 5º e 6º que $\pi_{\text{RVNC}:n}(\phi + \Delta, \theta) > \pi_{\text{RVC}:n}(\phi + \Delta, \theta)$ para $n = 35, 50, 150$.

- (iii) Quando $\phi = 0.8 + \Delta$, $\theta = 0.8$ e $\Delta = 0.1$ (Estimação MVC):

$$\begin{aligned} \pi_{\text{KL}:35;50}(\phi + \Delta, \theta) &> \pi_{\text{RVC}:35;50}(\phi + \Delta, \theta) > \pi_{\text{JS}:35;50}(\phi + \Delta, \theta) > \pi_{\text{RVNC}:35;50}(\phi + \Delta, \theta) \\ &> \pi_{\text{B}:35;50}(\phi + \Delta, \theta) > \pi_{\text{TKL}:35;50}(\phi + \Delta, \theta) > \pi_{\text{H}:35;50}(\phi + \Delta, \theta) \\ \pi_{\text{KL}:150}(\phi + \Delta, \theta) &> \pi_{\text{RVC}:150}(\phi + \Delta, \theta) > \pi_{\text{RVNC}:150}(\phi + \Delta, \theta) > \pi_{\text{JS}:150}(\phi + \Delta, \theta) \\ &> \pi_{\text{B}:150}(\phi + \Delta, \theta) > \pi_{\text{TKL}:150}(\phi + \Delta, \theta) > \pi_{\text{H}:150}(\phi + \Delta, \theta) \end{aligned}$$

- (iv) Quando $\phi = 0.8 + \Delta$, $\theta = 0.8$ e $\Delta = 0.1$ (Estimação MVNC ou MQNC):

$$\begin{aligned} \pi_{\text{RVC}:35;50}(\phi + \Delta, \theta) &> \pi_{\text{KL}:35;50}(\phi + \Delta, \theta) > \pi_{\text{RVNC}:35;50}(\phi + \Delta, \theta) > \pi_{\text{JS}:35;50}(\phi + \Delta, \theta) \\ &> \pi_{\text{TKL}:35;50}(\phi + \Delta, \theta) > \pi_{\text{B}:35;50}(\phi + \Delta, \theta) > \pi_{\text{H}:35;50}(\phi + \Delta, \theta) \\ \pi_{\text{KL}:150}(\phi + \Delta, \theta) &> \pi_{\text{RVNC}:150}(\phi + \Delta, \theta) > \pi_{\text{RVC}:150}(\phi + \Delta, \theta) > \pi_{\text{JS}:150}(\phi + \Delta, \theta) \\ &> \pi_{\text{B}:150}(\phi + \Delta, \theta) > \pi_{\text{TKL}:150}(\phi + \Delta, \theta) > \pi_{\text{H}:150}(\phi + \Delta, \theta) \end{aligned}$$

Tabela 4.14: Cenário (b): Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $ARMA(1, 1)$ com $\phi = 0.5 + \Delta$, $\theta = 0.5$ e $\sigma^2 = 1$ conhecido

η	Δ	Condicional						Não Condicional						Mínimos Quadrados						RVC	$RVNC$		
		KL	TKL	JS	B	H	$RVNC$	KL	TKL	JS	B	H	RVC	$RVNC$	KL	TKL	JS	B	H				
35	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000	
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000	
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.984	0.999	
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.986	1.000	
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.987	0.999	
	-0.9	0.995	0.995	0.995	0.994	0.986	0.988	0.980	0.996	1.000	0.997	0.996	0.996	0.964	0.986	0.996	0.997	0.997	0.995	0.995	0.968	0.954	0.977
	-0.8	0.991	0.991	0.992	0.989	0.891	0.978	0.970	0.991	0.995	0.994	0.993	0.905	0.947	0.979	0.988	0.994	0.995	0.991	0.942	0.975	0.901	0.942
	-0.7	0.950	0.968	0.961	0.961	0.949	0.951	0.931	0.955	0.976	0.968	0.959	0.773	0.899	0.940	0.951	0.967	0.962	0.954	0.773	0.881	0.938	0.938
	-0.6	0.887	0.919	0.910	0.884	0.595	0.844	0.849	0.901	0.928	0.920	0.908	0.615	0.806	0.857	0.752	0.919	0.919	0.900	0.619	0.816	0.853	0.853
	-0.5	0.759	0.796	0.780	0.749	0.376	0.701	0.697	0.761	0.816	0.794	0.763	0.382	0.668	0.720	0.752	0.795	0.784	0.749	0.403	0.672	0.712	0.712
50	-0.4	0.559	0.653	0.609	0.568	0.192	0.523	0.508	0.577	0.669	0.633	0.589	0.220	0.484	0.530	0.581	0.653	0.627	0.580	0.502	0.502	0.521	
	-0.3	0.378	0.444	0.416	0.344	0.081	0.314	0.310	0.391	0.469	0.432	0.395	0.104	0.315	0.341	0.423	0.476	0.452	0.420	0.319	0.351	0.329	
	-0.2	0.213	0.260	0.232	0.204	0.045	0.171	0.150	0.216	0.267	0.244	0.216	0.037	0.158	0.169	0.243	0.283	0.267	0.235	0.056	0.178	0.159	
	-0.1	0.131	0.171	0.156	0.131	0.021	0.092	0.073	0.130	0.180	0.165	0.139	0.023	0.090	0.090	0.186	0.225	0.206	0.189	0.030	0.125	0.085	
	0.1	0.169	0.169	0.171	0.134	0.018	0.110	0.081	0.138	0.156	0.152	0.117	0.014	0.107	0.086	0.262	0.247	0.263	0.214	0.041	0.145	0.082	
	0.2	0.342	0.190	0.297	0.214	0.031	0.299	0.275	0.324	0.196	0.281	0.217	0.035	0.298	0.285	0.442	0.307	0.413	0.330	0.073	0.315	0.278	
	0.3	0.638	0.409	0.581	0.492	0.166	0.589	0.569	0.630	0.407	0.573	0.491	0.145	0.576	0.590	0.734	0.551	0.690	0.621	0.277</			

Tabela 4.15: Cenário (b): Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $ARMA(1,1)$ com $\phi = 0.8 + \Delta$, $\theta = 0.8$ e $\sigma^2 = 1$ conhecido

n	Δ	Condicional						Não Condicional						Mínimos Quadrados						
		KL	TKL	JS	H	RVC	RVNC	KL	TKL	JS	H	RVC	RVNC	KL	TKL	JS	B	H	RVC	RVNC
35	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.988	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.992	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.974	1.000
	-1.2	0.999	1.000	1.000	1.000	0.999	0.998	1.000	1.000	1.000	1.000	0.999	0.999	1.000	1.000	1.000	1.000	0.999	0.972	0.998
	-1.1	1.000	1.000	1.000	1.000	0.999	0.997	1.000	1.000	1.000	1.000	0.998	0.987	1.000	1.000	1.000	1.000	0.999	0.968	0.998
	-1.0	1.000	1.000	1.000	1.000	1.000	0.997	1.000	1.000	1.000	1.000	1.000	0.980	1.000	1.000	1.000	1.000	1.000	0.958	0.997
	-0.9	1.000	1.000	1.000	1.000	0.989	0.994	0.999	1.000	1.000	1.000	0.995	0.971	0.996	1.000	1.000	1.000	0.992	0.937	0.995
	-0.8	0.995	1.000	1.000	0.997	0.963	0.983	0.998	1.000	1.000	0.999	0.976	0.944	0.989	0.997	1.000	0.999	0.975	0.923	0.984
50	-0.7	0.987	0.999	0.997	0.993	0.889	0.956	0.998	1.000	1.000	0.999	0.916	0.913	0.971	0.989	1.000	0.997	0.993	0.880	0.957
	-0.6	0.953	0.986	0.979	0.970	0.777	0.913	0.910	0.966	0.993	0.987	0.979	0.812	0.856	0.958	0.990	0.981	0.968	0.791	0.851
	-0.5	0.859	0.958	0.926	0.895	0.607	0.805	0.805	0.880	0.973	0.952	0.925	0.753	0.829	0.870	0.957	0.936	0.907	0.589	0.751
	-0.4	0.698	0.881	0.820	0.760	0.343	0.652	0.612	0.707	0.901	0.850	0.784	0.378	0.585	0.675	0.870	0.824	0.754	0.335	0.641
	-0.3	0.472	0.687	0.606	0.524	0.172	0.446	0.365	0.479	0.728	0.650	0.553	0.174	0.413	0.453	0.708	0.618	0.526	0.167	0.419
	-0.2	0.268	0.418	0.349	0.279	0.069	0.255	0.190	0.249	0.462	0.385	0.288	0.068	0.297	0.226	0.430	0.344	0.264	0.051	0.447
	-0.1	0.123	0.222	0.182	0.128	0.024	0.121	0.067	0.114	0.220	0.166	0.124	0.019	0.182	0.107	0.197	0.162	0.114	0.020	0.342
	0.1	0.321	0.101	0.253	0.124	0.020	0.258	0.192	0.241	0.074	0.188	0.053	0.002	0.274	0.449	0.293	0.423	0.272	0.060	0.398
	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	0.999	1.000
150	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	0.997	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	1.000	1.000	1.000	1.000	1.000	0.994	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.984	1.000	1.000	1.000	1.000	0.997	0.983	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.980	1.000	1.000	1.000	1.000	1.000	0.992	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	1.000	1.000	1.000	1.000	1.000	0.985	1.000
	-1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	1.000	1.000	1.000	1.000	1.000	0.983	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	1.000	1.000	1.000	1.000	1.000	0.968	1.000
	-0.8	1.000	1.000	1.000	1.000	0.998	0.998	0.999	1.000	1.000	1.000	0.999	0.975	0.999	1.000	1.000	1.000	1.000	0.967	0.999
	-0.7	0.999	1.000	1.000	1.000	0.990	0.996	0.995	1.000	1.000	1.000	0.995	0.977	0.998	1.000	1.000	1.000	0.994	0.945	0.993
	-0.6	0.992	1.000	0.999	0.995	0.948	0.980	0.984	0.996	1.000	1.000	0.998	0.956	0.986	0.993	1.000	0.998	0.997	0.916	0.980
350	-0.5	0.965	0.994	0.989	0.977	0.850	0.939	0.942	0.973	0.998	0.992	0.985	0.867	0.892	0.966	0.990	0.986	0.974	0.861	0.888
	-0.4	0.845	0.958	0.932	0.903	0.617	0.809	0.802	0.879	0.971	0.956	0.926	0.647	0.760	0.873	0.968	0.943	0.917	0.625	0.771
	-0.3	0.653	0.827	0.773	0.715	0.353	0.600	0.573	0.666	0.854	0.799	0.753	0.375	0.568	0.655	0.857	0.809	0.740	0.369	0.593
	-0.2	0.352	0.569	0.502	0.413	0.128	0.342	0.298	0.369	0.620	0.549	0.460	0.121	0.343	0.360	0.649	0.606	0.458	0.125	0.433
	-0.1	0.164	0.304	0.278	0.186	0.036	0.153	0.100	0.138	0.360	0.296	0.188	0.027	0.173	0.141	0.454	0.426	0.189	0.034	0.296
	0.1	0.340	0.165	0.321	0.163	0.017	0.327	0.259	0.324	0.204	0.304	0.158	0.007	0.333	0.475	0.377	0.551	0.336	0.062	0.274
	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	1.000
	-0.3	0.986	1.000	0.998	0.994	0.946	0.988	0.994	0.968	1.000	1.000	1.000	0.950	0.976	0.992	1.000	1.000	0.998	0.938	0.996
	-0.2	0.848	0.938	0.900	0.888	0.592	0.828	0.838	0.868	0.950	0.920	0.902	0.606	0.812	0.848	0.932	0.894	0.878	0.590	0.854
	-0.1	0.312	0.466	0.404	0.366	0.132	0.302	0.302	0.344	0.482	0.432	0.404	0.142	0.286	0.334	0.472	0.416	0.388	0.148	0.288
	0.1	0.615	0.388	0.551	0.493	0.152	0.607	0.592	0.625	0.389	0.557	0.499	0.140	0.586	0.679	0.485	0.632	0.578	0.201	0.578

Cenário(c): $\phi \neq \theta$ ($\phi, \theta + \Delta$)

Neste cenário avaliamos o poder empírico dos testes assumindo os valores verdadeiros dos parâmetros $\theta = 0.1 + \Delta, 0.5 + \Delta, 0.8 + \Delta$ e $\phi = 0.1, 0.6$ e 0.8 , apresentados respectivamente, nas Tabelas 4.16, 4.17 e 4.18.

Com base na Tabela 4.16, segue os resultados:

Considerando a estimação MVC, o poder empírico do teste RVC é maior quando $\Delta \downarrow -1$. Entretanto, quando $\Delta \uparrow 0.8$, o teste RVNC teve poder superior ao teste RVC. Como nos cenários anteriores, os testes são assintoticamente equivalentes.

Os testes segundo as medidas de distâncias quando $\Delta \uparrow 0.8$ apresentam poderes empíricos maiores quando são comparados com os testes da razão entre verossimilhanças, com exceção do teste induzido da distância de *Hellinger* que por sua vez apresentou o pior resultado.

As seguintes desigualdades podem ser verificadas:

(i) Quando $\phi = 0.1, \theta = 0.1 + \Delta$ e $\Delta \downarrow -1$ (Estimação MVC):

$$\begin{aligned} \pi_{\text{TKL}:35;50}(\phi, \theta + \Delta) &> \pi_{\text{JS}:35;50}(\phi, \theta + \Delta) > \pi_{\text{KL}:35;50}(\phi, \theta + \Delta) > \pi_{\text{B}:35;50}(\phi, \theta + \Delta) \\ &> \pi_{\text{RVC}:35;50}(\phi, \theta + \Delta) > \pi_{\text{RVNC}:35;50}(\phi, \theta + \Delta) > \pi_{\text{H}:35;50}(\phi, \theta + \Delta) \\ \pi_{\text{RVC}:150}(\phi, \theta + \Delta) &\approx \pi_{\text{RVNC}:150}(\phi, \theta + \Delta) \approx \pi_{\text{TKL}:150}(\phi, \theta + \Delta) \approx \pi_{\text{KL}:150}(\phi, \theta + \Delta) \\ &\approx \pi_{\text{B}:150}(\phi, \theta + \Delta) \approx \pi_{\text{JS}:150}(\phi, \theta + \Delta) > \pi_{\text{H}:150}(\phi, \theta + \Delta) \end{aligned}$$

(ii) Quando $\phi = 0.1, \theta = 0.1 + \Delta, \Delta \downarrow -1$ (Estimação MVNC ou MQNC) ou $\Delta \uparrow 0.8$: A relação de desigualdades acima assumindo para os termos 5º e 6º que $\pi_{\text{RVNC}:n}(\phi, \theta + \Delta) > \pi_{\text{RVC}:n}(\phi, \theta + \Delta)$ para $n = 35, 50, 150$.

De acordo com a Tabela 4.17, a comparação entre os testes da razão de verossimilhanças RVC e RVNC seguem o mesmo padrão citado nos casos anteriores. Quando $\Delta \downarrow -1.4$, o poder do teste baseado nas distâncias *Kullback-Leibler* e *Jensen-Shannon* tiveram maiores poderes (onde $\pi_{\text{KL}}(\phi + \Delta, \theta) > \pi_{\text{JS}}(\phi + \Delta, \theta)$) e quando $\Delta \uparrow 0.4$, os testes *Tipo Kullback-Leibler* e *Jensen-Shannon* tem melhores performances (onde $\pi_{\text{TKL}}(\phi + \Delta, \theta) > \pi_{\text{JS}}(\phi + \Delta, \theta)$).

As seguintes desigualdades de funções poder empíricas podem ser verificadas:

(i) Quando $\phi = 0.5, \theta = 0.5 + \Delta$ e $\Delta \downarrow -1.4$ (Estimação MVC):

$$\begin{aligned} \pi_{\text{KL}:35;50}(\phi, \theta + \Delta) &> \pi_{\text{JS}:35;50}(\phi, \theta + \Delta) > \pi_{\text{TKL}:35;50}(\phi, \theta + \Delta) > \pi_{\text{B}:35;50}(\phi, \theta + \Delta) \\ &> \pi_{\text{RVC}:35;50}(\phi, \theta + \Delta) > \pi_{\text{RVNC}:35;50}(\phi, \theta + \Delta) > \pi_{\text{H}:35;50}(\phi, \theta + \Delta) \\ \pi_{\text{KL}:150}(\phi, \theta + \Delta) &> \pi_{\text{JS}:150}(\phi, \theta + \Delta) \approx \pi_{\text{JS}:150}(\phi, \theta + \Delta) \approx \pi_{\text{TKL}:150}(\phi, \theta + \Delta) \\ &\approx \pi_{\text{RVC}:150}(\phi, \theta + \Delta) \approx \pi_{\text{RVNC}:150}(\phi, \theta + \Delta) > \pi_{\text{H}:150}(\phi, \theta + \Delta) \end{aligned}$$

- (ii) Quando $\phi = 0.5$, $\theta = 0.5 + \Delta$ e $\Delta \downarrow -1.4$ (Estimação MVNC ou MQNC): A relação de desigualdades acima assumindo para os termos 5º e 6º que $\pi_{RVNC:n}(\phi, \theta + \Delta) > \pi_{RVC:n}(\phi, \theta + \Delta)$ para $n = 35, 50, 150$.
- (iii) Quando $\phi = 0.5$, $\theta = 0.5 + \Delta$ e $\Delta \uparrow 0.4$ (Estimação MVC):

$$\begin{aligned} \pi_{TKL:35;50}(\phi, \theta + \Delta) &> \pi_{JS:35;50}(\phi, \theta + \Delta) > \pi_{B:35;50}(\phi, \theta + \Delta) > \pi_{KL:35;50}(\phi, \theta + \Delta) \\ &> \pi_{RVC:35;50}(\phi, \theta + \Delta) > \pi_{RVNC:35;50}(\phi, \theta + \Delta) > \pi_{H:35;50}(\phi, \theta + \Delta) \\ \pi_{TKL:150}(\phi, \theta + \Delta) &> \pi_{JS:150}(\phi, \theta + \Delta) \approx \pi_{B:150}(\phi, \theta + \Delta) \approx \pi_{KL:150}(\phi, \theta + \Delta) \\ &\approx \pi_{RVC:150}(\phi, \theta + \Delta) \approx \pi_{RVNC:150}(\phi, \theta + \Delta) > \pi_{H:150}(\phi, \theta + \Delta) \end{aligned}$$

- (iv) Quando $\phi = 0.5$, $\theta = 0.5 + \Delta$ e $\Delta \uparrow 0.4$ (Estimação MVNC ou MQNC): A relação de desigualdades acima assumindo para os termos 5º e 6º que $\pi_{RVNC:n}(\phi, \theta + \Delta) > \pi_{RVC:n}(\phi, \theta + \Delta)$ para $n = 35, 50, 150$.

Analisando a Tabela 4.18, pode-se chegar as seguintes conclusões:

Com base na estimação MVC, o poder empírico do teste RVC é maior em relação ao teste RVNC. Entretanto, segundo os métodos MVNC e MQNC, o teste RVNC teve poder superior ao teste RVC. Para tamanhos grandes de amostras os testes são aproximados. Os testes a luz das medidas de distância em geral tem desempenho superior aos testes RVC e RVNC. Quando $\Delta \downarrow -1.4$, o poder do teste baseado nas distâncias *Tipo Kullback-Leibler* e *Jensen-Shannon* tiveram maior poder (onde $\pi_{TKL}(\phi + \Delta, \theta) > \pi_{JS}(\phi + \Delta, \theta)$) e quando $\Delta \uparrow 0.4$, os testes *Kullback-Leibler* e *Jensen-Shannon* tem melhores performances (onde $\pi_{KL}(\phi + \Delta, \theta) > \pi_{JS}(\phi + \Delta, \theta)$).

As seguintes desigualdades de funções poder empíricas podem ser verificadas:

- (i) Quando $\phi = 0.8$, $\theta = 0.8 + \Delta$ e $\Delta \downarrow -1.7$ (Estimação MVC):

$$\begin{aligned} \pi_{KL:35;50}(\phi, \theta + \Delta) &> \pi_{JS:35;50}(\phi, \theta + \Delta) > \pi_{TKL:35;50}(\phi, \theta + \Delta) > \pi_{B:35;50}(\phi, \theta + \Delta) \\ &> \pi_{RVC:35;50}(\phi, \theta + \Delta) \approx \pi_{RVNC:35;50}(\phi, \theta + \Delta) > \pi_{H:35;50}(\phi, \theta + \Delta) \\ \pi_{KL:150}(\phi, \theta + \Delta) &> \pi_{JS:150}(\phi, \theta + \Delta) \approx \pi_{TKL:150}(\phi, \theta + \Delta) \approx \pi_{B:150}(\phi, \theta + \Delta) \\ &\approx \pi_{RVC:150}(\phi, \theta + \Delta) \approx \pi_{RVNC:150}(\phi, \theta + \Delta) > \pi_{H:150}(\phi, \theta + \Delta) \end{aligned}$$

- (ii) Quando $\phi = 0.8$, $\theta = 0.8 + \Delta$ e $\Delta \downarrow -1.7$ (Estimação MVNC ou MQNC):

$$\begin{aligned}\pi_{\text{KL}:35}(\phi, \theta + \Delta) &> \pi_{\text{JS}:35}(\phi, \theta + \Delta) > \pi_{\text{TKL}:35}(\phi, \theta + \Delta) > \pi_{\text{B}:35}(\phi, \theta + \Delta) \\ &> \pi_{\text{RVNC}:35}(\phi, \theta + \Delta) \approx \pi_{\text{RVC}:35}(\phi, \theta + \Delta) > \pi_{\text{H}:35}(\phi, \theta + \Delta)\end{aligned}$$

$$\begin{aligned}\pi_{\text{KL}:50}(\phi, \theta + \Delta) &> \pi_{\text{JS}:50}(\phi, \theta + \Delta) > \pi_{\text{RVNC}:50}(\phi, \theta + \Delta) > \pi_{\text{RVC}:50}(\phi, \theta + \Delta) \\ &> \pi_{\text{B}:50}(\phi, \theta + \Delta) \approx \pi_{\text{TKL}:50}(\phi, \theta + \Delta) > \pi_{\text{H}:50}(\phi, \theta + \Delta)\end{aligned}$$

$$\begin{aligned}\pi_{\text{KL}:150}(\phi, \theta + \Delta) &> \pi_{\text{JS}:150}(\phi, \theta + \Delta) \approx \pi_{\text{RVNC}:150}(\phi, \theta + \Delta) \approx \pi_{\text{RVC}:150}(\phi, \theta + \Delta) \\ &\approx \pi_{\text{B}:150}(\phi, \theta + \Delta) \approx \pi_{\text{TKL}:150}(\phi, \theta + \Delta) > \pi_{\text{H}:150}(\phi, \theta + \Delta)\end{aligned}$$

(iii) Quando $\phi = 0.8$, $\theta = 0.8 + \Delta$ e $\Delta = 0.1$ (Estimação MVC):

$$\begin{aligned}\pi_{\text{TKL}:35;50}(\phi, \theta + \Delta) &> \pi_{\text{JS}:35;50}(\phi, \theta + \Delta) > \pi_{\text{RVC}:35;50}(\phi, \theta + \Delta) > \pi_{\text{RVNC}:35;50}(\phi, \theta + \Delta) \\ &> \pi_{\text{B}:35;50}(\phi, \theta + \Delta) > \pi_{\text{KL}:35;50}(\phi, \theta + \Delta) > \pi_{\text{H}:35;50}(\phi, \theta)\end{aligned}$$

$$\begin{aligned}\pi_{\text{TKL}:150}(\phi, \theta + \Delta) &> \pi_{\text{JS}:150}(\phi, \theta + \Delta) > \pi_{\text{RVC}:150}(\phi, \theta + \Delta) > \pi_{\text{RVNC}:150}(\phi, \theta + \Delta) \\ &> \pi_{\text{B}:150}(\phi, \theta + \Delta) > \pi_{\text{KL}:150}(\phi, \theta + \Delta) > \pi_{\text{H}:150}(\phi, \theta + \Delta)\end{aligned}$$

(iv) Quando $\phi = 0.8$, $\theta = 0.8 + \Delta$ e $\Delta = 0.1$ (Estimação MVNC ou MQNC): A relação de desigualdades acima assumindo para os termos 5º e 6º que $\pi_{\text{RVNC}:n}(\phi, \theta + \Delta) > \pi_{\text{RVC}:n}(\phi, \theta + \Delta)$ para $n = 35, 50, 150$.

Tabela 4.16: Cenário (c): Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $ARMA(1,1)$ com $\phi = 0.1$, $\theta = 0.1 + \Delta$ e $\sigma^2 = 1$ conhecido

n	Δ	Condicional						Não Condicional						Mínimos Quadrados						
		KL	TKL	JS	B	RVC	RVNC	KL	TKL	JS	B	RVC	RVNC	KL	TKL	JS	B	H	RVC	RVNC
35	-1	0.993	0.995	0.995	0.993	0.983	0.987	1.000	1.000	1.000	0.999	0.943	0.995	1.000	1.000	1.000	1.000	1.000	0.927	0.992
	-0.9	0.992	0.989	0.991	0.989	0.978	0.986	0.997	0.995	0.997	0.995	0.985	0.994	1.000	0.998	1.000	1.000	0.993	0.994	
	-0.8	0.972	0.969	0.970	0.967	0.918	0.954	0.968	0.966	0.968	0.965	0.918	0.932	0.954	0.985	0.985	0.982	0.944	0.929	
	-0.7	0.962	0.973	0.970	0.963	0.837	0.925	0.976	0.989	0.985	0.981	0.854	0.922	0.945	0.987	0.992	0.994	0.984	0.908	
	-0.6	0.856	0.866	0.864	0.844	0.614	0.827	0.885	0.884	0.894	0.874	0.634	0.833	0.850	0.906	0.921	0.918	0.900	0.682	
	-0.5	0.705	0.717	0.726	0.687	0.360	0.721	0.758	0.745	0.771	0.725	0.361	0.685	0.700	0.784	0.798	0.807	0.775	0.442	
	-0.4	0.483	0.506	0.514	0.461	0.208	0.431	0.491	0.459	0.487	0.441	0.201	0.521	0.467	0.575	0.601	0.605	0.549	0.293	
	-0.3	0.300	0.358	0.343	0.291	0.102	0.358	0.338	0.344	0.349	0.297	0.087	0.328	0.288	0.398	0.430	0.443	0.376	0.159	
	-0.2	0.191	0.261	0.239	0.201	0.049	0.208	0.201	0.268	0.249	0.204	0.040	0.200	0.157	0.300	0.366	0.360	0.293	0.119	
	-0.1	0.113	0.196	0.163	0.119	0.038	0.125	0.101	0.164	0.145	0.094	0.020	0.182	0.084	0.166	0.260	0.233	0.174	0.082	
50	0.1	0.108	0.161	0.152	0.110	0.020	0.098	0.092	0.163	0.149	0.098	0.015	0.110	0.072	0.166	0.234	0.233	0.173	0.055	
	0.2	0.207	0.307	0.295	0.225	0.055	0.169	0.209	0.310	0.288	0.225	0.039	0.213	0.160	0.284	0.394	0.385	0.314	0.101	
	0.3	0.371	0.437	0.434	0.364	0.120	0.340	0.368	0.449	0.425	0.371	0.128	0.347	0.338	0.436	0.518	0.516	0.440	0.176	
	0.4	0.559	0.644	0.629	0.573	0.228	0.526	0.571	0.653	0.636	0.587	0.237	0.510	0.530	0.624	0.706	0.699	0.643	0.291	
	0.5	0.767	0.832	0.824	0.788	0.439	0.715	0.775	0.852	0.834	0.794	0.423	0.683	0.723	0.825	0.885	0.876	0.845	0.514	
	0.6	0.914	0.959	0.957	0.928	0.704	0.879	0.929	0.968	0.962	0.944	0.725	0.836	0.896	0.951	0.978	0.978	0.963	0.803	
	0.7	0.973	0.987	0.984	0.976	0.867	0.937	0.981	0.993	0.992	0.987	0.909	0.894	0.964	0.988	0.999	0.998	0.993	0.943	
	0.8	0.992	0.990	0.993	0.988	0.940	0.971	0.995	1.000	1.000	0.999	0.970	0.982	0.982	0.999	1.000	1.000	0.988	0.858	
	-1	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000	0.978	1.000	1.000	1.000	1.000	1.000	1.000	0.953	
	-0.9	0.990	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	0.998	0.978	1.000	1.000	1.000	1.000	1.000	0.999	0.969	
150	-0.8	0.997	0.997	0.997	0.997	0.996	0.995	0.995	0.998	0.998	0.998	0.966	0.983	0.996	0.998	0.998	0.998	0.997	0.972	
	-0.7	0.991	0.990	0.990	0.990	0.948	0.986	0.999	0.998	0.999	0.999	0.965	0.993	0.999	0.998	0.999	0.999	0.973	0.963	
	-0.6	0.965	0.960	0.965	0.960	0.848	0.938	0.984	0.979	0.984	0.976	0.864	0.926	0.963	0.988	0.987	0.987	0.986	0.930	
	-0.5	0.870	0.887	0.885	0.871	0.620	0.855	0.896	0.914	0.913	0.892	0.639	0.824	0.856	0.912	0.931	0.938	0.920	0.668	
	-0.4	0.685	0.723	0.719	0.689	0.378	0.689	0.721	0.766	0.755	0.715	0.376	0.639	0.665	0.747	0.814	0.801	0.763	0.431	
	-0.3	0.460	0.569	0.544	0.484	0.179	0.506	0.409	0.569	0.524	0.450	0.177	0.504	0.430	0.510	0.714	0.665	0.580	0.291	
	-0.2	0.218	0.244	0.240	0.200	0.058	0.276	0.246	0.223	0.233	0.190	0.049	0.263	0.210	0.315	0.331	0.336	0.284	0.124	
	-0.1	0.106	0.169	0.152	0.120	0.034	0.133	0.101	0.150	0.144	0.094	0.022	0.150	0.087	0.172	0.244	0.229	0.177	0.086	
	0.1	0.137	0.171	0.164	0.136	0.035	0.105	0.119	0.163	0.143	0.111	0.024	0.115	0.080	0.176	0.220	0.219	0.172	0.066	
	0.2	0.250	0.284	0.282	0.251	0.062	0.235	0.241	0.287	0.276	0.244	0.061	0.225	0.227	0.296	0.328	0.324	0.279	0.087	

Tabela 4.17: Cenário (c): Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $ARMA(1, 1)$ com $\phi = 0.5$, $\theta = 0.5 + \Delta$ e $\sigma^2 = 1$ conhecido

n	Δ	Condicional						Não Condicional						Mínimos Quadrados					
		KL	TKL	JS	H	RVC	RVNC	KL	TKL	JS	H	RVC	RVNC	KL	TKL	JS	H	RVC	RVNC
35	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.993	1.000	1.000	1.000	1.000	1.000	0.985
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000	0.982
	-1.1	1.000	1.000	1.000	1.000	0.998	0.998	1.000	1.000	1.000	1.000	0.998	0.990	0.999	1.000	1.000	1.000	0.998	0.979
	-1	0.999	0.999	0.999	0.998	0.991	0.995	0.996	0.999	0.999	0.999	0.992	0.976	0.996	0.997	0.999	0.999	0.992	0.971
	-0.9	0.996	0.997	0.997	0.996	0.962	0.990	0.986	0.996	0.998	0.997	0.969	0.981	0.990	0.995	0.995	0.995	0.963	0.969
	-0.8	0.984	0.985	0.985	0.982	0.910	0.979	0.985	0.987	0.987	0.982	0.920	0.960	0.982	0.983	0.985	0.981	0.911	0.954
	-0.7	0.966	0.967	0.971	0.960	0.786	0.945	0.939	0.973	0.973	0.973	0.969	0.792	0.937	0.947	0.970	0.969	0.964	0.782
	-0.6	0.920	0.901	0.914	0.896	0.634	0.876	0.864	0.921	0.909	0.916	0.901	0.623	0.861	0.877	0.910	0.893	0.904	0.888
	-0.5	0.801	0.757	0.785	0.754	0.411	0.760	0.749	0.817	0.770	0.797	0.767	0.410	0.743	0.760	0.801	0.752	0.785	0.746
50	-0.4	0.625	0.572	0.603	0.568	0.237	0.599	0.587	0.631	0.593	0.616	0.581	0.231	0.582	0.604	0.612	0.568	0.602	0.567
	-0.3	0.426	0.402	0.413	0.382	0.124	0.379	0.373	0.426	0.383	0.412	0.376	0.117	0.367	0.384	0.409	0.381	0.402	0.361
	-0.2	0.292	0.261	0.280	0.252	0.080	0.227	0.225	0.292	0.273	0.281	0.258	0.077	0.235	0.224	0.450	0.425	0.442	0.417
	-0.1	0.122	0.122	0.123	0.099	0.019	0.085	0.072	0.115	0.126	0.126	0.103	0.013	0.088	0.076	0.167	0.160	0.161	0.141
	0.1	0.132	0.186	0.170	0.125	0.019	0.096	0.074	0.146	0.218	0.199	0.146	0.014	0.116	0.097	0.260	0.303	0.300	0.253
	0.2	0.212	0.329	0.282	0.232	0.039	0.155	0.126	0.230	0.382	0.331	0.278	0.034	0.177	0.161	0.403	0.502	0.462	0.395
	0.3	0.368	0.575	0.517	0.427	0.088	0.364	0.340	0.428	0.653	0.594	0.512	0.078	0.349	0.401	0.633	0.771	0.726	0.659
	0.4	0.603	0.785	0.747	0.666	0.205	0.597	0.563	0.730	0.923	0.881	0.815	0.244	0.562	0.647	0.897	0.966	0.958	0.921
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	1.000	1.000	1.000	1.000	0.989
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	1.000	1.000	1.000	1.000	0.997
150	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	1.000	1.000	1.000	1.000	0.995
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.987	1.000	1.000	1.000	1.000	0.985
	-1	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	0.992
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	0.992
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.993	1.000	1.000	1.000	1.000	0.993
	-0.7	0.996	0.995	0.996	0.994	0.957	0.992	0.992	0.997	0.995	0.996	0.962	0.962	0.991	0.994	0.996	0.995	0.995	0.987
	-0.6	0.978	0.979	0.979	0.976	0.850	0.960	0.962	0.977	0.976	0.978	0.974	0.852	0.957	0.963	0.975	0.973	0.974	0.848
	-0.5	0.917	0.900	0.912	0.900	0.651	0.880	0.883	0.917	0.907	0.912	0.904	0.658	0.874	0.884	0.907	0.891	0.902	0.891
	-0.4	0.768	0.744	0.766	0.733	0.372	0.738	0.734	0.773	0.738	0.767	0.741	0.375	0.729	0.741	0.757	0.718	0.747	0.719
	-0.3	0.565	0.505	0.542	0.512	0.174	0.528	0.526	0.562	0.513	0.546	0.523	0.174	0.524	0.533	0.545	0.496	0.522	0.491

Tabela 4.18: Cenário (c): Poder do Teste induzidos das medidas de divergências e razão de verossimilhanças para o Modelo $ARMA(1, 1)$ com $\phi = 0.8$, $\theta = 0.8 + \Delta$ e $\sigma^2 = 1$ conhecido

n	Δ	Condicional					Não Condicional					Mínimos Quadrados						
		KL	TKL	JS	H	RVNC	KL	TKL	JS	H	RVNC	KL	TKL	JS	H	RVNC	RVNC	
35	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	0.999	1.000	1.000	1.000	0.998	0.999	1.000	1.000	0.998	0.998	0.998	0.998	0.999	1.000	1.000	0.999	0.998
	-1.1	1.000	1.000	1.000	1.000	0.998	0.998	1.000	1.000	0.999	0.998	0.999	0.997	1.000	1.000	1.000	0.994	0.997
	-1	1.000	1.000	1.000	1.000	0.997	1.000	1.000	1.000	1.000	0.998	0.997	0.999	1.000	1.000	1.000	0.996	0.997
	-0.9	0.998	0.998	0.998	0.996	0.971	0.986	0.987	1.000	0.999	0.982	0.986	0.987	0.997	0.994	0.996	0.986	0.987
	-0.8	0.990	0.986	0.992	0.985	0.934	0.975	0.969	0.989	0.986	0.991	0.982	0.941	0.972	0.970	0.983	0.977	0.921
	-0.7	0.989	0.980	0.986	0.981	0.886	0.975	0.970	0.985	0.982	0.983	0.970	0.972	0.984	0.969	0.980	0.972	0.970
	-0.6	0.954	0.907	0.948	0.922	0.704	0.925	0.917	0.958	0.921	0.948	0.926	0.704	0.920	0.923	0.938	0.878	0.929
	-0.5	0.900	0.807	0.878	0.834	0.524	0.855	0.839	0.898	0.813	0.877	0.833	0.529	0.854	0.823	0.872	0.781	0.855
	-0.4	0.988	0.999	0.996	0.990	0.943	0.977	0.978	0.990	0.999	0.997	0.994	0.950	0.965	0.981	0.988	0.996	0.993
	-0.3	0.938	0.959	0.957	0.936	0.702	0.901	0.894	0.937	0.967	0.961	0.943	0.700	0.879	0.908	0.912	0.944	0.947
	-0.2	0.656	0.698	0.705	0.629	0.238	0.624	0.576	0.649	0.715	0.711	0.631	0.249	0.608	0.599	0.580	0.638	0.638
	-0.1	0.258	0.291	0.296	0.231	0.047	0.242	0.185	0.245	0.311	0.288	0.227	0.043	0.253	0.205	0.189	0.276	0.255
	0.1	0.097	0.103	0.100	0.063	0.013	0.095	0.029	0.047	0.072	0.063	0.032	0.043	0.256	0.054	0.125	0.158	0.133
50	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.4	0.896	0.804	0.870	0.838	0.565	0.862	0.852	0.890	0.808	0.871	0.840	0.558	0.863	0.856	0.870	0.780	0.855
	-0.3	0.739	0.591	0.700	0.627	0.259	0.713	0.659	0.722	0.578	0.674	0.609	0.251	0.709	0.663	0.672	0.534	0.636
	-0.2	0.461	0.303	0.402	0.320	0.084	0.447	0.398	0.436	0.294	0.379	0.316	0.088	0.441	0.402	0.392	0.286	0.382
	-0.1	0.184	0.149	0.184	0.117	0.023	0.175	0.120	0.138	0.148	0.153	0.102	0.020	0.190	0.133	0.140	0.221	0.245
	0.1	0.088	0.248	0.205	0.074	0.006	0.086	0.040	0.043	0.445	0.313	0.063	0.002	0.269	0.077	0.114	0.672	0.669
150	-1.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1					

Conclusivamente,

para processos $ARMA(1, 1)$, os testes com estatísticas S_{KL} , e S_B combinados com a estimação MVNC se mostraram bastante competitivos e em alguns casos mais robustos do que o teste com base em S_{RVC} e S_{RVNC} .

4.3 Quantificação de *Overfitting*

Nesta seção, apresenta-se um estudo de comparação entre critérios de seleção clássicos discutidos na Seção 2.3 e outros propostos induzidos das distâncias de *Kullback-Leibler*, tipo *Kullback-Leibler*, *Jensen-Shannon* e *Battacharrya*. Para tal fim, utilizou-se experimentos Monte Carlo com 1000 replicações para tamanhos de amostras $n = 15, 35, 50, 100$. Adicionalmente, assumiremos $\{X_t\} \sim AR(1) : [\phi, \sigma^2]$ sob a especificação paramétrica $\phi = -0.99, -0.98, -0.97, \dots, 0.97, 0.98, 0.99$ e $\sigma^2 = 1$ como modelo verdadeiro, totalizando 199 pontos discretizados sobre o intervalo de estacionaridade $\phi \in (-1, 1)$. Adicionalmente apresentamos estudos de casos exibidos nas Tabelas 4.19, 4.21, 4.20 e 4.22. Consideramos dois cenários: Seja $X_t^* = \phi_1^* X_{t-1} + Z_t^*$ o modelo verdadeiro.

Cenário 1: Modelo verdadeiro é um processo *Autoregressive* de ordem 1 enquanto que o modelo proposto segue um *Autoregressive* de ordem 2; isto é, checagem de

$$\mathcal{H}_0 : X_t^* = \phi_1^* X_{t-1} + Z_t^* \text{ contra } \mathcal{H}_1 : X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t.$$

Cenário 2: Modelo verdadeiro é de ordem 1 e o modelo proposto de ordem 3; isto é, checagem de

$$\mathcal{H}_0 : X_t^* = \phi_1^* X_{t-1} + Z_t^* \text{ contra } \mathcal{H}_1 : X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + Z_t.$$

Ambos os cenários são analisados sob as estimações MVC e MVNC. A probabilidade de *overfitting* foi utilizada como critério de comparação dos métodos. Os resultados obtidos para os dois cenários são apresentados nas seções a seguir.

4.3.1 *Overfitting* sobre o Cenário 1

Avaliação dos métodos clássicos de seleção de modelos

Na Figura 4.35, apresentam-se os desempenhos dos critérios clássicos discutidos na Seção 2.3 para os tamanhos de amostras $n = 15$ e $n = 35$ e sob as estimações condicional e não condicional. Em geral, estes critérios não sofrem modificação quanto ao tipo de estimação utilizada. As probabilidade de *overfitting* obedeceram as seguintes desigualdades para os tamanhos 15 e 35, respectivamente:

$$\Pr_{15}(AICu) < \Pr_{15}(AICc) < \Pr_{15}(SIC) < \Pr_{15}(AIC) < \Pr_{15}(HQc) < \Pr_{15}(FEPu) < \Pr_{15}(HQ) < \Pr_{15}(FPE)$$

e

$$\Pr_{35}(AICu) = \Pr_{35}(SIC) < \Pr_{35}(AICc) < \Pr_{35}(AIC) < \Pr_{35}(HQc) < \Pr_{35}(HQ) < \Pr_{35}(FEPu) < \Pr_{35}(FPE).$$

Os critérios AICc (para $n = 15, 35$) e SIC (para $n = 35$) apresentaram os melhores desempenhos dentre oito métodos indicados na literatura. Para o tamanho $n = 35$, as probabilidades de *overfitting* são menores quando aproximadamente $|\phi| < 0,6$, com exceção do critério FEP. Este fato fornece evidências à conclusão de que a probabilidade de erro de classificação assume uma relação inversa com a correlação do processo ($\phi = \rho_X(1) = \text{Cor}(X_{t+1}, X_t)$ para $\{X_t\} \sim \text{AR}(1)$). Como obteve-se a melhora dos critérios SIC e HQ do tamanho $n = 15$ para $n = 35$, o estudo também foi estendido para os tamanhos $n = 50, 100$. A Figura 4.36 exibe o resultado do estudo. Nota-se que aproximações entre os seguintes pares de critérios (HQ, HQc) e (AIC, AICc) quando o tamanho de amostra aumenta como esperado. Por outro lado, as correções AICu e FPEu se mostraram expressivas mesmo no tamanho $n = 100$. Os desempenhos dos critérios obedeceram as seguintes desigualdades para os tamanhos 50 e 100, respectivamente:

$$\Pr_{50}(\text{SIC}) < \Pr_{50}(\text{AICu}) < \Pr_{50}(\text{AICc}) < \Pr_{50}(\text{AIC}) < \Pr_{50}(\text{HQc}) < \Pr_{50}(\text{HQ}) < \Pr_{50}(\text{FEPu}) < \Pr_{50}(\text{FPE})$$

e

$$\Pr_{100}(\text{SIC}) = \Pr_{100}(\text{AICu}) < \Pr_{100}(\text{AICc}) < \Pr_{100}(\text{AIC}) < \Pr_{100}(\text{HQc}) < \Pr_{100}(\text{HQ}) < \Pr_{100}(\text{FEPu}) < \Pr_{100}(\text{FPE}).$$

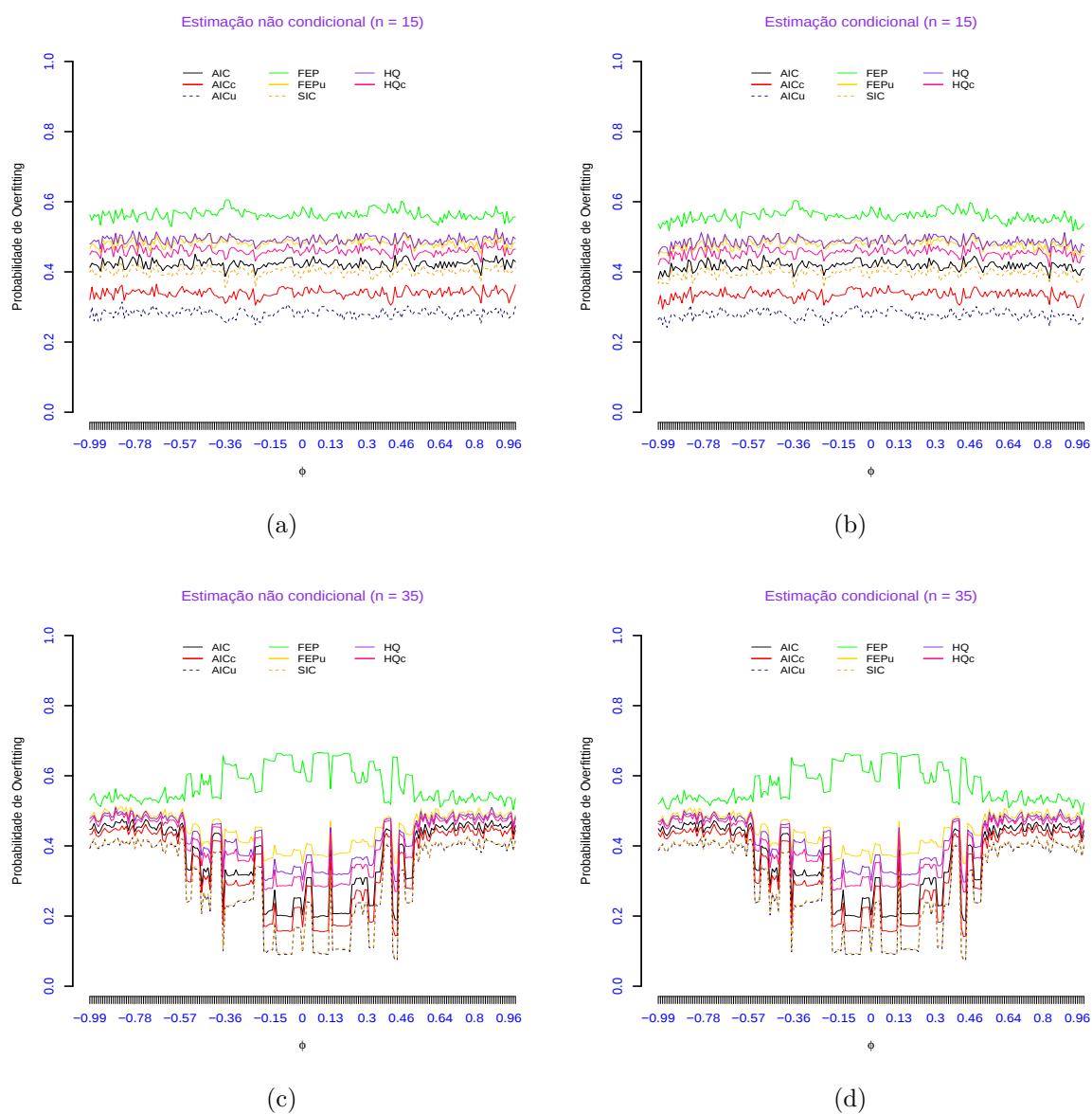
Para ambos os tamanhos de amostra, o critério SIC apresentou o melhor desempenho. Em geral, os métodos clássicos apresentam altas probabilidades de *overfitting* na fronteira de estabilidade, indicando a importância da proposta de critérios que trabalhem bem na região paramétrica $|\phi| > 0.6$ para o processo $\{X_t\} \sim \text{AR}(1)$.

Avaliação dos novos métodos de seleção de modelos

É conveniente começar dizendo que os métodos propostos são condicionados ao uso de um bom estimador não paramétrico para uma das densidades nas medidas de divergências, representadas de um modo geral por $K(f||g)$. Este estudo se constitui importante pois revela que podem existir discretizações (via estimação não-paramétrica de densidade) da estatística (h, ϕ) de divergências com desempenho melhor do que as medidas de seleção clássica.

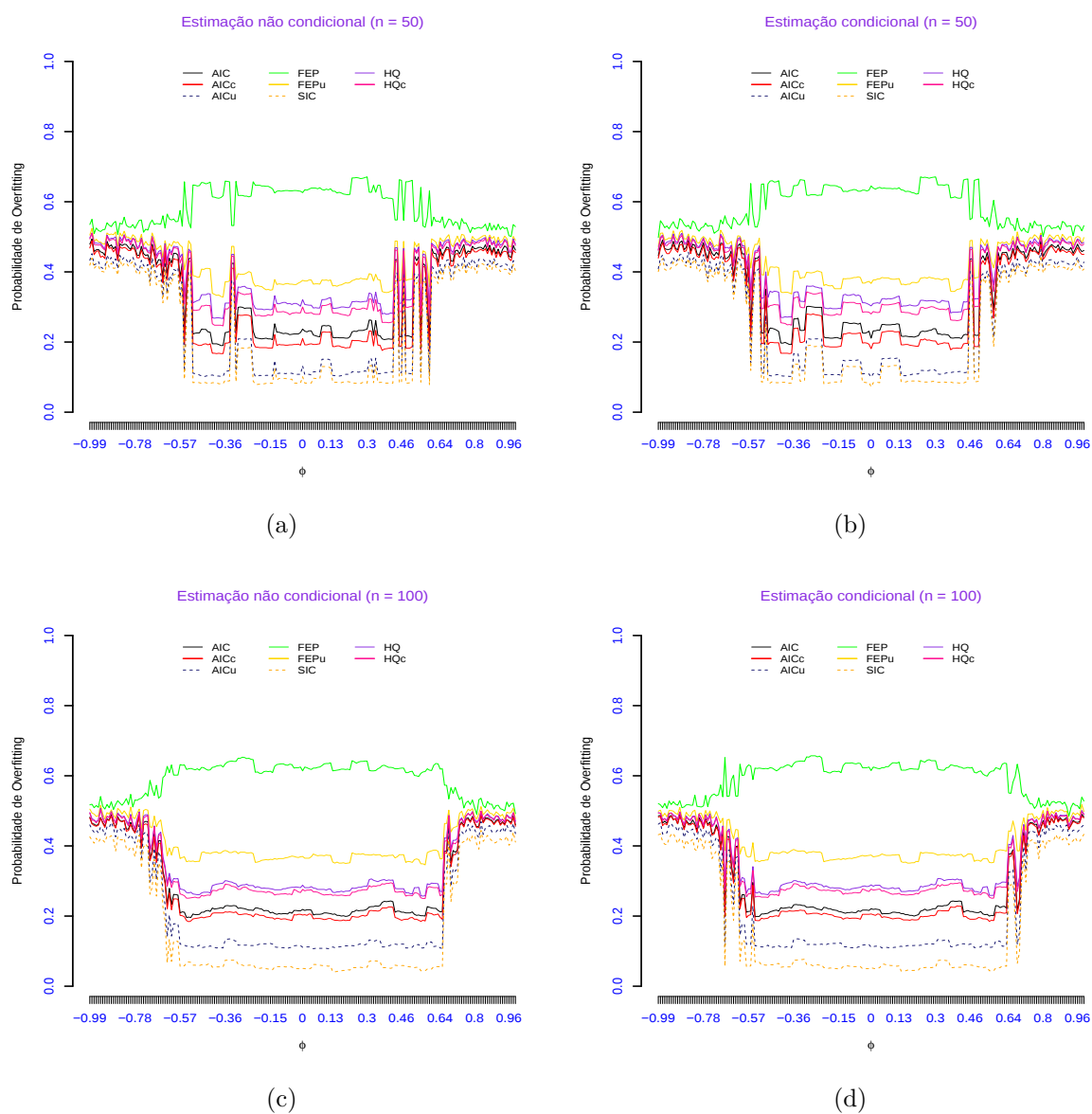
A Figura 4.37 apresenta o resultado do estudo de *overfitting* para os novos critérios. Para os tamanhos $n = 15$ e $n = 35$, observa-se que os métodos em função das distâncias *Kullback-Leibler* e *Battacharrya* apresentaram menores probabilidades de *overfitting* próximo a fronteira de estabilidade para estimação MVNC. Ainda considerando os intervalos limites, os métodos definidos nas medidas tipo *Kullback-Leibler* e *Battacharrya* obtiveram os melhores desempenhos na estimação MVC. Para valores de ϕ próximo a origem (experimentalmente $|\phi| < 0.7$) os novos critérios induzidos da *Battacharrya*, *Kullback-Leibler* e *Jensen-Shannon* apresenram desempenhos equivalentes. Os resultados para tamanhos $n = 50$ e $n = 100$ são apresentados na Figura 4.38. Os novos critérios apresentaram baixas probabilidades de *overfitting* na região de estabilidade próxima a origem (aproximadamente para $|\phi| < 0,6$). Com decaimento dessa probabilidade quando $|\phi| \uparrow 1$. O método induzido da distância de *Kullback-Leibler* teve o melhor desempenho entre as novas propostas na região próxima a origem e próximo a fronteira os métodos mostraram-se equivalentes. É importante ressaltar que a estimação por MVNC reduz significativamente a probabilidade de *overfitting* dos

Figura 4.35: Cenário 1: Probabilidades de *overfitting* dos critérios clássicos de seleção de modelos para amostras pequenas.



Fonte: Autoria própria.

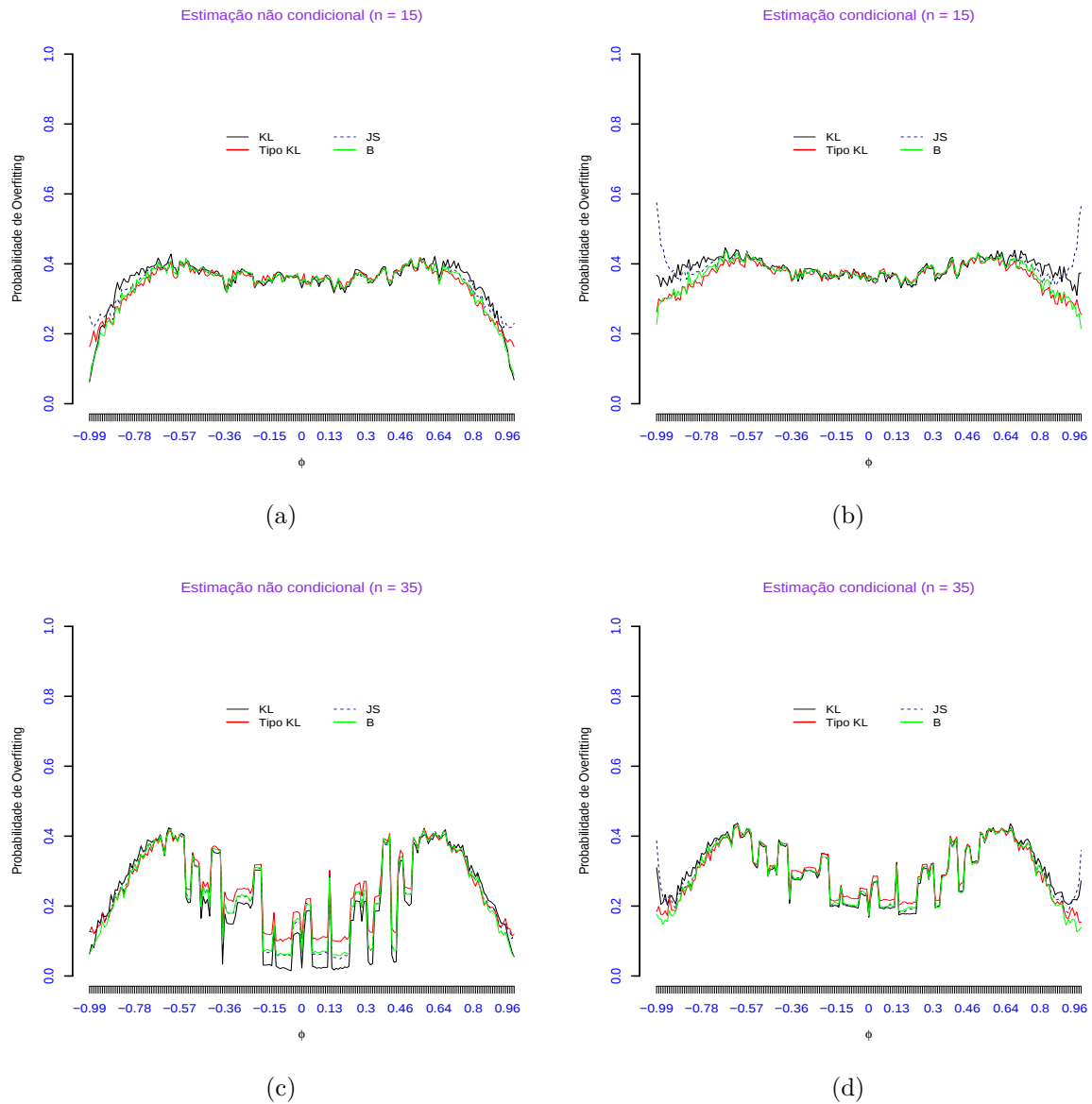
Figura 4.36: Cenário 1: Probabilidades de *overfitting* dos critérios clássicos de seleção de modelos para amostras grandes.



Fonte: Autoria própria.

critérios propostos induzidos das medidas de distâncias (vê o comparativo entre as Figuras 4.37 e 4.38).

Figura 4.37: Cenário 1: Probabilidades de *overfitting* dos novos critérios propostos induzidos das medidas de divergências para amostras pequenas.

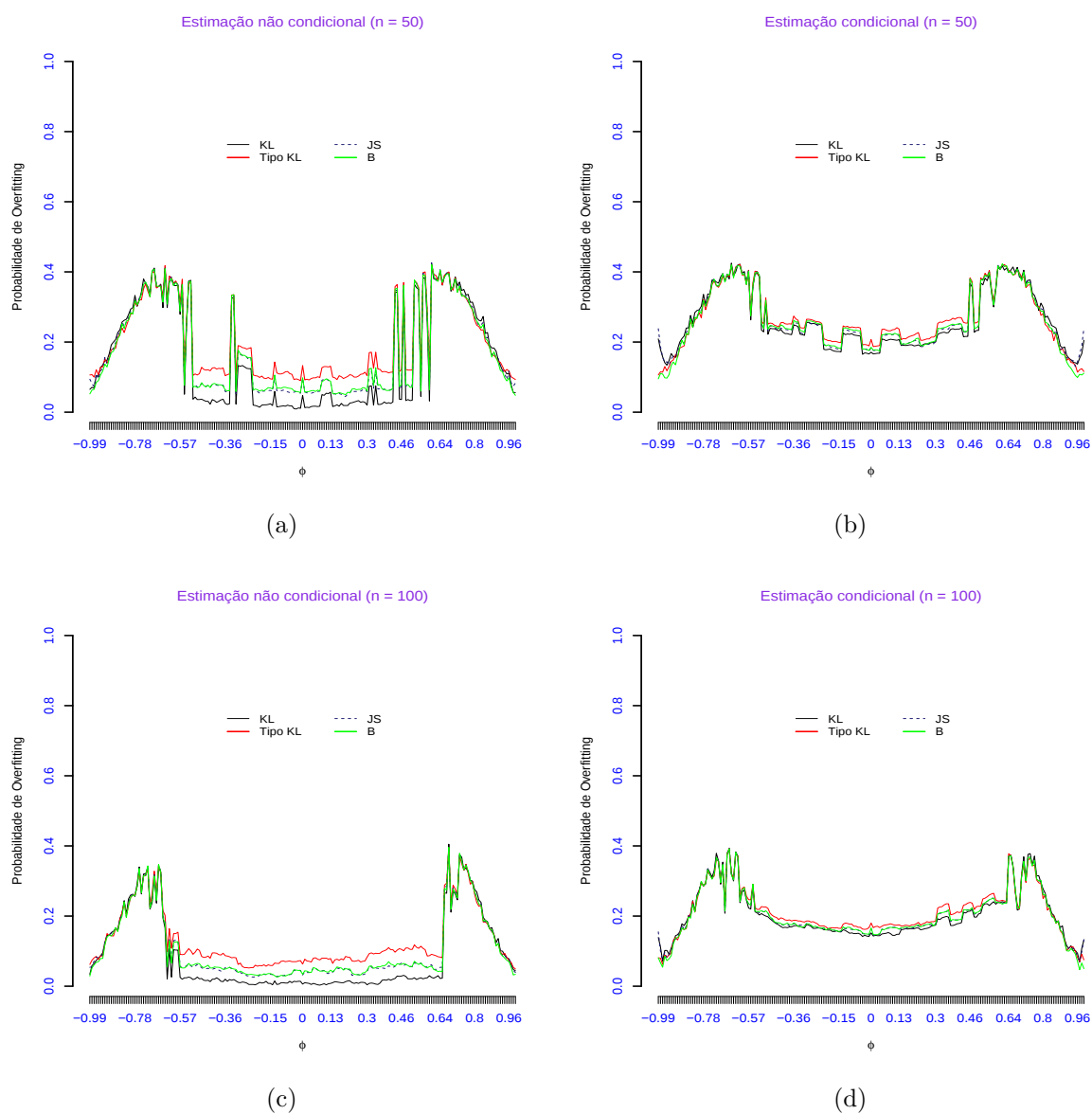


Fonte: Autoria própria.

Comparação entre os métodos clássicos e métodos novos

As Figuras 4.39 e 4.40 comparam os quatro melhores métodos clássicos—AIC, AICc e AICu—with os propostos. Comparando os novos critérios para seleção de modelos com os clássicos no tamanho de amostra $n = 15$ (baseando-se nas Figuras 4.37, vê-se que o uso dos novos métodos combinados com a estimação MVNC é mais recomendado para identificações na fronteira da estabilidade. Entretanto, próximo a origem o critério AICu

Figura 4.38: Cenário 1: Probabilidades de *overfitting* dos novos critérios propostos induzidos das medidas de divergências para amostras grandes.



Fonte: Autoria própria.

teve o melhor desempenho. Em termos de desigualdades,

$$\left\{ \begin{array}{l} \text{Para } |\phi| < 0.8 : \\ \Pr_{15}(\text{AICu}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{B}) < \Pr_{15}(\text{JS}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{KL}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}) \\ \text{Para } [-0.9 < \phi < -0.8] \cap [0.8 < \phi < 0.9] : \\ \Pr_{15}(\text{AICu}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{B}) < \Pr_{15}(\text{JS}) < \Pr_{15}(\text{KL}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}) \\ \text{Para } |\phi| > 0.9 : \\ \Pr_{15}(\text{KL}) = \Pr_{15}(\text{B}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{JS}) < \Pr_{15}(\text{AICu}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}). \end{array} \right.$$

Utilizando a estimação MVC, o critério AICu foi melhor em todos os casos. Para tamanho de amostra $n = 35$ (correspondente a Figura 4.37, em ambos métodos de estimação, os novos critérios apresentaram menores taxas de *overfitting* para $|\phi| > 0,6$ em comparação com os critérios clássicos. Na região de estabilidade próxima a origem os novos critérios induzidos das distâncias *Kullback-Leibler*, *Battacharrya* e *Jensen-Shanon* foram melhores quando o método de estimação é MVNC. Entretanto, os critérios AICu e SIC são melhores para a estimação por MVC.

Sob estimação MVC com tamanho $n = 15$, baseado na Figura 4.39 vale-se a seguinte desigualdade:

$$\left\{ \begin{array}{l} \text{Para } |\phi| < 0.7 : \\ \Pr_{15}(\text{AICu}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{KL}) \approx \Pr_{15}(\text{JS}) \approx \Pr_{15}(\text{B}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}) \\ \text{Para } [-0.9 < \phi < -0.7] \cap [0.7 < \phi < 0.9] : \\ \Pr_{15}(\text{AICu}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{B}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{JS}) < \Pr_{15}(\text{KL}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}) \\ \text{Para } |\phi| > 0.9 : \\ \Pr_{15}(\text{AICu}) = \Pr_{15}(\text{B}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{KL}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}) < \Pr_{15}(\text{JS}). \end{array} \right.$$

e para $n = 35$, segue-se:

$$\left\{ \begin{array}{l} \text{Para } |\phi| < 0.3 : \\ \Pr_{35}(\text{AICu}) < \Pr_{35}(\text{SIC}) < \Pr_{35}(\text{AICc}) < \Pr_{35}(\text{KL}) < \Pr_{35}(\text{B}) \approx \Pr_{35}(\text{JS}) < \Pr_{35}(\text{TKL}) < \Pr_{35}(\text{AIC}) \\ \text{Para } [-0.9 < \phi < -0.7] \cap [0.7 < \phi < 0.9] : \\ \Pr_{35}(\text{B}) < \Pr_{35}(\text{TKL}) < \Pr_{35}(\text{JS}) < \Pr_{35}(\text{KL}) < \Pr_{35}(\text{SIC}) < \Pr_{35}(\text{AICu}) < \Pr_{35}(\text{AICc}) < \Pr_{35}(\text{AIC}) \\ \text{Para } |\phi| > 0.9 : \\ \Pr_{35}(\text{B}) = \Pr_{35}(\text{TKL}) < \Pr_{35}(\text{KL}) < \Pr_{35}(\text{JS}) < \Pr_{35}(\text{SIC}) \approx \Pr_{35}(\text{AICu}) < \Pr_{35}(\text{AICc}) < \Pr_{35}(\text{AIC}). \end{array} \right.$$

A Figura 4.40 apresenta os resultados para tamanhos de amostras grandes ($n = 50$ e $n = 100$). No tocante a estimação MVNC, os critérios induzidos das distâncias *Kullback-Leibler* e *Battacharrya* tiveram os melhores desempenhos. Mas, considerando a estimação por MVC, próximo a origem os critérios SIC e AICu tiveram os melhores desempenhos.

Quando o método de EMVC é considerado as seguintes relações são obtidas(Vê Figura 4.39,

$$\left\{ \begin{array}{l} \text{Para } |\phi| < 0.7 : \\ \Pr_{15}(\text{AICu}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{KL}) < \Pr_{15}(\text{JS}) \approx \Pr_{15}(\text{B}) \approx \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}) \\ \text{Para } [-0.9 < \phi < -0.7] \cap [0.7 < \phi < 0.9] : \\ \Pr_{15}(\text{AICu}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{B}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{JS}) < \Pr_{15}(\text{KL}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}) \\ \text{Para } |\phi| > 0.9 : \\ \Pr_{15}(\text{AICu}) = \Pr_{15}(\text{B}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{KL}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}) < \Pr_{15}(\text{JS}). \end{array} \right.$$

agora considerando tamanho de amostra $n = 50$ (Figura 4.40), tem-se que:

$$\left\{ \begin{array}{l} \text{Para } |\phi| < 0.3 : \\ \Pr_{50}(\text{SIC}) < \Pr_{50}(\text{AICu}) < \Pr_{50}(\text{KL}) < \Pr_{50}(\text{B}) \approx \Pr_{50}(\text{TKL}) \approx \Pr_{50}(\text{JS}) < \Pr_{50}(\text{AICc}) < \Pr_{50}(\text{AIC}) \\ \text{Para } [-0.7 < \phi < -0.3] \cap [0.3 < \phi < 0.7] : \\ \Pr_{50}(\text{SIC}) < \Pr_{50}(\text{AICu}) < \Pr_{50}(\text{JS}) < \Pr_{50}(\text{KL}) < \Pr_{50}(\text{B}) < \Pr_{50}(\text{TKL}) < \Pr_{50}(\text{AICc}) < \Pr_{50}(\text{AIC}) \\ \text{Para } |\phi| > 0.7 : \\ \Pr_{50}(\text{B}) < \Pr_{50}(\text{TKL}) < \Pr_{50}(\text{JS}) < \Pr_{50}(\text{KL}) < \Pr_{50}(\text{SIC}) < \Pr_{50}(\text{AICu}) < \Pr_{50}(\text{AICc}) < \Pr_{50}(\text{AIC}). \end{array} \right.$$

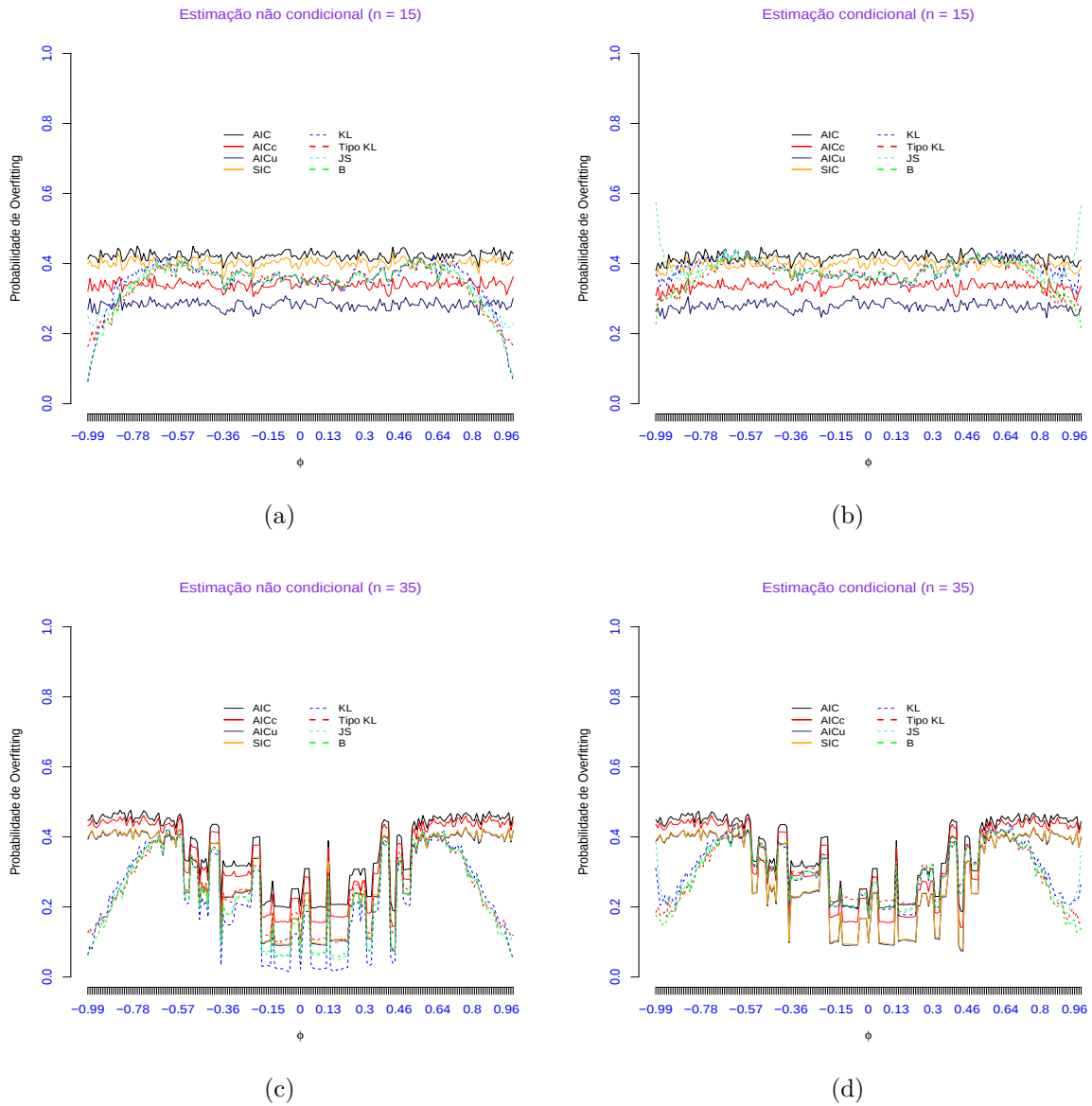
Para tamanho de amostra $n = 100$ (Figura 4.40), tem-se que:

$$\left\{ \begin{array}{l} \text{Para } |\phi| < 0.7 : \\ \Pr_{50}(\text{SIC}) < \Pr_{50}(\text{AICu}) < \Pr_{50}(\text{KL}) < \Pr_{50}(\text{B}) \approx \Pr_{50}(\text{JS}) < \Pr_{50}(\text{TKL}) < \Pr_{50}(\text{AICc}) < \Pr_{50}(\text{AIC}) \\ \text{Para } |\phi| > 0.7 : \\ \Pr_{50}(\text{B}) < \Pr_{50}(\text{TKL}) < \Pr_{50}(\text{JS}) < \Pr_{50}(\text{KL}) < \Pr_{50}(\text{SIC}) < \Pr_{50}(\text{AICu}) < \Pr_{50}(\text{AICc}) < \Pr_{50}(\text{AIC}). \end{array} \right.$$

Conclusivamente, para ambos os métodos de estimação, os novos critérios para tamanhos de amostras grandes apresentaram menores taxas de *overfitting* para $|\phi| > 0,7$

(correspondente a Figura 4.40). Vale salientar que a estimação MVNC diminui a probabilidade de *overfitting* dos critérios de seleção (com exceção do critério FEP); principalmente na região de estabilidade próxima a origem, para $|\phi| \downarrow 0$.

Figura 4.39: Cenário 1: Comparação entre os novos métodos propostos e os métodos clássicos de seleção de modelos para pequenas amostras

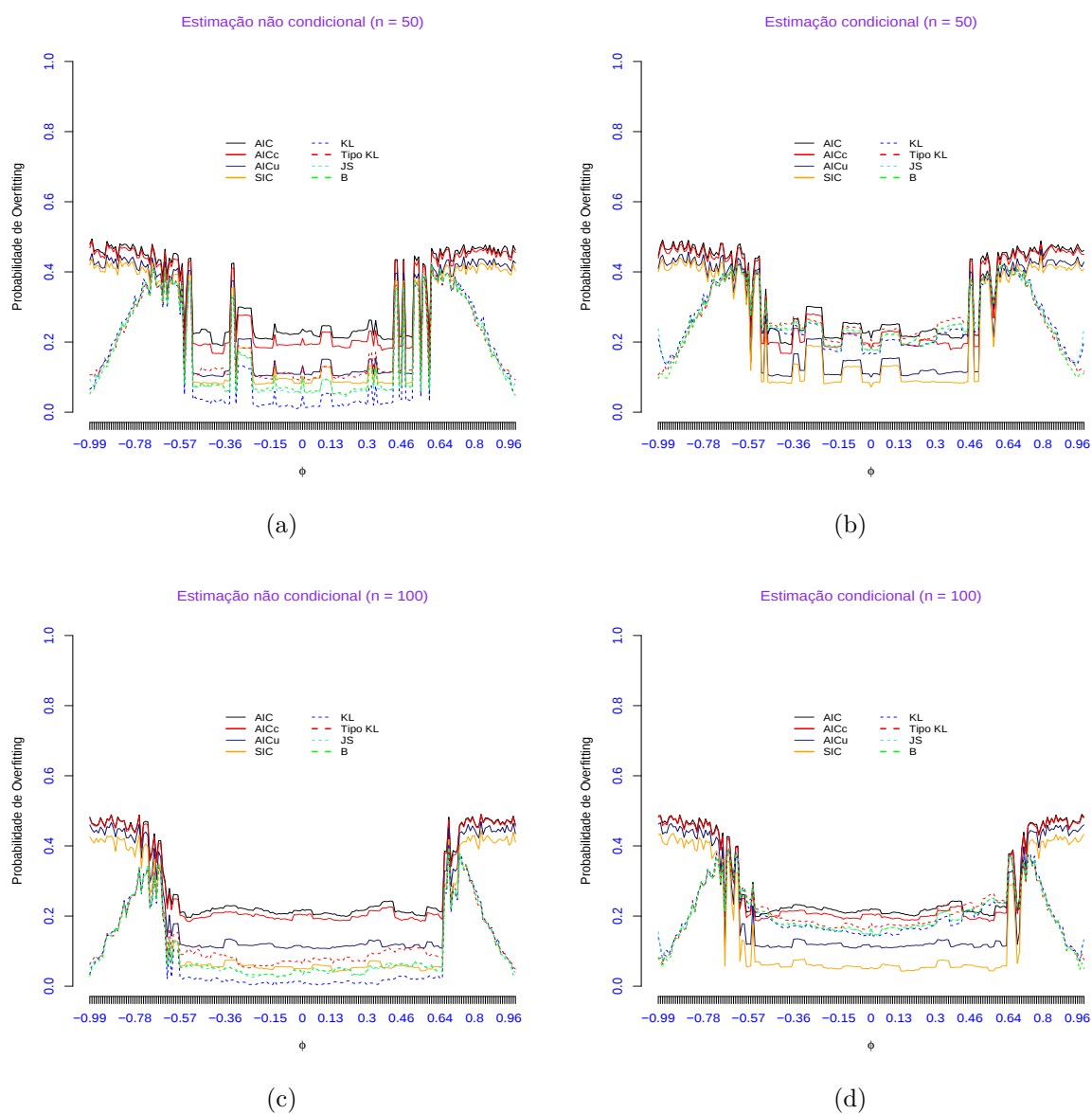


Fonte: Autoria própria.

4.3.2 *Overfitting* sobre o Cenário 2

Neste cenário, o modelo *overfitting* foi um autoregressivo de ordem 3, enquanto que modelo verdadeiro foi *autoregressive* de ordem 1. As taxas de erros de classificação foram utilizadas como critério de classificação.

Figura 4.40: Cenário 1: Comparação entre os novos métodos propostos e os métodos clássicos de seleção de modelos para grandes amostras



Fonte: Autoria própria.

Avaliação dos métodos clássicos de seleção de modelos

Para amostras de tamanhos $n = 15$ e 35 (como apresentado na Figura 4.41), em ambos os métodos de estimação, os critérios AICu e AICc tiveram as menores probabilidades de *overfitting* em toda região de estabilidade. Para estimação MVC (similarmente à discussão no cenário 1), os critérios clássicos apresentaram probabilidades altas quando o valor do parâmetro ϕ está próximo da fronteira de estabilidade. Para tamanhos de amostras $n = 50$ e 100 (na Figura 4.42), o critério SIC seguido do critério AICu apresentaram as menores taxas de *overfitting* para ambos os métodos de estimação e sobre toda região de estabilidade discretizada.

De acordo com as Figuras 4.41 e 4.42, a seguinte relação é mantida:

$$\left\{ \begin{array}{l} \text{Para } MVNC \\ \text{Pr}_{15}(\text{AICu}) < \text{Pr}_{15}(\text{AICc}) < \text{Pr}_{15}(\text{SIC}) < \text{Pr}_{15}(\text{AIC}) \approx \text{Pr}_{15}(\text{HQc}) \approx \text{Pr}_{15}(\text{FEPu}) < \text{Pr}_{15}(\text{HQ}) < \text{Pr}_{15}(\text{FEP}) \\ \text{Pr}_{35}(\text{AICu}) < \text{Pr}_{35}(\text{SIC}) < \text{Pr}_{35}(\text{AICc}) < \text{Pr}_{35}(\text{AIC}) < \text{Pr}_{35}(\text{HQc}) < \text{Pr}_{35}(\text{HQ}) < \text{Pr}_{35}(\text{FEPu}) < \text{Pr}_{35}(\text{FEP}) \\ \text{Pr}_{50;100}(\text{SIC}) < \text{Pr}_{50;100}(\text{AICu}) < \text{Pr}_{50;100}(\text{AIC}) < \text{Pr}_{50;100}(\text{HQc}) < \text{Pr}_{50;100}(\text{HQ}) \\ < \text{Pr}_{50;100}(\text{HQu}) < \text{Pr}_{50;100}(\text{FEPu}) < \text{Pr}_{50;100}(\text{FEP}). \end{array} \right.$$

Avaliação dos novos métodos propostos de seleção de modelos

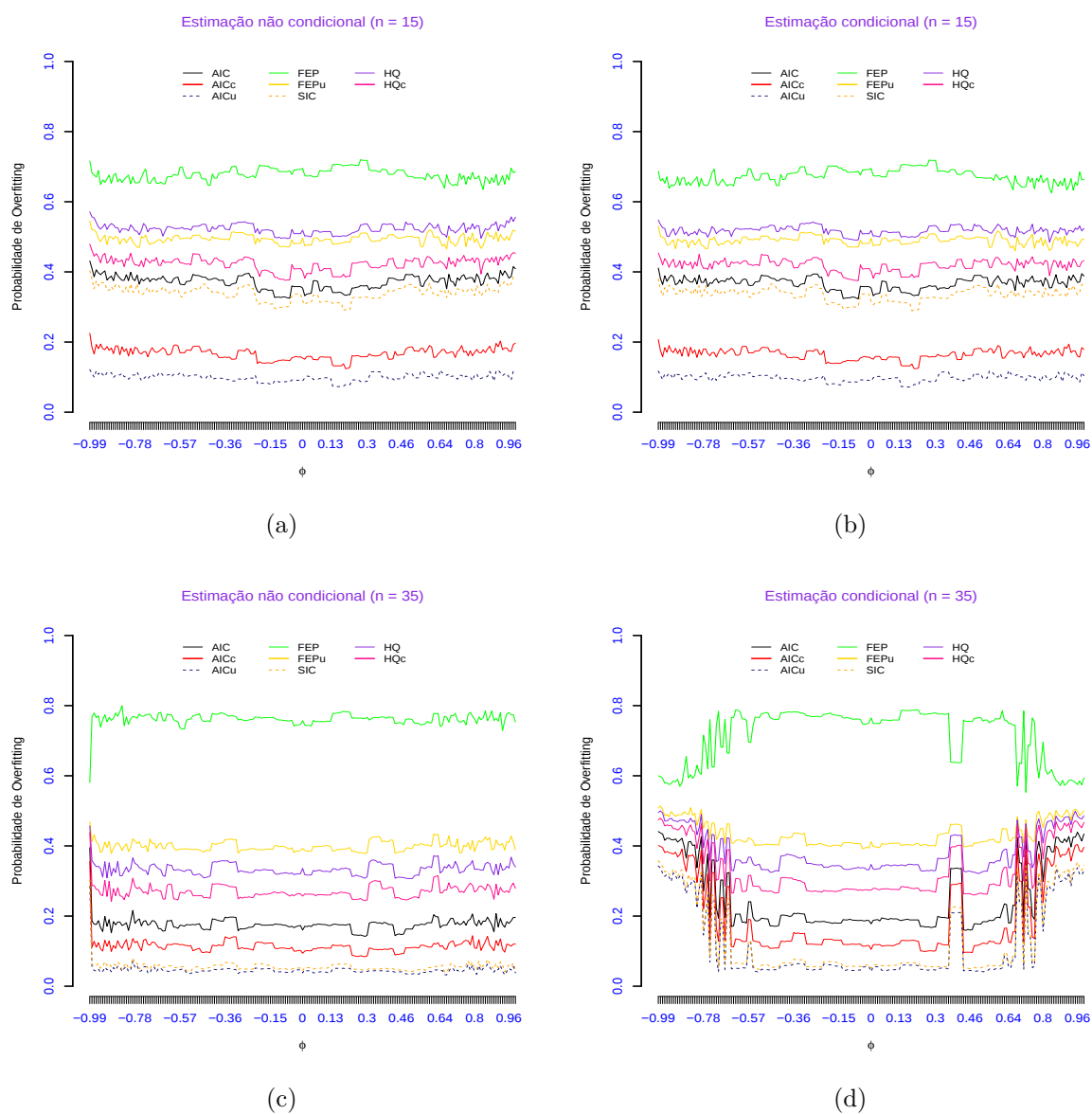
Em geral, os critérios induzidos das medidas de divergência apresentaram comportamento similares em termos de probabilidades de *overfitting* sobre a discretização de pontos na região de estabilidade. Para estimação por MVNC, os novos critérios apresentaram probabilidades de *overfitting* próximas de zero para tamanhos de amostras superior a $n = 35$ (como apresentado nas Figuras 4.45 e 4.46). Para $n = 15$ e estimação MVNC, as taxas de *overfitting* aproximam-se de 20% e, considerando a MVC, essa taxa pode ultrapassar 40% na região próxima a fronteira (vê Figura 4.45). Considerando tamanhos de amostras superior a $n = 35$, as taxas de *overfitting* são próximas de zero para estimação de MVNC. No caso da estimação MVC, na região próxima de zero, a probabilidade de *overfitting* é pequena, enquanto que na fronteira da estabilidade essa probabilidade aumenta.

Com base nas Figuras 4.41 e 4.42, a seguinte relação é mantida para $n = 15, 35, 50$ e 100 :

$$\left\{ \begin{array}{l} \text{Para : } MVNC \\ \text{Pr}_n(\text{AICu}) < \text{Pr}_n(\text{AICc}) < \text{Pr}_n(\text{SIC}) < \text{Pr}_n(\text{AIC}) \approx \text{Pr}_n(\text{HQc}) \approx \text{Pr}_n(\text{FEPu}) < \text{Pr}_n(\text{HQ}) < \text{Pr}_n(\text{FEP}) \end{array} \right.$$

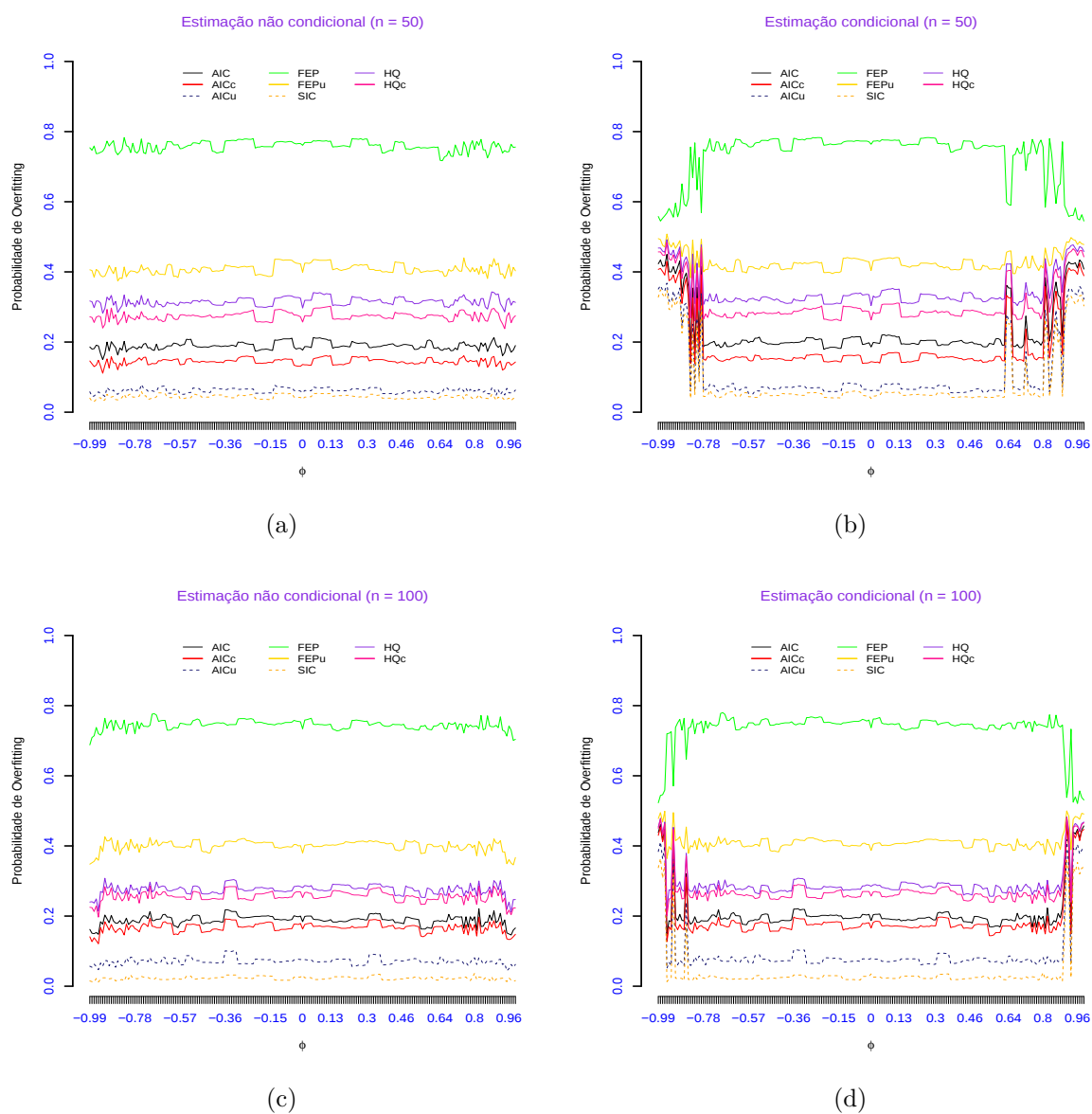
segundo estimação MVC, tem-se:

Figura 4.41: Cenário 2: Probabilidades de *overfitting* dos métodos de seleção usuais para amostras pequenas.



Fonte: Autoria própria.

Figura 4.42: Cenário 2: Probabilidades de *overfitting* dos métodos de seleção usuais para amostras grandes.



Fonte: Autoria própria.

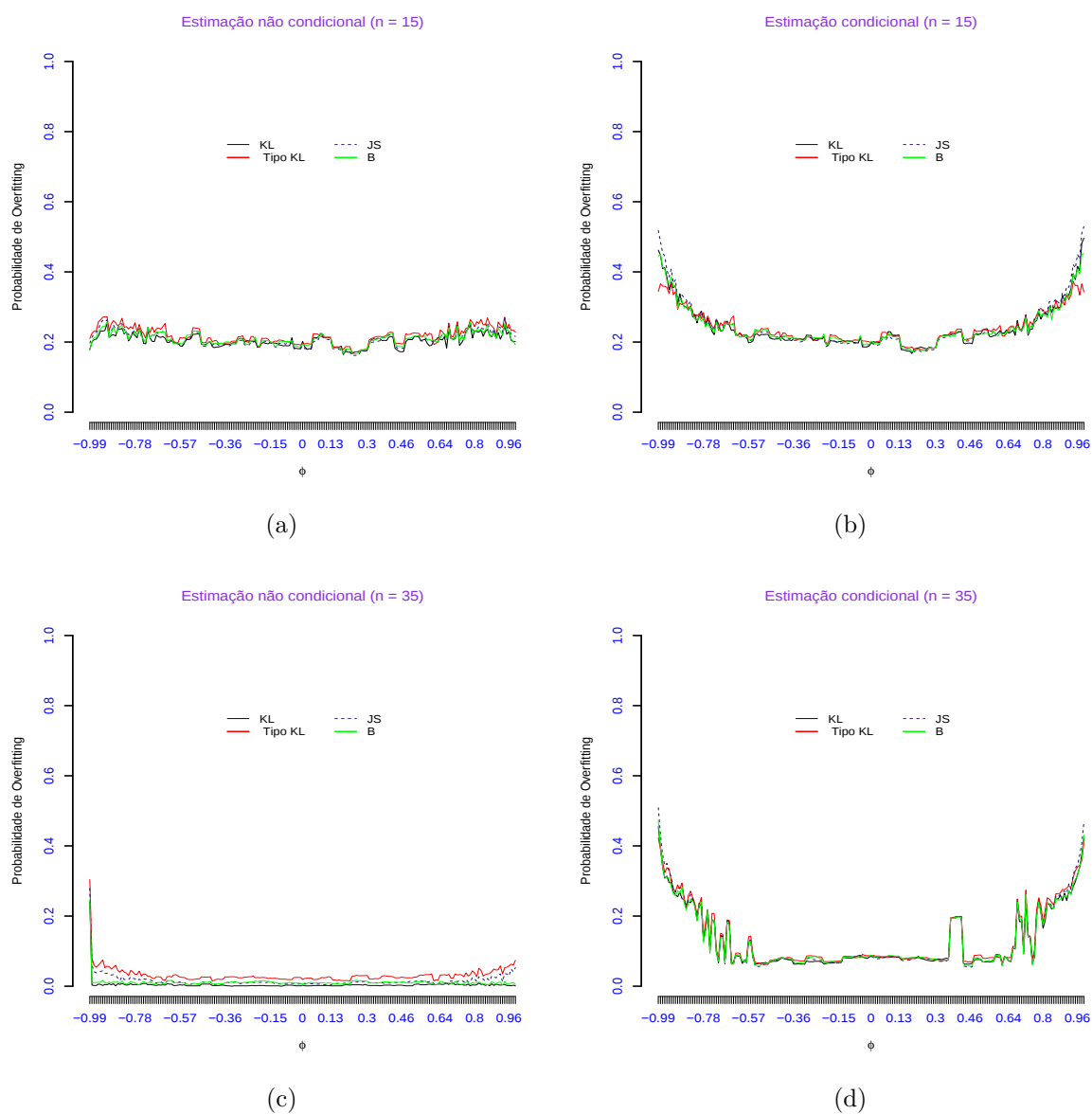
$$\left\{ \begin{array}{l}
\text{Para } -0.7 < \phi < 0.7 : \\
\Pr_{15;35}(\text{KL}) \approx \Pr_{15;35}(\text{B}) \approx \Pr_{15;35}(\text{JS}) \approx \Pr_{15;35}(\text{TKL}) \\
\Pr_{50;100}(\text{KL}) < \Pr_{50;100}(\text{JS}) < \Pr_{50;100}(\text{B}) < \Pr_{50;100}(\text{TKL}) \\
\text{Para } -0.9 < \phi < -0.7 \cap 0.7 < \phi < 0.9 : \\
\Pr_{15}(\text{TKL}) < \Pr_{15}(\text{B}) < \Pr_{15}(\text{JS}) < \Pr_{15}(\text{KL}) \\
\Pr_{35}(\text{TKL}) \approx \Pr_{35}(\text{B}) \approx \Pr_{35}(\text{JS}) \approx \Pr_{35}(\text{KL}) \\
\Pr_{50;100}(\text{TKL}) \approx \Pr_{50;100}(\text{B}) \approx \Pr_{50;100}(\text{JS}) \approx \Pr_{50;100}(\text{KL}) \\
\text{Para } |\phi| > 0.9 \\
\Pr_{15;35}(\text{B}) < \Pr_{15;35}(\text{TKL}) < \Pr_{15;35}(\text{KL}) < \Pr_{15;35}(\text{JS}) \\
\Pr_{50}(\text{B}) < \Pr_{50}(\text{TKL}) < \Pr_{50}(\text{JS}) \approx \Pr_{50}(\text{KL}) \\
\Pr_{100}(\text{TKL}) \approx \Pr_{100}(\text{B}) \approx \Pr_{100}(\text{JS}) \approx \Pr_{100}(\text{KL})
\end{array} \right.$$

Comparação entre os métodos clássicos e métodos novos

Para tamanho de amostra $n = 15$, o critério AICu teve o melhor desempenho em toda a região de estabilidade discretizada, considerando ambos os métodos de estimação (como podemos ver na Figura 4.45). Os novos critérios apresentaram menores taxas de *overfitting* em comparação com os critérios clássicos AIC e SIC para estimação MVNC. Com o uso da estimação MVC, os novos critérios pioram o desempenho para $|\phi| \uparrow 1$. A partir do tamanho de amostra $n = 35$, pode-se identificar as taxas de *overfitting* dos critérios considerados diminuem a medida que aumentamos o tamanho da amostra. Os métodos propostos combinados com a estimação MVNC apresentaram os melhores desempenhos. Na estimação por MVC, para valores de $|\phi|$ próximos de 1, os critérios apresentaram probabilidades de *overfitting* altas em relação a estimação MVNC. Mas, para valores de ϕ próximo a origem, essa probabilidade é pequena com destaque para os novos critérios e os critérios SIC e AICu. Portanto, os novos critérios se configuram como uma nova alternativa para a identificação de modelos em séries temporais, em especial a classe de modelos autoregressivos. No qual, a probabilidade de escolher o modelo *overfitting* diminui a medida que a ordem desse modelo se afasta da verdadeira ordem, além disso, a medida que aumentamos o tamanho das amostras essa probabilidade tende a decrescer. No geral, a estimação por Máxima Verossimilhança não condicional mostrou-se mais robusta para identificação de modelos quando existe forte correlação nos dados, isto é, para valores de $|\phi|$ próximos de 1.

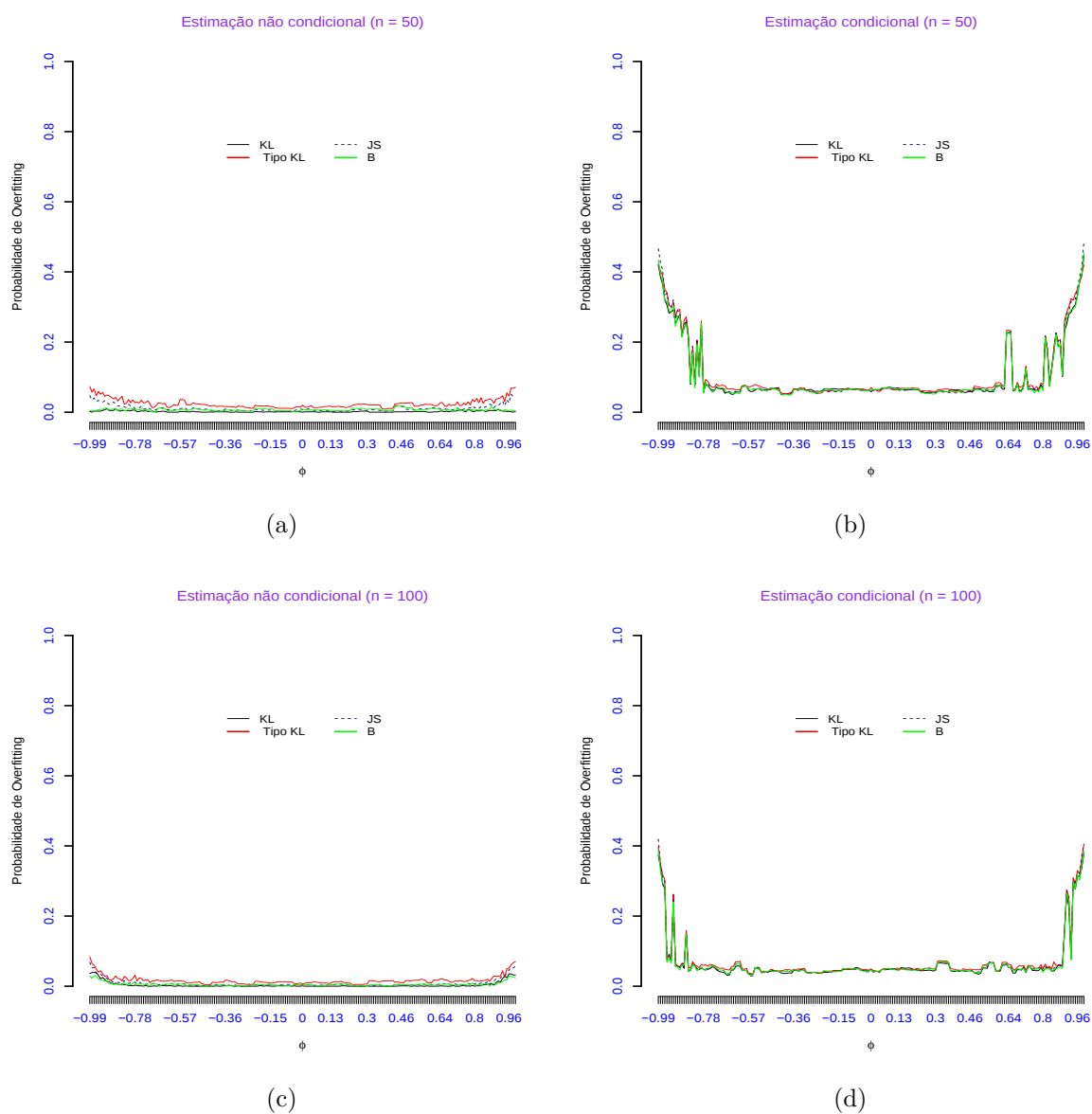
Alusivo aos resultados extraídos nas Figuras 4.41 e 4.42, a seguinte relação é encontrada:

Figura 4.43: Cenário 2: Probabilidades de *overfitting* dos métodos propostos induzidos das medidas de divergências para amostras pequenas.



Fonte: Autoria própria.

Figura 4.44: Cenário 2: Probabilidades de *overfitting* dos métodos propostos induzidos das medidas de divergências para amostras grandes.



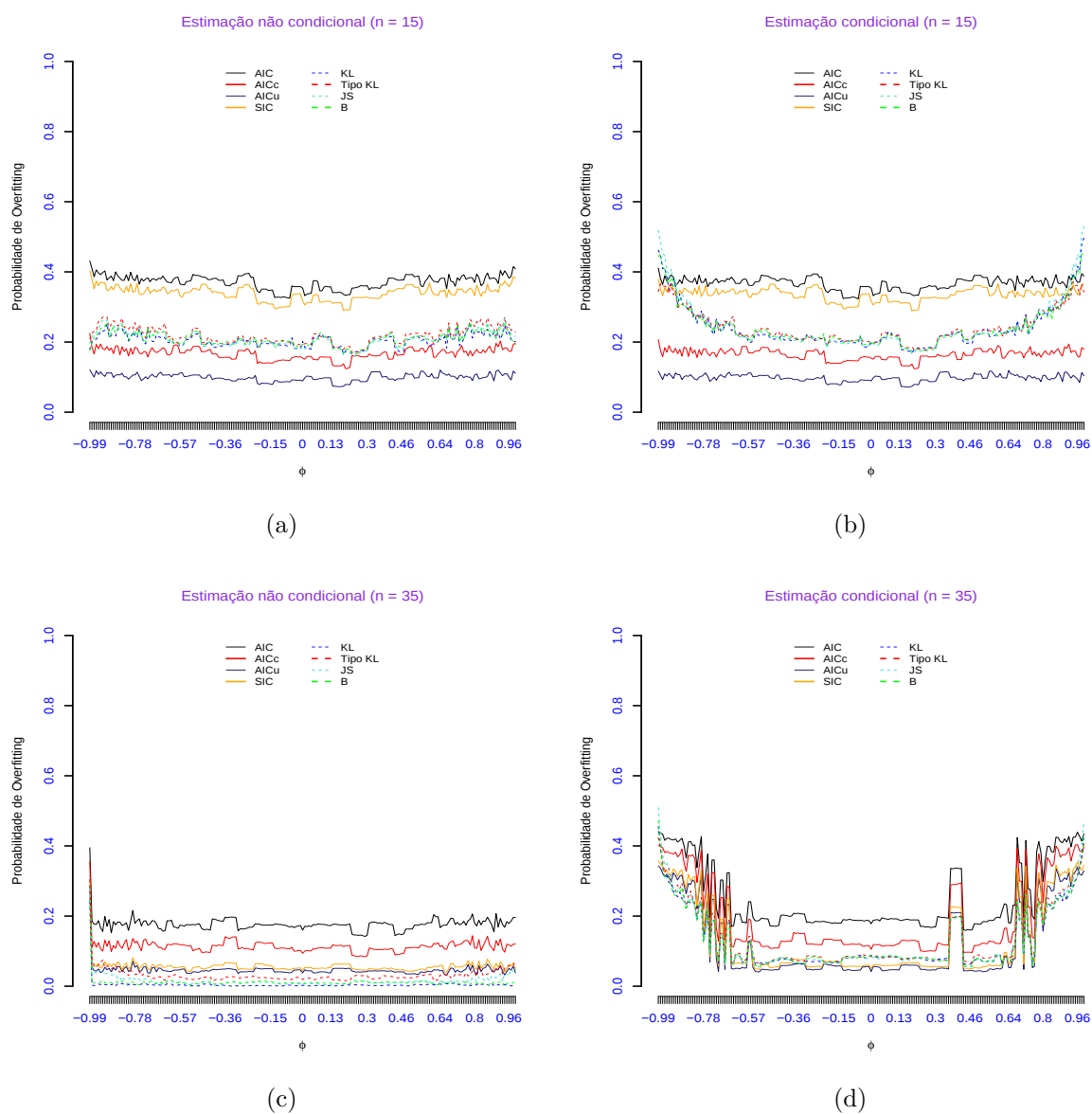
Fonte: Autoria própria.

$$\left\{ \begin{array}{l} \text{Para } : n = 15(MVNC) \\ \Pr_{15}(\text{AICu}) < \Pr_{15}(\text{AICc}) < \Pr_{15}(\text{KL}) < \Pr_{15}(\text{B}) \approx \Pr_{15}(\text{JS}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{AIC}) \\ \text{Para } : n = 25, 50, 100(MVNC) \\ \Pr_n(\text{KL}) < \Pr_n(\text{B}) \approx \Pr_n(\text{JS}) < \Pr_n(\text{TKL}) < \Pr_n(\text{SIC}) \approx \Pr_n(\text{AICu}) < \Pr_n(\text{AICc}) < \Pr_n(\text{AIC}) \end{array} \right.$$

com base na estimação condicional MVC, obtemos a desigualdade:

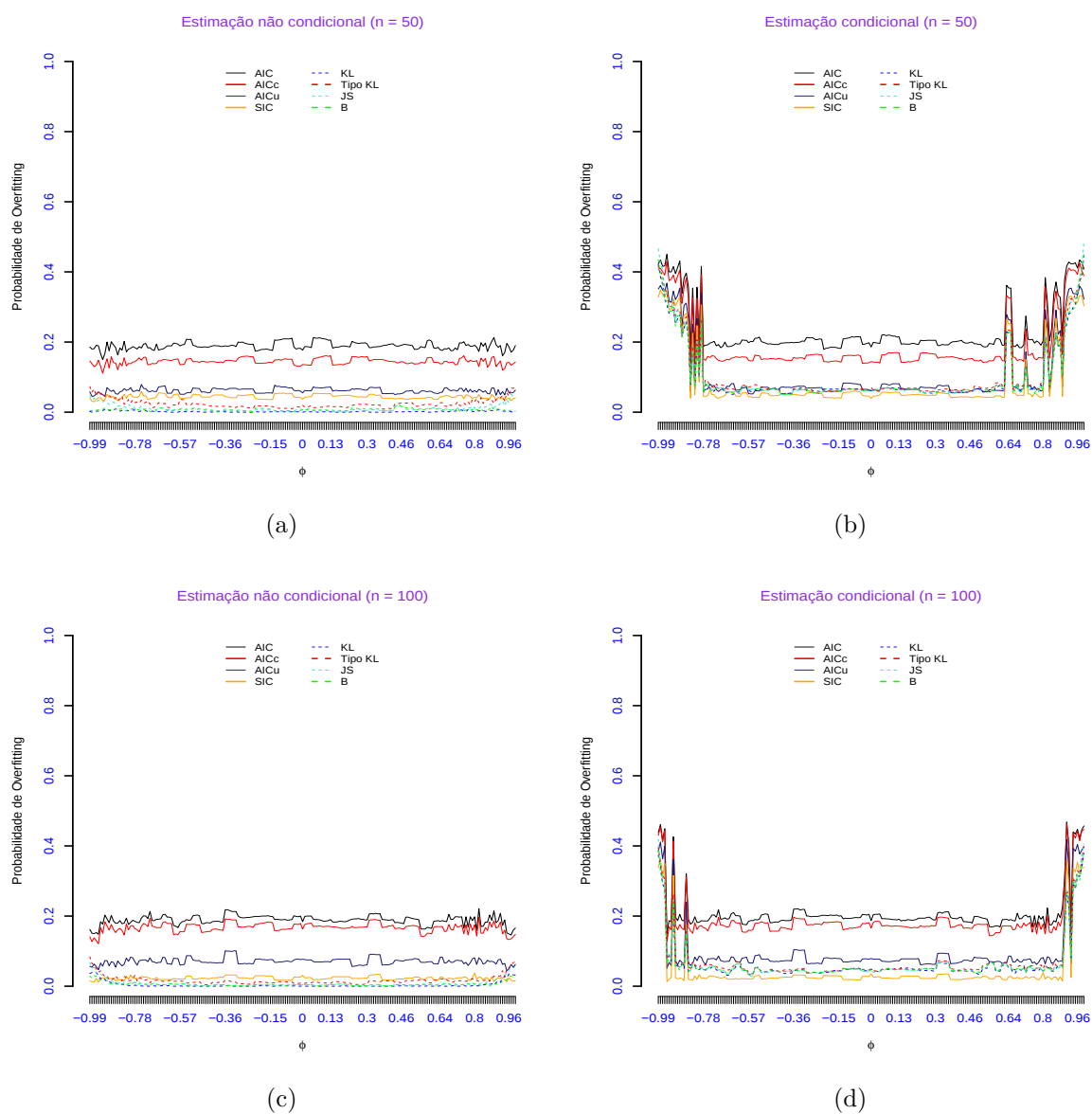
$$\left\{ \begin{array}{l} \text{Para } |\phi| < 0.9 : \\ \Pr_{15;35}(\text{AICu}) < \Pr_{15;35}(\text{AICc}) < \Pr_{15;35}(\text{KL}) \approx \Pr_{15;35}(\text{B}) \\ \approx \Pr_{15;35}(\text{JS}) < \Pr_{15;35}(\text{TKL}) < \Pr_{15;35}(\text{SIC}) < \Pr_{15;35}(\text{AIC}) \\ \text{Para } : |\phi| > 0.9 \\ \Pr_{15}(\text{AICu}) < \Pr_{15}(\text{AICc}) \approx \Pr_{15}(\text{SIC}) < \Pr_{15}(\text{TKL}) < \Pr_{15}(\text{AIC}) \approx \Pr_{15}(\text{B}) < \Pr_{15}(\text{KL}) < \Pr_{15}(\text{JS}) \\ \text{Para } -0.95 < \phi < -0.9 \cap 0.9 < \phi < 0.95 : \\ \Pr_{35}(\text{KL}) \approx \Pr_{35}(\text{B}) \approx \Pr_{35}(\text{JS}) < \Pr_{35}(\text{TKL}) < \Pr_{35}(\text{AICu}) < \Pr_{35}(\text{SIC}) < \Pr_{35}(\text{AICc}) < \Pr_{35}(\text{AIC}) \\ \text{Para } |\phi| < 1 : \\ \Pr_{50;100}(\text{KL}) \approx \Pr_{50;100}(\text{B}) \approx \Pr_{50;100}(\text{JS}) < \Pr_{50;100}(\text{TKL}) < \Pr_{50;100}(\text{AICu}) \\ < \Pr_{50;100}(\text{SIC}) < \Pr_{50;100}(\text{AICc}) < \Pr_{50;100}(\text{AIC}) \end{array} \right.$$

Figura 4.45: Cenário 2: Comparação entre os novos métodos propostos e os métodos clássicos de seleção de modelos para pequenas amostras



Fonte: Autoria própria.

Figura 4.46: Cenário 2: Comparação entre os novos métodos propostos e os métodos clássicos de seleção de modelos para grandes amostras



Fonte: Autoria própria.

Tabela 4.19: Cenário (1): Probabilidade de *overfitting* dos critérios clássicos de seleção de modelos na classe ARMA

n	ϕ	Condicional								Não condicional							
		AIC	AICc	AICu	FEP	FEPu	SIC	HQ	HQc	AIC	AICc	AICu	FEP	FEPu	SIC	HQ	HQc
15	-0.90	0.41	0.32	0.27	0.56	0.47	0.39	0.48	0.45	0.43	0.34	0.28	0.57	0.49	0.40	0.49	0.46
	-0.70	0.44	0.35	0.30	0.55	0.48	0.41	0.48	0.46	0.44	0.36	0.30	0.55	0.49	0.41	0.49	0.47
	-0.30	0.41	0.34	0.29	0.57	0.48	0.40	0.49	0.46	0.41	0.34	0.29	0.58	0.48	0.40	0.49	0.46
	-0.10	0.43	0.36	0.29	0.55	0.48	0.40	0.48	0.46	0.43	0.36	0.29	0.55	0.48	0.40	0.48	0.46
	0.10	0.42	0.34	0.30	0.54	0.48	0.40	0.49	0.46	0.42	0.34	0.30	0.55	0.48	0.40	0.49	0.46
	0.30	0.43	0.35	0.28	0.56	0.49	0.41	0.51	0.46	0.43	0.35	0.28	0.56	0.49	0.41	0.51	0.47
	0.70	0.42	0.34	0.27	0.55	0.48	0.40	0.49	0.46	0.43	0.34	0.28	0.56	0.49	0.41	0.50	0.46
	0.90	0.42	0.34	0.28	0.57	0.50	0.40	0.51	0.47	0.45	0.36	0.30	0.59	0.51	0.42	0.52	0.49
35	-0.90	0.45	0.44	0.40	0.54	0.49	0.40	0.49	0.48	0.46	0.44	0.40	0.55	0.50	0.41	0.49	0.49
	-0.70	0.45	0.43	0.39	0.53	0.48	0.39	0.48	0.47	0.45	0.43	0.39	0.53	0.48	0.39	0.48	0.47
	-0.30	0.32	0.29	0.24	0.59	0.41	0.24	0.38	0.36	0.32	0.29	0.24	0.60	0.41	0.24	0.38	0.36
	-0.10	0.20	0.16	0.09	0.66	0.38	0.09	0.33	0.29	0.20	0.16	0.09	0.66	0.38	0.09	0.33	0.29
	0.10	0.20	0.16	0.09	0.66	0.37	0.09	0.32	0.28	0.20	0.16	0.09	0.66	0.37	0.09	0.32	0.28
	0.30	0.31	0.28	0.24	0.58	0.41	0.24	0.37	0.35	0.31	0.28	0.24	0.58	0.41	0.24	0.37	0.35
	0.70	0.47	0.45	0.41	0.56	0.50	0.41	0.50	0.48	0.47	0.45	0.41	0.56	0.50	0.41	0.50	0.49
	0.90	0.46	0.44	0.40	0.53	0.49	0.40	0.48	0.47	0.46	0.45	0.40	0.54	0.49	0.41	0.48	0.47
50	-0.90	0.46	0.46	0.43	0.52	0.49	0.42	0.48	0.48	0.47	0.47	0.43	0.53	0.51	0.41	0.49	0.49
	-0.70	0.48	0.48	0.44	0.55	0.52	0.42	0.51	0.50	0.48	0.47	0.44	0.55	0.51	0.42	0.50	0.50
	-0.30	0.30	0.28	0.21	0.62	0.40	0.19	0.36	0.34	0.30	0.28	0.21	0.62	0.40	0.18	0.36	0.34
	-0.10	0.25	0.23	0.15	0.63	0.38	0.13	0.33	0.31	0.23	0.19	0.11	0.63	0.36	0.10	0.31	0.29
	0.10	0.25	0.23	0.15	0.64	0.38	0.13	0.32	0.31	0.25	0.23	0.15	0.63	0.38	0.13	0.32	0.30
	0.30	0.24	0.20	0.12	0.67	0.38	0.09	0.32	0.30	0.24	0.20	0.12	0.67	0.38	0.08	0.32	0.29
	0.70	0.46	0.45	0.42	0.52	0.49	0.40	0.48	0.47	0.43	0.42	0.39	0.53	0.47	0.37	0.46	0.45
	0.90	0.47	0.46	0.43	0.54	0.50	0.41	0.49	0.48	0.45	0.44	0.42	0.52	0.48	0.41	0.47	0.47
100	-0.90	0.46	0.46	0.44	0.52	0.49	0.41	0.48	0.47	0.47	0.47	0.45	0.51	0.49	0.42	0.48	0.48
	-0.70	0.40	0.39	0.33	0.57	0.47	0.29	0.42	0.41	0.39	0.38	0.32	0.56	0.45	0.28	0.41	0.40
	-0.30	0.22	0.21	0.12	0.65	0.38	0.06	0.29	0.28	0.22	0.20	0.12	0.65	0.38	0.06	0.28	0.28
	-0.10	0.21	0.19	0.11	0.64	0.36	0.06	0.28	0.26	0.21	0.19	0.11	0.63	0.36	0.06	0.28	0.26
	0.10	0.21	0.19	0.11	0.64	0.37	0.06	0.28	0.26	0.21	0.19	0.11	0.63	0.37	0.06	0.28	0.26
	0.30	0.22	0.20	0.12	0.64	0.38	0.05	0.29	0.28	0.22	0.20	0.12	0.64	0.38	0.05	0.28	0.28
	0.70	0.39	0.38	0.34	0.55	0.46	0.30	0.42	0.41	0.39	0.38	0.34	0.55	0.45	0.29	0.42	0.41
	0.90	0.48	0.47	0.45	0.52	0.50	0.43	0.49	0.48	0.47	0.47	0.44	0.52	0.50	0.41	0.49	0.49

Tabela 4.20: Cenário (2): Probabilidade de *overfitting* dos critérios clássicos de seleção de modelos na classe ARMA

n	ϕ	Condicional								Não condicional							
		AIC	AICc	AICu	FEP	FEPu	SIC	HQ	HQc	AIC	AICc	AICu	FEP	FEPu	SIC	HQ	HQc
15	-0.90	0.37	0.17	0.11	0.67	0.49	0.34	0.52	0.42	0.38	0.18	0.11	0.68	0.50	0.35	0.53	0.43
	-0.70	0.37	0.18	0.11	0.66	0.48	0.34	0.52	0.43	0.38	0.18	0.11	0.66	0.48	0.34	0.52	0.43
	-0.30	0.39	0.17	0.10	0.68	0.51	0.35	0.54	0.43	0.39	0.18	0.10	0.68	0.51	0.35	0.54	0.43
	-0.10	0.33	0.15	0.09	0.68	0.47	0.30	0.49	0.38	0.33	0.15	0.09	0.69	0.47	0.30	0.50	0.38
	0.10	0.36	0.16	0.10	0.69	0.49	0.32	0.51	0.41	0.36	0.16	0.10	0.69	0.49	0.32	0.52	0.41
	0.30	0.36	0.16	0.09	0.72	0.51	0.33	0.53	0.43	0.36	0.16	0.09	0.72	0.51	0.33	0.54	0.43
	0.70	0.37	0.17	0.10	0.64	0.48	0.33	0.51	0.43	0.37	0.17	0.10	0.65	0.48	0.34	0.52	0.43
	0.90	0.36	0.19	0.11	0.67	0.50	0.34	0.52	0.42	0.38	0.20	0.12	0.68	0.51	0.35	0.54	0.44
35	-0.90	0.41	0.37	0.30	0.59	0.49	0.31	0.47	0.44	0.19	0.13	0.05	0.77	0.41	0.06	0.35	0.30
	-0.70	0.30	0.25	0.18	0.68	0.45	0.19	0.42	0.36	0.19	0.13	0.05	0.76	0.41	0.06	0.35	0.27
	-0.30	0.18	0.12	0.05	0.77	0.40	0.06	0.35	0.27	0.16	0.11	0.04	0.76	0.38	0.05	0.32	0.25
	-0.10	0.19	0.12	0.05	0.76	0.41	0.06	0.33	0.28	0.18	0.11	0.04	0.76	0.40	0.05	0.32	0.27
	0.10	0.19	0.12	0.05	0.76	0.41	0.06	0.34	0.28	0.18	0.11	0.04	0.75	0.40	0.05	0.33	0.26
	0.30	0.17	0.10	0.05	0.78	0.40	0.06	0.33	0.27	0.14	0.09	0.04	0.77	0.38	0.05	0.31	0.25
	0.70	0.35	0.30	0.24	0.64	0.44	0.26	0.43	0.40	0.19	0.13	0.05	0.77	0.40	0.06	0.34	0.28
	0.90	0.42	0.37	0.31	0.57	0.49	0.33	0.48	0.46	0.21	0.13	0.05	0.76	0.42	0.06	0.35	0.29
50	-0.90	0.41	0.39	0.33	0.56	0.48	0.32	0.45	0.44	0.17	0.13	0.06	0.76	0.42	0.04	0.32	0.26
	-0.70	0.18	0.15	0.06	0.75	0.41	0.04	0.31	0.27	0.18	0.14	0.06	0.74	0.39	0.04	0.30	0.26
	-0.30	0.21	0.16	0.07	0.78	0.42	0.06	0.33	0.29	0.20	0.15	0.06	0.78	0.41	0.05	0.32	0.28
	-0.10	0.21	0.16	0.08	0.77	0.44	0.06	0.33	0.30	0.21	0.15	0.07	0.77	0.43	0.05	0.33	0.29
	0.10	0.22	0.17	0.08	0.78	0.44	0.06	0.35	0.31	0.21	0.16	0.07	0.77	0.43	0.05	0.34	0.30
	0.30	0.20	0.16	0.07	0.78	0.42	0.05	0.33	0.29	0.20	0.16	0.06	0.78	0.42	0.04	0.32	0.28
	0.70	0.18	0.15	0.06	0.74	0.41	0.05	0.32	0.27	0.18	0.14	0.06	0.73	0.40	0.04	0.30	0.26
	0.90	0.38	0.35	0.28	0.59	0.46	0.26	0.43	0.42	0.19	0.13	0.05	0.76	0.42	0.03	0.34	0.29
100	-0.90	0.19	0.16	0.06	0.75	0.42	0.02	0.28	0.27	0.18	0.16	0.06	0.75	0.41	0.02	0.28	0.26
	-0.70	0.19	0.17	0.07	0.78	0.41	0.02	0.27	0.26	0.19	0.17	0.07	0.78	0.40	0.02	0.27	0.25
	-0.30	0.20	0.17	0.06	0.77	0.42	0.02	0.28	0.26	0.20	0.17	0.06	0.76	0.42	0.02	0.28	0.25
	-0.10	0.19	0.17	0.07	0.75	0.40	0.02	0.27	0.26	0.19	0.16	0.07	0.75	0.39	0.02	0.27	0.25
	0.10	0.19	0.17	0.07	0.75	0.40	0.02	0.28	0.26	0.18	0.17	0.07	0.74	0.40	0.02	0.27	0.25
	0.30	0.19	0.17	0.06	0.76	0.42	0.02	0.28	0.26	0.19	0.17	0.06	0.76	0.41	0.02	0.28	0.26
	0.70	0.19	0.17	0.07	0.75	0.40	0.02	0.27	0.26	0.19	0.16	0.07	0.75	0.40	0.02	0.27	0.25
	0.90	0.31	0.29	0.21	0.64	0.44	0.16	0.37	0.36	0.18	0.17	0.06	0.74	0.41	0.02	0.29	0.28

Tabela 4.21: Cenário (1): Probabilidade de *overfitting* dos novos critérios de seleção de modelos induzidos das medidas de distância

n	ϕ	Condicional				Não condicional			
		KL	TKL	JS	B	KL	TKL	JS	B
15	-0.90	0.36	0.29	0.36	0.30	0.27	0.24	0.27	0.24
	-0.70	0.42	0.39	0.42	0.41	0.39	0.38	0.39	0.39
	-0.30	0.39	0.38	0.38	0.39	0.38	0.38	0.38	0.38
	-0.10	0.38	0.38	0.38	0.38	0.38	0.37	0.37	0.37
	0.10	0.36	0.36	0.36	0.37	0.35	0.36	0.35	0.35
	0.30	0.39	0.39	0.39	0.39	0.37	0.37	0.36	0.36
	0.70	0.41	0.38	0.39	0.39	0.40	0.37	0.38	0.38
	0.90	0.35	0.29	0.35	0.29	0.24	0.22	0.25	0.22
35	-0.90	0.21	0.19	0.19	0.19	0.18	0.17	0.16	0.17
	-0.70	0.39	0.37	0.37	0.37	0.38	0.35	0.36	0.36
	-0.30	0.30	0.31	0.30	0.30	0.21	0.25	0.23	0.23
	-0.10	0.20	0.23	0.20	0.20	0.02	0.11	0.06	0.07
	0.10	0.19	0.22	0.21	0.20	0.02	0.11	0.07	0.07
	0.30	0.32	0.32	0.32	0.32	0.21	0.27	0.24	0.24
	0.70	0.38	0.36	0.37	0.37	0.36	0.35	0.36	0.36
	0.90	0.24	0.21	0.23	0.21	0.20	0.19	0.19	0.18
50	-0.90	0.18	0.16	0.17	0.16	0.17	0.15	0.16	0.15
	-0.70	0.39	0.37	0.38	0.37	0.40	0.40	0.40	0.40
	-0.30	0.26	0.26	0.26	0.26	0.13	0.19	0.17	0.18
	-0.10	0.22	0.24	0.23	0.24	0.02	0.11	0.06	0.07
	0.10	0.21	0.23	0.22	0.22	0.05	0.13	0.09	0.09
	0.30	0.20	0.22	0.20	0.20	0.02	0.10	0.06	0.06
	0.70	0.38	0.39	0.38	0.38	0.36	0.35	0.35	0.36
	0.90	0.19	0.17	0.17	0.16	0.17	0.15	0.16	0.16
100	-0.90	0.15	0.13	0.14	0.14	0.14	0.14	0.15	0.15
	-0.70	0.29	0.30	0.30	0.29	0.24	0.26	0.25	0.24
	-0.30	0.18	0.19	0.18	0.18	0.01	0.06	0.04	0.04
	-0.10	0.15	0.18	0.16	0.17	0.01	0.07	0.03	0.03
	0.10	0.14	0.18	0.17	0.17	0.01	0.08	0.04	0.04
	0.30	0.17	0.19	0.18	0.18	0.00	0.08	0.04	0.04
	0.70	0.32	0.32	0.31	0.31	0.27	0.29	0.28	0.28
	0.90	0.13	0.13	0.13	0.13	0.12	0.12	0.12	0.12

Tabela 4.22: Cenário (2): Probabilidade de *overfitting* dos novos critérios de seleção de modelos induzidos das medidas de distância

n	ϕ	Condicional				Não condicional			
		KL	TKL	JS	B	KL	TKL	JS	B
15	-0.90	0.31	0.29	0.32	0.30	0.21	0.23	0.22	0.21
	-0.70	0.23	0.26	0.24	0.23	0.20	0.24	0.22	0.22
	-0.30	0.22	0.21	0.21	0.22	0.21	0.21	0.20	0.21
	-0.10	0.20	0.20	0.20	0.20	0.19	0.20	0.20	0.20
	0.10	0.22	0.22	0.21	0.21	0.21	0.21	0.21	0.22
	0.30	0.18	0.18	0.18	0.18	0.18	0.18	0.17	0.18
	0.70	0.24	0.23	0.23	0.23	0.21	0.21	0.21	0.21
	0.90	0.32	0.30	0.33	0.32	0.21	0.22	0.21	0.21
35	-0.90	0.26	0.29	0.28	0.27	0.01	0.06	0.03	0.01
	-0.70	0.14	0.15	0.14	0.14	0.00	0.02	0.01	0.01
	-0.30	0.07	0.09	0.08	0.08	0.00	0.02	0.01	0.01
	-0.10	0.08	0.08	0.08	0.08	0.00	0.02	0.01	0.01
	0.10	0.08	0.08	0.08	0.09	0.00	0.02	0.00	0.01
	0.30	0.07	0.08	0.08	0.08	0.00	0.03	0.01	0.01
	0.70	0.18	0.20	0.20	0.19	0.01	0.02	0.01	0.01
	0.90	0.26	0.28	0.27	0.26	0.01	0.05	0.02	0.01
50	-0.90	0.27	0.29	0.29	0.28	0.00	0.05	0.02	0.01
	-0.70	0.07	0.08	0.07	0.07	0.00	0.01	0.01	0.00
	-0.30	0.06	0.06	0.06	0.06	0.00	0.02	0.01	0.00
	-0.10	0.07	0.07	0.06	0.06	0.00	0.01	0.00	0.00
	0.10	0.07	0.06	0.06	0.07	0.00	0.02	0.01	0.01
	0.30	0.06	0.06	0.05	0.05	0.00	0.02	0.00	0.01
	0.70	0.06	0.07	0.06	0.06	0.00	0.02	0.01	0.00
	0.90	0.23	0.26	0.25	0.24	0.00	0.04	0.02	0.01
100	-0.90	0.06	0.06	0.06	0.05	0.01	0.02	0.02	0.01
	-0.70	0.04	0.05	0.05	0.05	0.00	0.01	0.01	0.01
	-0.30	0.04	0.04	0.04	0.04	0.00	0.01	0.00	0.00
	-0.10	0.05	0.05	0.05	0.05	0.00	0.01	0.00	0.00
	0.10	0.05	0.05	0.05	0.05	0.00	0.01	0.00	0.00
	0.30	0.05	0.05	0.05	0.05	0.00	0.01	0.00	0.00
	0.70	0.05	0.06	0.05	0.05	0.00	0.01	0.00	0.00
	0.90	0.16	0.16	0.16	0.15	0.01	0.03	0.02	0.00

4.4 Conclusão dos estudos computacionais

Nesta Seção apresenta-se os principais resultados obtidos neste capítulo. Na Seção 4.1, (i) referente ao estudo comparativo entre os três métodos de estimação clássicos em termos do viés e EQM e sobre a densidade conjunta de processo ARMA gaussiano via medidas de divergência, pode-se concluir que: (i) Para os processos AR(1), MA(1) e ARMA(1,1), em relação ao Erro Quadrático Médio das estimativas dos parâmetros, tem-se que os métodos MQNC e MVNC são melhores quando o processo se aproxima da não estabilidade, além disso para região onde $\phi \downarrow 0$ o método MVNC é melhor para tamanhos de amostras pequenos, (ii) o método de MVNC é o mais indicado entre termos de influência da estimação sobre a densidade conjunta de processos Autorregressivos de Médias Móveis gaussiano, especificamente, AR(1), MA(1) e ARMA(1,1) via medidas de divergência, (iii) na estimação dos parâmetros do processo de Médias Móveis, identifica-se dois comportamentos incomuns. O primeiro, para a estimação do parâmetro θ nota-se que para uma determinada região onde $0.7 < \theta < 0.9 \cap -0.9 < \theta < -0.7$, o viés da estimação não diminui com o tamanho da amostra. O segundo, em relação a estimação de σ^2 , quando $|\theta| \uparrow 1$ a estimação de σ^2 torna-se altamente viesada. Como sugestão para trabalho futuro, segure-se uma investigação mais profunda nesses aspectos e (iv) as divergências na forma reduzida apresentaram maior poder na capacidade de identificar dissimilaridades entre os métodos de estimação em relação a forma completa. Este fato, pode ser de suma importância na escolha do método de estimação dos parâmetros.

Na Seção 4.2, sobre a proposta e discussão de cinco testes de hipóteses baseados nas medidas de divergências para identificação em modelos ARMA e o estudo comparativo entre testes da razão entre verossimilhanças condicionais e não condicionais, tem-se que: (i) Com base na função poder empírica, tem-se que o teste da RVNC tem melhor desempenho em comparação a RVC para tamanho de amostra pequeno ou moderado, quando os processos aproximam-se da não estabilidade, (ii) Para grandes amostras, ambos os testes em geral são assintoticamente equivalentes, como esperado, (iii) Considerando o método de estimação MVNC, os poderes empíricos dos testes baseados nas medidas são maiores em relação ao método de estimação MVC. (iv) Para os processos AR(1) e ARMA(1,1) quando os processos são próximos da não estabilidade combinado com a estimação MVNC, os testes baseados nas medidas Δ_{KL} , Δ_{JS} e Δ_B apresentam taxas maiores ou equivalentes ao testes clássicos da RVNC e RVC. Entretanto, considerando o processo MA(1), os testes clássicos tiveram os melhores desempenhos. Vale, salientar que, a estimação do parâmetro θ apresenta alto viés em determinada região, o que pode influenciar diretamente no resultado dos testes empíricos.

Por fim, na Seção 4.3 foi apresentado as novas propostas para critérios de seleção

de modelos baseados nas distâncias Δ_{KL} , Δ_{TKL} , Δ_{JS} , Δ_B e Δ_H em comparação aos critérios clássicos (AIC, AICc, AICu, HQ, HQu, FEP, FEPu, SIC). Assim, as novas propostas apresentaram as menores taxas de *Overfitting* para um processo Autoregressivo próximo de um passeio aleatório, quando o tamanho da amostra é grande. Além disso, a EMVNC diminui a probabilidade de *overfitting* dos critérios de seleção, principalmente na região de estabilidade próxima a origem ($|\phi| \downarrow 0$). Deste modo, a estimação por Máxima Verossimilhança Não Condicional, fornece maior robustez para as novas propostas baseadas nas medidas de distância para seleção de modelos.

Capítulo 5

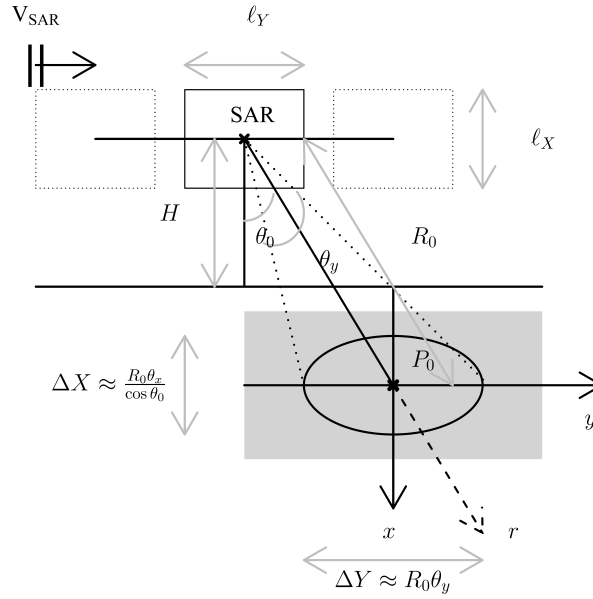
Classificadores baseados em um modelo ARMA espacial para dados PolSAR

5.1 Radar de abertura sintética e modelo espacial para imagem SAR

Radar de Abertura Sintética (*Synthetic Aperture Radar-SAR*) tem sido apresentado como uma poderosa ferramenta de sensoriamento remoto (LEE; POTTIER, 2009). A procura pelo SAR pode ser justificada basicamente por dois motivos: (1) sua capacidade de produzir imagens em alta resolução espacial e (2) sua habilidade de operar independentemente da luminosidade e das condições atmosféricas. Entretanto, as imagens SAR são fortemente afetadas pelo ruído *speckle*, o que dificulta sua análise e interpretação. Assim as imagens SAR carecem de processamento especializado. Nesta aplicação, partimos do suposto que a inclusão de correlação no modelo combinada ao uso de medidas de divergências pode gerar metodologias eficientes no tratamento de tais imagens.

As imagens SAR são produzidas obedecendo à seguinte dinâmica (explicada por meio da Figura 5.1): Um radar é acoplado à uma plataforma aerotransportada, V_{SAR} , a uma altura H e emite pulsos a um ângulo de incidência θ_0 . A abertura física do radar (a saber, largura ℓ_X e comprimento ℓ_Y) produz um feixe elíptico com centro em P_0 e eixos com tamanhos ΔX e ΔY . A área imagem da é definida em termos dos ângulos de abertura θ_x e θ_y .

Os pulsos emitidos são polarizado nas direções vertical “V” e horizontal “H”. Ao final, cada pixel é associado a três valores complexos, representando os canais de polarização HH, HV $\approx$ VH e VV. A frente que trabalha com todas as informações dos canais é chamada de imagens PolSAR (*Polarimetric Synthetic Aperture Radar-PolSAR*). Assim, a entrada

Figura 5.1: Geometria de obtenção da imagem SAR

FONTE: Nascimento (2012)

(i, j) de uma imagem PolSAR pode ser representada por

$$z(i, j) = [Z_{HH}(i, j), Z_{HV}(i, j), Z_{VV}(i, j)]^T \in \mathbb{C}^3.$$

Um dado desta natureza é chamado de *retorno singelook*. No sentido univariado, a inferência e modelagem de dados SAR costuma ser feita sobre as *intensidades* relacionadas aos canais (NASCIMENTO; CINTRA; FRERY, 2010),

$$I_{\mathcal{M}}(i, j) = \|Z_{\mathcal{M}}(i, j)\|^2 = Z_{\mathcal{M}}(i, j)^* Z_{\mathcal{M}}(i, j) \text{ para } \mathcal{M} \in \{HH, HV, VV\},$$

em que $\|z\| = \sqrt{z^* z}$ é a norma de um argumento complexo. Nesta aplicações, o mecanismo estocástico visa descrever intensidades.

A grande maioria dos métodos tem trabalhado sobre a suposição de independência espacial entre os pixels de uma amostra retirada. Entretanto, recentemente, Bustos, Ojeda e Vallejos (2009) apresentaram um trabalho de revisão definindo modelos ARMA espaciais para filtragem e estimação robusta de imagens SAR. Conforme (YAO; BROCKWELL, 2006), este modelo pode ser definido como

$$Z(s) = \sum_{i \in \mathcal{I}_1} \phi_i Z(s - i) + \sum_{i \in \mathcal{I}_2} \theta_i \epsilon(s - i) + \epsilon(s) \Leftrightarrow \Phi(B) Z(s) = \Theta(B) \epsilon(s), \quad (5.1)$$

em que $B^i \mathbf{X}(s) \triangleq B_1^i B_2^k \mathbf{X}(u, v) = \mathbf{X}(u - j, v - k) = \mathbf{X}(s - i)$, $\Phi(z) \triangleq 1 - \sum_{i \in \mathcal{I}_1} \phi_i z^i$, e $\Theta(z) \triangleq 1 + \sum_{i \in \mathcal{I}_2} \theta_i z^i$. Nesta dissertação, consideraremos o seguinte submodelo ARMA Espacial de (5.1) para a intensidade de um canal $\mathcal{M} \in \{HH, HV, VV\}$ associada à entrada (i, j) , $I_{\mathcal{M}}(i, j)$:

$$I_{\mathcal{M}}(i, j) = \sum_{k=1}^p \phi_k I_{\mathcal{M}}(i, j - k) + \epsilon(i, j) + \sum_{k=1}^q \phi_k \epsilon(i, j - k). \quad (5.2)$$

Embora seja uma condição forte, adotaremos a suposição gaussiana a fim de que a incorporação de correlação no modelo seja analiticamente tratável.

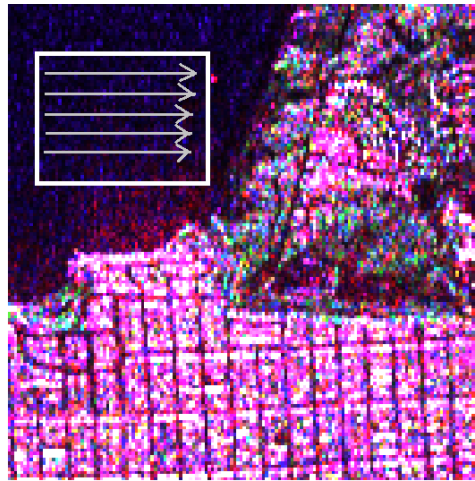
5.2 Esquema de classificação de imagem SAR

Neste ponto, as medidas de identificação discutidas na Seção 4.2 são adaptadas para resolver problemas de classificação em imagens SAR, cujas intensidades associadas a cada *pixel* seguem o modelo (5.2). Esta aplicação é pioneira no sentido de

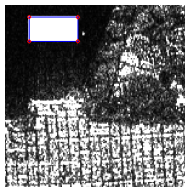
Classificar assumindo uma estrutura de correlação espacial e com base em distâncias estocásticas (a saber, S_{KL} , S_{TKL} , S_{JS} , S_B e S_H).

Dutra (1989) também considerou a distância Δ_{JS} , mas a medida não foi utilizada como uma estatística de teste. Esta aplicação também considera esta medida. Os resultados foram comparados com aqueles dos testes das Razões entre verossimilhanças condicional e não condicional S_{RVC} e S_{RVNC} , respectivamente. O algoritmo 2 apresenta os passos para realização da classificação em imagens SAR. Note que o primeiro passo, supõe que um bom método de segmentação (ou agrupamento) já foi utilizado, como é de costume em problemas de classificação.

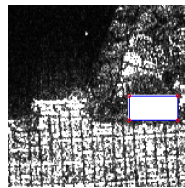
Figura 5.2: Imagem real PolSAR



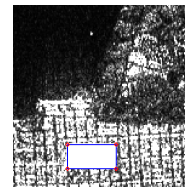
(a) Representação RGB



(b) Região selecionada de oceano



(c) Região selecionada de floresta



(d) Região selecionada de cidades

FONTE: Autoria Própria.

Algoritmo 2: Algoritmo para identificação e *performance*.**Identificação:** Fase de treinamento

Passo 1: Assuma que uma imagem SAR apresenta g grupos (valores de intensidades com as mesmas propriedades estatísticas) distintos. Selecione uma região retangular de tamanho N_k , para $k = 1, 2, \dots, g$, de cada grupo e concatene suas linhas, resultando em $\{X_t^{(k)}; t = 1, \dots, N_k\}$ para $k = 1, 2, \dots, g$ (DUTRA, 1989, para outro uso de concatenação de linhas na definição de séries).

Passo 2: Em caso de não normalidade, defina $X_t^{(k)}$ como a transformação Box-Cox (BROCKWELL; DAVIS, 2001, p. 284) mas próxima na normalidade. Como toda a abordagem foi desenvolvida para processos com média zero, retire a locação de cada série, $\dot{X}_t^{(k)} = X_t^{(k)} - N_k^{-1} \sum_{t=1}^{N_k} X_t^{(k)}$. Identifique e ajuste modelos ARMA para $\dot{X}_t^{(k)}$:

1. Utilizando o AIC e/ou as funções de autocorrelação e sua versão parcial, determine o submodelo ARMA para $\dot{X}_t^{(k)}$;
2. Realizar as estimações preliminares (usada como chute inicial) e estimar via (MVC; MVNC) para cada modelo selecionado no item anterior. Adicionalmente, procurar coeficientes não significativos via intervalos de confiança assintótico;
3. Ratificar a escolha através de análise diagnóstico, verificando se os resíduos resultantes são ruídos branco.

Para cada $\dot{X}_t^{(k)}$ defina a hipótese de treinamento $\mathcal{H}_{0,k}: \beta = \beta_k$ para $k = 1, \dots, g$, em que β_k é o vetor de coeficientes ARMA obtido para o modelo após realização dos itens 1.–3..

Performance: Fase de teste

Passo 3: Particione $\{\dot{X}_t^{(k)}; t = 1, \dots, N_k\}$ em $B_k = \lfloor N_k/n \rfloor$ sub séries de tamanho $n = 25, 35, 50$, diga-se $\{\dot{X}_t^{(k:v:n)}; t = 1, \dots, n\}$ para $v = 1, \dots, B_k$, em que $\lfloor \cdot \rfloor$ representa o maior inteiro menor ou igual o argumento. Defina $\hat{\beta}_{k:v:n}$ como o estimador de MVC ou MVNC para β_k como base na v -ésima série de tamanho n obtida de $\dot{X}_t^{(k)}$.

Passo 4: Obtenha para cada série a quantificação do erro do tipo I para $\mathcal{H}_{0,k}: \beta = \beta_k$ (ou taxa de alarme falso da classe k) dada por

$$\hat{\alpha}_{\mathcal{M},k,n} = \frac{\sum_{v=1}^{B_k} \mathbb{I}(S_{\mathcal{M}}(\hat{\beta}_{k:v:n}, \beta_k) > \chi_{p+q,\alpha}^2)}{B_k},$$

em que $\mathbb{I}(A)$ é zero para A como uma condição verdadeira e zero caso contrário. Por outro lado, o poder do teste para $\mathcal{H}_{0,k_1}: \beta = \beta_{k_1} \times \mathcal{H}_{1,k}: \beta = \beta_{k_2}$ (ou a probabilidade de detecção entre as classe k_1 e k_2) é dado por

$$1 - \hat{\beta}_{\mathcal{M},k_1,n} = \frac{\sum_{v=1}^{B_{k_2}} \mathbb{I}(S_{\mathcal{M}}(\hat{\beta}_{k_2:v:n}, \beta_{k_1}) > \chi_{p+q,\alpha}^2)}{B_{k_2}}.$$

5.3 Aplicação a imagem AIRSAR de São Francisco

Na Figura 5.3(a), apresenta-se uma imagem de 150×150 *pixels* obtida pelo sensor AIRSAR sob um *número de looks* quatro. O *número de looks* L pode ser entendido como um parâmetro de entrada o qual reflete que um *pixel* é o resultado da média de L mensurações de uma imagem sob estudo. Este parâmetro impõe um controle do efeito do

ruído speckle sobre a imagem SAR. Na Figura 5.3(a), vê-se também os sentidos em que as linhas serão concatenadas a formar séries do tipo (5.2). Note ainda que existem três regiões bem definidas: A saber, áreas de oceano, de floresta e de cidades. Desta forma, tem-se para cada canal de polarização amostras de treinamento de tamanho $N_k = N = 861$ representando as populações de Oceano ($k = 1$), Floresta ($k = 2$) e Cidades ($k = 3$), ilustradas nas Figuras 5.3(b)–5.3(d), respectivamente.

O processo de identificação foi feito conforme o Algoritmo 2. Preliminarmente ao processo de identificação, aplicamos os testes usuais de Shapiro e Wilk (1965) e Lilliefors (1967) para verificar a suposição de normalidade nas séries temporais, visto que esta dissertação é focada em processos gaussianos. Na Tabela 5.1, apresentam-se os resultados dos testes de *Shapiro-Wilker* (SW) e *Lilliefors* para séries obtidas de acordo com Passo 1.

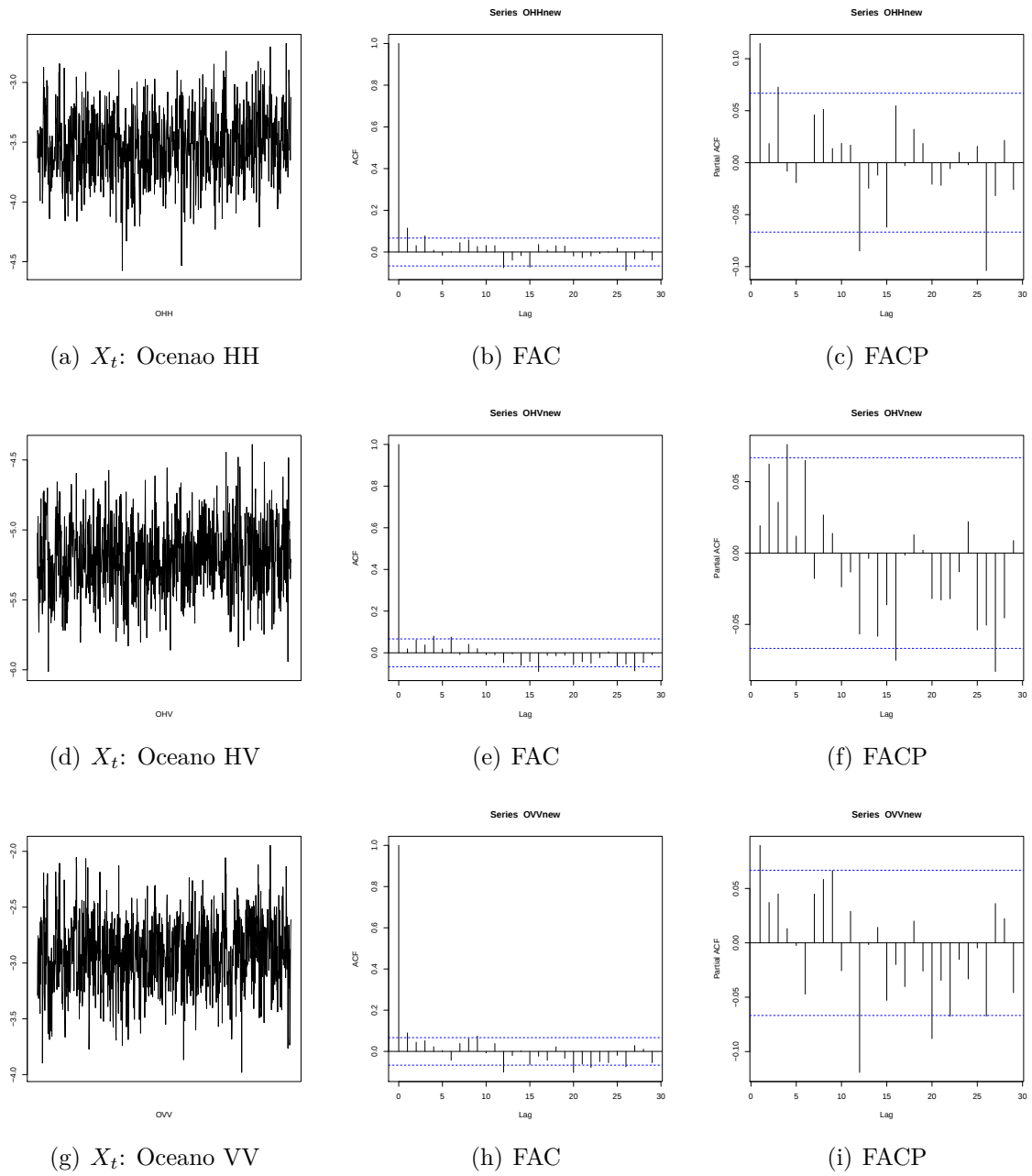
Tabela 5.1: *P*-valores dos testes de normalidade nos dados PolSAR.

(Região, Canal)	Dados originais		λ	Dados transformados	
	Shapiro-Wilk	Lilliefors		Shapiro-Wilk	Lilliefors
(Floresta, HH)	2.04×10^{-30}	4.66×10^{-30}	0.15	0.0315	0.0074
(Floresta, HV)	6.19×10^{-30}	3.36×10^{-58}	0.15	0.0002	0.0489
(Floresta, VV)	2.02×10^{-31}	2.46×10^{-51}	0.05	0.6203	0.5668
(Oceano, HH)	1.50×10^{-23}	2.11×10^{-24}	0.15	0.9635	0.8223
(Oceano, HV)	3.06×10^{-24}	2.52×10^{-21}	0.10	0.6607	0.3423
(Oceano, VV)	2.83×10^{-23}	1.89×10^{-25}	0.15	0.6484	0.2048
(Cidade, HH)	5.20×10^{-49}	8.32×10^{-273}	-0.15	0.6553	0.1779
(Cidade, HV)	5.51×10^{-48}	3.23×10^{-308}	-0.10	0.2351	0.1344
(Cidade, VV)	1.55×10^{-47}	1.46×10^{-269}	-0.10	0.0045	0.0148

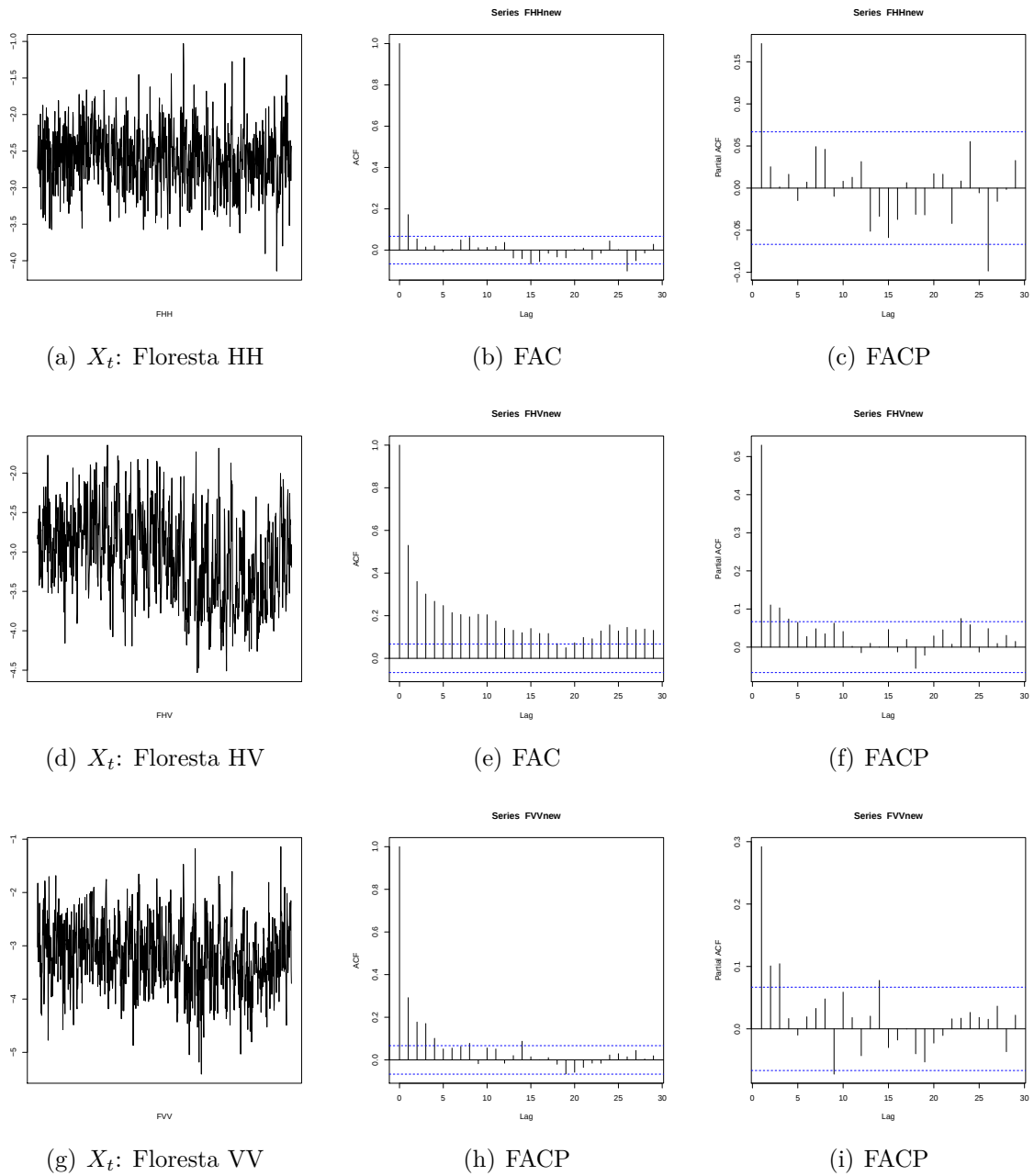
Como esperado, identificou-se que os processos são fortemente não gaussianos, pois os *p*-valores dos testes são menores do que o nível de nominal adotado, 5%. As séries então foram submetidas à transformação de Box-Cox (Box e Cox, 1964), a fim de normalizar os resíduos resultantes, pós processo de ajuste.

Nas Figuras 5.3, 5.4 e 5.5, apresentam-se as séries transformadas e suas respectivas funções de autocorrelação (FAC) e autocorrelação parcial (FACP).

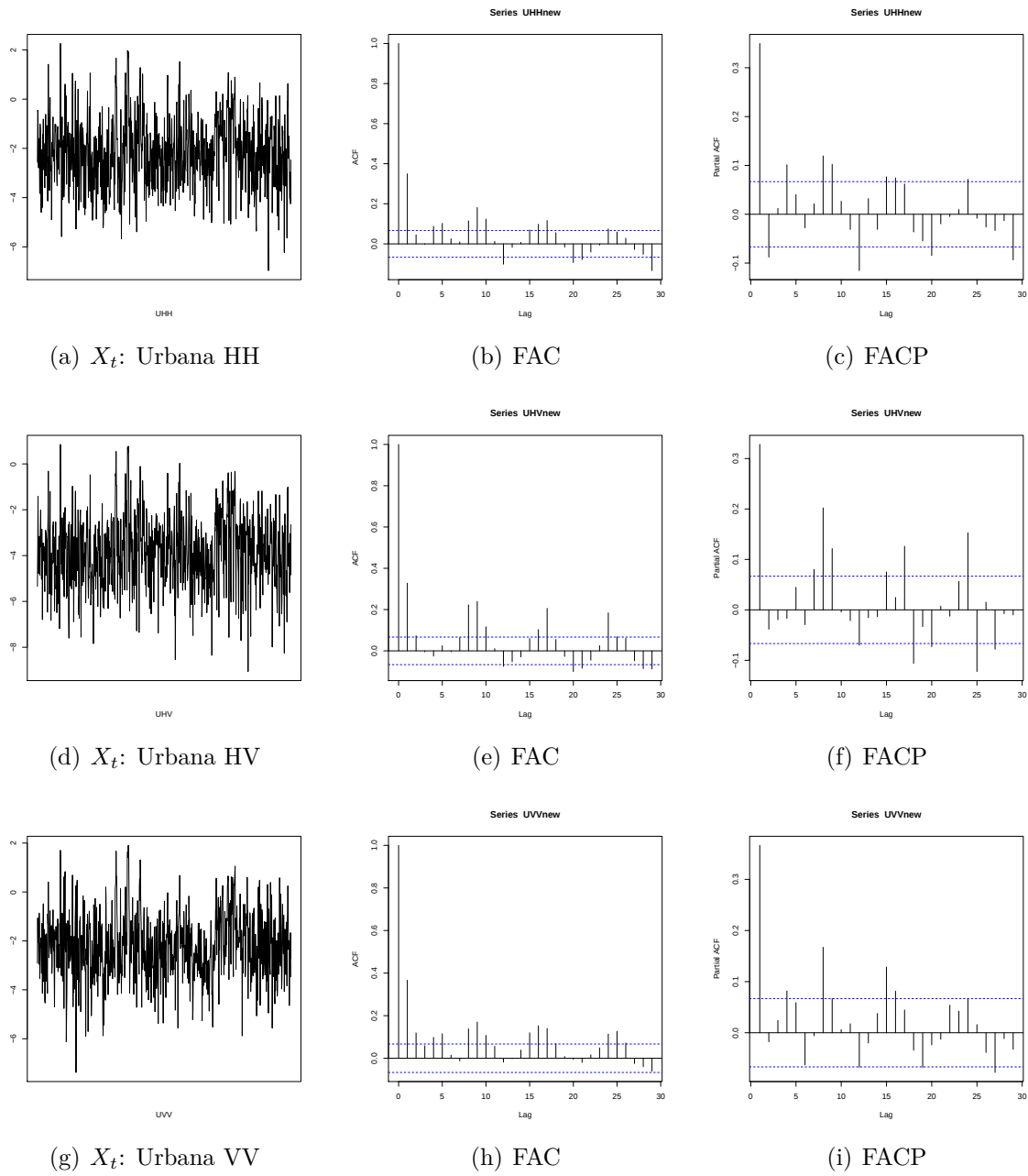
Com base nas FAC e FACP das séries combinadas com o uso de critérios de seleção, parece razoável sugerir os seguintes modelos ARMA para as sequencias de floresta e oceano: (1) Autoregressivos e (2) Ruído Branco. Para a região urbana na Figura 5.5, pôde-se identificar segundo FAC e FACP que para as séries transformadas possui um comportamento sazonal. Como o ferramental desta dissertação não foi desenvolvido nem para modelos ARMA sazonal nem para modelos ARMA não gaussianos, tínhamos que decidir entre trabalhar ou sob não normalidade (a melhor transformação de Box-Cox normalizava as séries, mas impõe a elas um comportamento sazonal) ou sob a sazonalidade. Um estudo piloto prévio mostrou que trabalhar com sazonalidade normal é mais adequado. Para este último caso,

Figura 5.3: Séries transformadas e suas FAC e FACP para região de oceano.

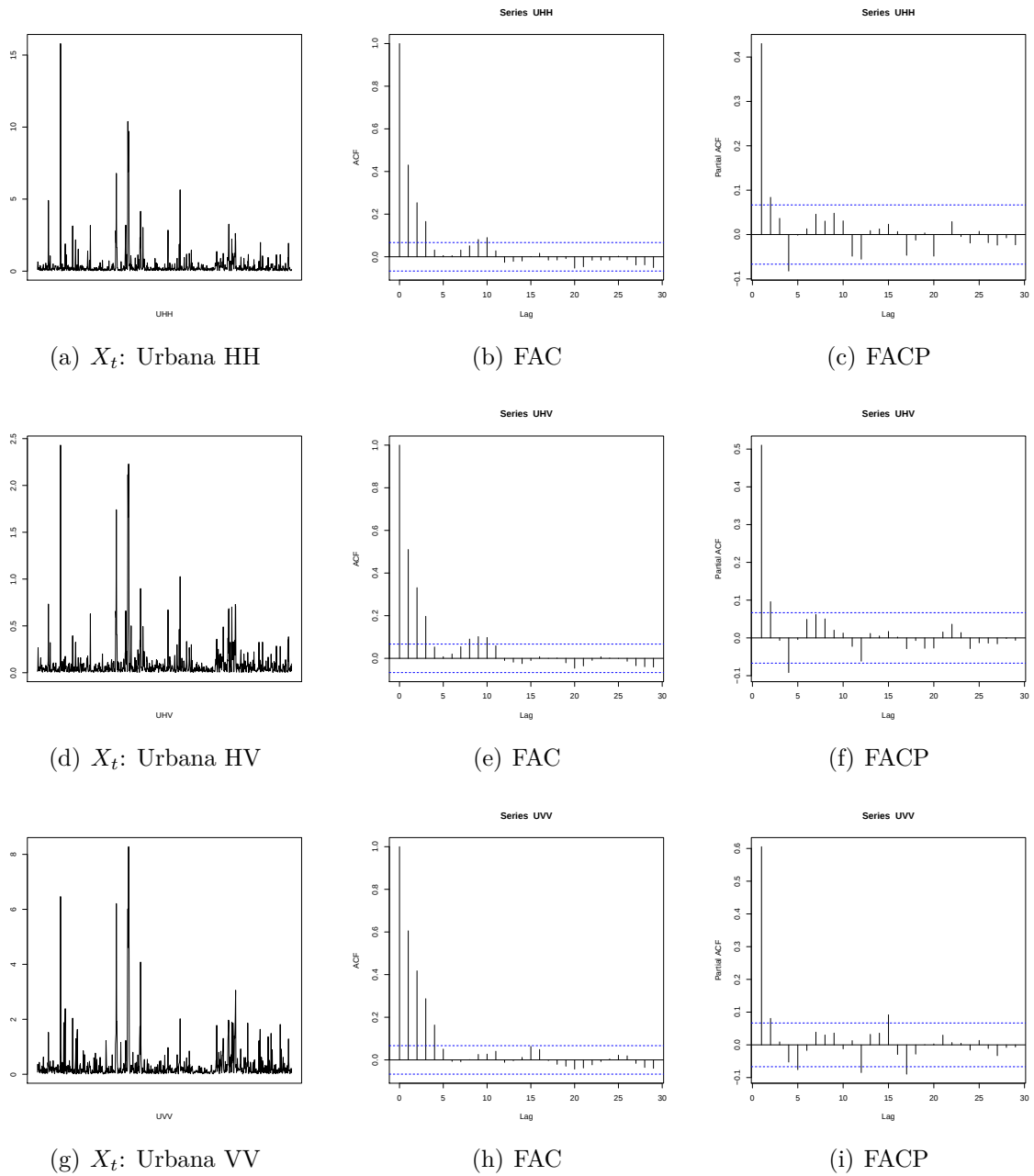
FONTE: Autoria própria

Figura 5.4: Séries transformadas e suas FAC e FACP para região de floresta.

FONTE: Autoria própria

Figura 5.5: Séries transformadas e suas FAC e FACP para região de áreas urbanas.

FONTE: Autoria própria

Figura 5.6: Séries originais e suas FAC e FACP para região de áreas urbanas.

FONTE: Autoria própria

objetiva-se responder à questão “Qual técnica é menos influenciada pela imposição da sazonalidade?”

Após a etapa de identificação baseada na FAC e FACP das séries e no critério AIC, os modelos foram ajustados. Os modelos identificados para (região,canal) foram

- AR(1) para: (Floresta,HH), (Oceano,HH), (Oceano,VV), (Cidade,HV) e (Cidade,VV);
- AR(3) para: (Floresta,HV) e (Floresta,VV);
- ARMA(1,1) para: (Cidade,HH);
- RB para: (Oceano,HV).

As estimativas de MVC e MVNC com seus respectivos erros padrão (EPs) são apresentados nas Tabelas 5.2 e 5.3, respectivamente. Tal como investigado no estudo de simulação, verifica-se um ganho da EMVNC sobre EMVC em relação a média das estimativas dos parâmetros entre as amostras de teste e treinamento.

Tabela 5.2: Estimação dos parâmetros por MVC em séries PolSAR

(Região, canal)	Modelo	n	$\hat{\phi}_1$	EP($\hat{\phi}_1$)	$\hat{\phi}_2$	EP($\hat{\phi}_2$)	$\hat{\phi}_3$	EP($\hat{\phi}_3$)	$\hat{\theta}_1$	EP($\hat{\theta}_1$)	$\hat{\sigma}^2$	EP($\hat{\sigma}^2$)
(Floresta, HH)	AR(1)	25	0.084	0.049	-	-	-	-	-	-	0.147	0.010
		35	0.112	0.035	-	-	-	-	-	-	0.151	0.007
		50	0.123	0.006	-	-	-	-	-	-	0.156	0.001
		861	0.172	0.034	-	-	-	-	-	-	0.163	-
(Oceano, HH)	AR(1)	25	0.078	0.039	-	-	-	-	-	-	0.072	0.004
		35	0.076	0.024	-	-	-	-	-	-	0.073	0.003
		50	0.089	0.018	-	-	-	-	-	-	0.075	0.002
		861	0.115	0.034	-	-	-	-	-	-	0.078	-
(Cidade,HH)	ARMA(1,1)	25	-0.078	0.075	-	-	-	-	0.387	0.088	1.474	0.107
		35	-0.0004	0.073	-	-	-	-	0.306	0.083	1.487	0.077
		50	-0.015	0.047	-	-	-	-	0.340	0.037	1.567	0.052
		861	0.166	0.083	-	-	-	-	0.212	0.080	1.700	-
(Floresta, HV)	AR(3)	25	0.346	0.048	-0.073	0.043	-0.0210	0.041	-	-	0.175	0.013
		35	0.356	0.033	-0.015	0.028	0.0007	0.028	-	-	0.184	0.010
		50	0.376	0.028	-0.036	0.017	0.013	0.019	-	-	0.190	0.0056
		861	0.460	0.040	0.062	0.027	0.103	0.034	-	-	0.215	-
(Oceano, HV)	RB	25	-	-	-	-	-	-	-	-	0.061	0.003
		35	-	-	-	-	-	-	-	-	0.0613	0.002
		50	-	-	-	-	-	-	-	-	0.062	0.002
		861	-	-	-	-	-	-	-	-	0.064	-
(Cidade, HV)	AR(1)	25	0.2135	0.0300	-	-	-	-	-	-	1.907	0.144
		35	0.249	0.024	-	-	-	-	-	-	1.936	0.098
		50	0.250	0.016	-	-	-	-	-	-	1.971	0.087
		861	0.328	0.032	-	-	-	-	-	-	2.072	-
(Floresta, VV)	AR(3)	25	0.185	0.041	-0.025	0.049	0.009	0.038	-	-	0.3009	0.022
		35	0.183	0.025	0.025	0.032	0.057	0.023	-	-	0.320	0.020
		50	0.199	0.012	0.036	0.024	0.038	0.022	-	-	0.326	0.011
		861	0.252	0.034	0.074	0.035	0.104	0.034	-	-	0.358	-
(Oceano, VV)	AR(1)	25	0.063	0.039	-	-	-	-	-	-	0.095	0.004
		35	0.061	0.028	-	-	-	-	-	-	0.096	0.003
		50	0.066	0.013	-	-	-	-	-	-	0.099	0.002
		861	0.090	0.034	-	-	-	-	-	-	0.103	-
(Cidade, VV)	AR(1)	25	0.221	0.040	-	-	-	-	-	-	1.377	0.117
		35	0.268	0.028	-	-	-	-	-	-	1.428	0.076
		50	0.271	0.020	-	-	-	-	-	-	1.436	0.053
		861	0.367	0.032	-	-	-	-	-	-	1.544	-

O teste de *Ljung-Box* (LB) também foi usado para avaliar suposições dos modelos ARMA; isto é, assegurar que os resíduos são ruídos branco. Resultados dos testes de LB e SW são dados na Tabela 5.4. Os p -valores do teste de *Ljung-Box* indicam que as autocorrelações dos resíduos são do tipo $\rho(h) = \mathbb{I}_0(h)$. Adicionalmente, os valores do teste de normalidade de *Shapiro-Wilker* indicam normalidade dos resíduos para as regiões de oceano e floresta.

Tabela 5.3: Estimação dos parâmetros por MVNC em séries PolSAR

(Região, canal)	Modelo	n	ϕ_1	EP(ϕ_1)	ϕ_2	EP(ϕ_2)	ϕ_3	EP(ϕ_3)	θ_1	EP(θ_1)	$\hat{\sigma}^2$	EP($\hat{\sigma}^2$)
(Floresta, HH)	AR(1)	25	0.1424	0.0485	-	-	-	-	-	-	0.1543	0.0099
		35	0.1553	0.0355	-	-	-	-	-	-	0.1559	0.0067
		50	0.1511	0.0246	-	-	-	-	-	-	0.1596	0.0050
		861	0.1715	0.0336	-	-	-	-	-	-	0.1633	-
(Oceano, HH)	AR(1)	25	0.1264	0.0359	-	-	-	-	-	-	0.0755	0.0038
		35	0.1153	0.0259	-	-	-	-	-	-	0.0760	0.0027
		50	0.1141	0.0164	-	-	-	-	-	-	0.0770	0.0017
		861	0.1148	0.0339	-	-	-	-	-	-	0.0783	-
(Cidade, HH)	ARMA(1,1)	25	0.1270	0.1032	-	-	-	-	0.2573	0.1900	1.3777	0.1281
		35	0.0627	0.0819	-	-	-	-	0.3666	0.0976	1.5184	0.0944
		50	0.0538	0.0560	-	-	-	-	0.3131	0.0452	1.6159	0.0526
		861	0.1648	-	-	-	-	-	0.2138	-	1.7000	-
(Floresta, HV)	AR(3)	25	0.4603	0.0419	-0.0039	0.0400	0.0993	0.0480	-	-	0.1922	0.0125
		35	0.4467	0.0305	0.0453	0.0292	0.0930	0.0297	-	-	0.1996	0.0100
		50	0.4588	0.0244	0.0155	0.0163	0.0988	0.0231	-	-	0.2042	0.0057
		861	0.4596	0.0339	0.0619	0.0273	0.1028	0.0338	-	-	0.2147	-
(Oceano, HV)	RB	25	-	-	-	-	-	-	-	-	0.0614	0.003
		35	-	-	-	-	-	-	-	-	0.0613	0.0022
		50	-	-	-	-	-	-	-	-	0.0623	0.0019
		861	-	-	-	-	-	-	-	-	0.0641	-
(Cidade, HV)	AR(1)	25	0.2853	0.0317	-	-	-	-	-	-	2.0119	0.1456
		35	0.2978	0.0252	-	-	-	-	-	-	2.0134	0.1005
		50	0.3014	0.0180	-	-	-	-	-	-	2.0459	0.0898
		861	0.3282	0.0322	-	-	-	-	-	-	2.0716	-
(Floresta, VV)	AR(3)	25	0.2566	0.0399	0.0339	0.0470	0.0834	0.0379	-	-	0.3201	0.0223
		35	0.2345	0.0233	0.0668	0.0302	0.1121	0.0233	-	-	0.3348	0.0196
		50	0.2422	0.0169	0.0713	0.0224	0.0801	0.0236	-	-	0.3391	0.0117
		861	0.2515	-	0.0739	-	0.1044	-	-	-	0.3581	-
(Oceano, VV)	AR(1)	25	0.0970	0.0337	-	-	-	-	-	-	0.0990	0.0046
		35	0.0938	0.0277	-	-	-	-	-	-	0.0990	0.0035
		50	0.0842	0.0113	-	-	-	-	-	-	0.1015	0.0024
		861	0.0896	-	-	-	-	-	-	-	0.1034	-
(Cidade, VV)	AR(1)	25	0.3047	0.0446	-	-	-	-	-	-	1.4567	0.1152
		35	0.3357	0.0336	-	-	-	-	-	-	1.4876	0.0723
		50	0.3341	0.0250	-	-	-	-	-	-	1.4940	0.0513
		861	0.3665	-	-	-	-	-	-	-	1.5443	-

Tabela 5.4: P -valores dos testes de *Shapiro-Wilker* e de *Ljung e Box* sobre os resíduos dos modelos ajustados aos dados PolSAR.

Região/Canal	Shapiro-Wilker	Ljung-Box				
		$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
(Floresta,HH)	0.0683	0.902	0.767	0.911	0.927	0.961
(Floresta,HV)	0.5158	0.834	0.894	0.541	0.670	0.491
(Floresta,VV)	0.5738	0.969	0.999	0.999	0.997	0.992
(Oceano,HH)	0.5738	0.953	0.957	0.179	0.296	0.394
(Oceano,HV)	0.9635	0.573	0.156	0.175	0.061	0.053
(Oceano,VV)	0.6877	0.921	0.624	0.421	0.543	0.681
(Cidade,HH)	0.1166	0.959	0.991	0.683	0.197	0.039
(Cidade,HV)	0.1036	0.709	0.665	0.729	0.6247	0.534
(Cidade,VV)	0.0115	0.837	0.757	0.861	0.492	0.027

¹ h : representa a defasagem (lag)

As Tabelas 5.5 e 5.6 apresentam, respectivamente, os tamanhos empíricos dos testes mediante os métodos MVC e MVNC (de acordo com o Passo 4 do Algoritmo 2) para as regiões selecionadas. Em geral, os testes apresentaram taxas de alarme falso pequenas em todas as regiões consideradas. Para a estimação MVNC, o tamanho empírico dos testes ($\hat{\alpha}$) foram reduzidos (relativamente aos resultados para MVC) nas regiões (Floresta, HH), (Floresta, HV), (Cidade, HV) e (Floresta, VV). Em relação aos testes sob TKL e JS, apresentou-se também uma redução na taxa de alarme falso com o uso da MVNC para a região (Cidade, VV). Portanto, os tamanhos empíricos podem ser melhorados via estimação não condicional.

Os testes baseados na distância *Kullback-Leibler* e na RV apresentam em geral as menores taxas $\hat{\alpha}$. Comparando resultados entre regiões, as menores taxas de alarme falso foram verificadas na região de oceano, com taxas muito próximas de zero mesmo em tamanhos de amostra pequenos.

Tabela 5.5: Tamanho dos testes de hipóteses para séries SAR assumindo estimação MVC

(Região, canal)	Modelo	n	$\hat{\alpha}(KL)$	$\hat{\alpha}(TKL)$	$\hat{\alpha}(JS)$	$\hat{\alpha}(B)$	$\hat{\alpha}(H)$	$\hat{\alpha}(RVC)$	$\hat{\alpha}(RVNC)$
(Floresta, HH)	$AR(1)$	25	0.118	0.206	0.147	0.118	0.029	0.118	0.088
		35	0.125	0.125	0.083	0.083	0.000	0.125	0.125
		50	0.118	0.059	0.059	0.059	0.059	0.118	0.118
(Oceano, HH)	$AR(1)$	25	0.029	0.059	0.059	0.059	0.000	0.000	0.000
		35	0.042	0.083	0.083	0.083	0.000	0.042	0.042
		50	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(Cidade, HH)	$ARMA(1, 1)$	25	0.088	0.235	0.118	0.118	0.000	0.176	0.088
		35	0.167	0.292	0.333	0.250	0.000	0.167	0.167
		50	0.000	0.118	0.059	0.059	0.000	0.059	0.059
(Floresta, HV)	$AR(3)$	25	0.147	0.353	0.324	0.294	0.000	0.088	0.088
		35	0.167	0.375	0.250	0.208	0.000	0.125	0.125
		50	0.176	0.294	0.294	0.235	0.000	0.176	0.176
(Oceano, HV)	RB	25	0.000	0.000	0.000	0.000	0.000	0.000	0.000
		35	0.000	0.000	0.000	0.042	0.000	0.000	0.000
		50	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(Cidade, HV)	$AR(1)$	25	0.029	0.265	0.235	0.206	0.000	0.059	0.088
		35	0.042	0.167	0.167	0.083	0.000	0.083	0.083
		50	0.176	0.176	0.118	0.059	0.000	0.118	0.176
(Floresta, VV)	$AR(3)$	25	0.059	0.294	0.176	0.088	0.000	0.147	0.147
		35	0.042	0.208	0.167	0.167	0.000	0.083	0.083
		50	0.000	0.176	0.059	0.000	0.000	0.059	0.059
(Oceano, VV)	$AR(1)$	25	0.059	0.059	0.059	0.029	0.000	0.000	0.000
		35	0.000	0.042	0.000	0.000	0.000	0.000	0.000
		50	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(Cidade, VV)	$AR(1)$	25	0.206	0.294	0.265	0.265	0.029	0.206	0.206
		35	0.125	0.250	0.208	0.208	0.042	0.125	0.125
		50	0.118	0.176	0.118	0.118	0.000	0.118	0.118

Os poderes empíricos dos testes (ou probabilidade de detecção) para os métodos de estimação MVC e MVNC são apresentados nas Tabelas 5.7 e 5.8, respectivamente.

Para o poder empírico quando $(\mathcal{H}_0, \mathcal{H}_1) = (\text{Floresta-HH}, \text{Oceano-HH})$, as medidas de distância com exceção de *Hellinger* apresentaram poder superior aos testes da Razão entre verossimilhanças em todos os tamanhos amostrais para EMVC. Em contra partida, com o uso da EMVNC os testes TKL, JS e B tiveram maiores poderes em relação aos testes RV e KL (em que $KL \equiv RVC \equiv RVNC$).

Em particular, em ambos os métodos de estimação para $n = 25$, os testes com base nas medidas de divergência Obtiver melhores *performance* nos cenários $(\mathcal{H}_1, \mathcal{H}_0) = (\text{Cidade-VV}, \text{Floresta-VV})$ ou $= (\text{Floresta-VV}, \text{Oceano-VV})$ ou $= (\text{Cidade-HH}, \text{Floresta-HH})$, De maneira que

Tabela 5.6: Tamanhos dos testes de hipóteses para séries SAR assumindo estimação MVNC.

(Região, canal)	Modelo	n	$\hat{\alpha}(KL)$	$\hat{\alpha}(TKL)$	$\hat{\alpha}(JS)$	$\hat{\alpha}(B)$	$\hat{\alpha}(H)$	$\hat{\alpha}(RVC)$	$\hat{\alpha}(RVNC)$
(Floresta,HH)	$AR(1)$	25	0.059	0.176	0.147	0.118	0.000	0.118	0.088
		35	0.125	0.125	0.083	0.083	0.000	0.125	0.125
		50	0.118	0.059	0.059	0.059	0.059	0.118	0.118
(Oceano,HH)	$AR(1)$	25	0.029	0.029	0.059	0.029	0.000	0.000	0.000
		35	0.042	0.083	0.083	0.083	0.000	0.042	0.042
		50	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(Cidade,HH)	$ARMA(1, 1)$	25	0.088	0.235	0.118	0.118	0.000	0.176	0.088
		35	0.167	0.292	0.333	0.250	0.000	0.167	0.167
		50	0.000	0.118	0.059	0.059	0.000	0.059	0.059
(Floresta,HV)	$AR(3)$	25	0.059	0.206	0.147	0.118	0.000	0.029	0.059
		35	0.000	0.167	0.083	0.042	0.000	0.000	0.000
		50	0.118	0.118	0.118	0.118	0.000	0.059	0.118
(Oceano,HV)	RB	25	0.000	0.000	0.000	0.000	0.000	0.000	0.000
		35	0.000	0.000	0.000	0.042	0.000	0.000	0.000
		50	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(Cidade,HV)	$AR(1)$	25	0.0597	0.176	0.118	0.059	0.000	0.059	0.059
		35	0.042	0.125	0.125	0.042	0.000	0.042	0.042
		50	0.118	0.118	0.118	0.118	0.000	0.059	0.118
(Floresta,VV)	$AR(3)$	25	0.059	0.147	0.059	0.059	0.000	0.059	0.059
		35	0.042	0.167	0.125	0.000	0.000	0.042	0.042
		50	0.000	0.059	0.059	0.000	0.000	0.000	0.000
(Oceano,VV)	$AR(1)$	25	0.000	0.000	0.000	0.000	0.000	0.000	0.000
		35	0.000	0.042	0.000	0.000	0.000	0.000	0.000
		50	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(Cidade,VV)	$AR(1)$	25	0.235	0.235	0.265	0.265	0.029	0.265	0.235
		35	0.167	0.208	0.208	0.208	0.042	0.208	0.167
		50	0.059	0.118	0.118	0.118	0.000	0.059	0.059

Tabela 5.7: Poder dos testes de hipóteses para séries SAR assumindo estimação MVC

[illegible]

Tabela 5.8: Poder dos testes de hipóteses para séries SAR assumindo estimação MVNC

[illegible]

$$\left\{ \begin{array}{l} \text{Para } (\mathcal{H}_1, \mathcal{H}_0) = (\text{Cidade-VV}, \text{Floresta-VV}) : \\ \pi(TKL) \approx \pi(B) \approx \pi(JS) > \pi(KL) \approx \pi(H) > \pi(RVNC) \approx \pi(RVC) \\ \text{Para } (\mathcal{H}_1, \mathcal{H}_0) = (\text{Floresta-VV}, \text{Oceano-VV}) : \\ \pi(TKL) \approx \pi(B) \approx \pi(JS) > \pi(KL) \approx \pi(RVNC) > \pi(RVC) > \pi(H) \\ \text{Para } (\mathcal{H}_1, \mathcal{H}_0) = (\text{Cidade-HH}, \text{Floresta-HH}) : \\ \pi(TKL) \approx \pi(B) \approx \pi(JS) \approx \pi(KL) \approx \pi(H) > \pi(RVC) > \pi(RVNC) \end{array} \right. \quad (5.3)$$

Nas demais regiões e/ou com tamanhos de amostras grandes os testes são aproximados.

Conclusivamente,

a proposta de classificadores em dados PolSAR via medidas de divergências em alguns casos pode se configurar como uma alternativa mais precisa àquelas com base na razão entre verossimilhanças (amplamente usado na literatura).

Além disso, esta aplicação assumiu uma estrutura de correlação espacial para classificação em dados PolSAR, até então não utilizada na literatura e uma comparação entre os métodos de estimação MVC e MVNC.

Como sugestão de trabalhos futuros, a fim de lidar com as evidências de (i) não normalidade dos resíduos e (ii) sazonalidade dos dados, requer-se a proposta de medidas de divergências para modelos gaussianos sazonais e modelos não gaussianos.

Capítulo 6

Considerações finais

Esta dissertação avança na proposta de métodos para modelos ARMA gaussianos. Especificamente,

- (i) um estudo comparativo entre três métodos de estimação clássicos em termos da influência de variações na região paramétrica de estabilidade sobre viés e EQM e sobre a densidade conjunta de processo ARMA gaussiano via medidas de divergência,
- (ii) proposta e discussão de cinco testes de hipótese baseados nas medidas de divergências para identificação em modelos ARMA e um estudo comparativo entre testes da razão entre verossimilhanças condicionais e não condicionais,
- (iii) proposta de novos critérios para seleção de modelos baseados nas distâncias *Kullback-Leibler*, Tipo *Kullback-Leibler*, *Jensen-Shannon*, *Bhattacharyya* e comparação com oito critérios clássicos (AIC, AICc, AICu, HQ, HQu, FEP, FEPu, SIC) e
- (iv) proposta dos testes de identificação propostos como classificadores em imagens PolSAR com *pixels* seguindo um processo ARMA espacial.

Quanto ao item (i), identificou-se que a estimação por MVNC é a que menos promove distorção a densidade conjunta de um processo autoregressivo gaussiano de ordem 1 quando $|\phi|$ aumenta para tamanhos pequenos e moderados de amostra. Para processos MA(1), identificou-se que todos os métodos tendem a produzir estimativas viesadas para os parâmetros θ e σ^2 quando $|\theta|$ cresce. Este último fenômeno indica que novos métodos de estimação e correções são requeridos. Para processos ARMA(1,1), as medidas de divergências *Kullback-Leibler*, *Jensen-Shannon* e *Bhattacharyya* apontam a estimação MVNC como a melhor para tamanhos pequenos e moderados na maioria dos pontos considerados e o método MQNC apresentou melhor desempenho quando $-0.7 < \phi < 0.7$ e $|\theta| \uparrow 1$ de acordo com a divergência Tipo *Kullback-Leibler*.

Com respeito ao item (ii), as evidências indicam que—para processos AR(1)—os testes com base nas estatísticas S_{KL} , S_{TKL} , S_{JS} e S_B combinados com a estimação MVNC se

mostraram bastante competitivos e, em alguns casos (principalmente quando os processos se aproximam de passeios aleatórios), significativamente melhores do que aqueles munidos por S_{RVC} e S_{RVNC} , que são altamente utilizados na literatura. Em processos MA(1), os testes em geral apresentaram um tamanho empírico acima do nível nominal para identificação. Dentre eles, o teste da RVNC combinado com a estimação de MVC apresentou o melhor resultado. Finalmente, para processos ARMA(1,1), os testes segundo as distâncias *Kullback-Leibler* e *Bhattacharyya* combinados com a estimação MVNC obtiveram os melhores desempenhos. Quanto à comparação entre testes das RVs, identificou-se uma diferença expressiva entre os testes RVNC e RVC em processos MA(1) e ARMA(1,1). Para os processos MA(1) e ARMA(1,1), quando o parâmetro $\theta \uparrow 1$ e os métodos de estimação MVNC e MQNC são considerados o teste RVNC tem maior poder empírico e menor tamanho empírico do teste, especificamente para processos próximos de um passeio aleatório. Para grandes amostras os testes são aproximados.

Quanto ao item (iii), para ambos os métodos de estimação, os novos critérios para seleção de modelos com tamanhos de amostras grandes apresentam taxas menores de *overfitting*. Para a estimação MVNC, as probabilidades de *overfitting* dos critérios de seleção diminuem (com maior velocidade nos novos critérios) para uma diminuição de $|\phi|$. A probabilidade de *overfitting* dos novos critérios diminui à medida que a ordem do modelo se afasta da verdadeira ordem. Além disso, a medida que aumentamos o tamanho das amostras essa probabilidade tende a decrescer.

Com respeito ao item (iv), vários classificadores foram planejados com base na combinação processo ARMA espacial + medidas de contraste (medidas de divergência vs. razão entre verossimilhanças). Nesta dissertação, assumimos uma estrutura de correlação espacial para classificação, até então não utilizada na literatura. Resultados têm mostrado evidências em favor do uso das medidas de distâncias, em particular *Kullback-Leibler* e *Bhattacharyya*, combinadas com o EMVNC.

Referências

AKAIKE, H. Fitting autoregressives models for prediction. In: *The Institute of Statistical Mathematics*. [S.l.: s.n.], 1969. v. 21.

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: *2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR*. [S.l.: s.n.], 1973. p. 267–281.

AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, v. 19, p. 716–723, 1974.

AKAIKE, H. A bayesian analysis of the minimum AIC procedure. *Annals of the Institute of Statistical Mathematics*, v. 30, p. 9–14, 1978.

ALI, S. M.; SILVEY, S. D. A general class of coefficients of divergence of one distribution from another. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 16, p. 131–142, 1996.

ANDERSON, T.; MENTZ, R. P. The generalized variance of a stationary autoregressive process. *Journal of Multivariate Analysis*, v. 7, n. 4, p. 584 – 588, 1977. ISSN 0047-259X.

BHANSALI, R.; DOWNHAM, D. Some properties of the order of an autoregressive model select by a generalization of akaike’s epf criterion. *Biometrika*, v. 64, p. 547–551, 1977.

BOX, G.; JENKINS, G. *Time series analysis forecasting and control*. [S.l.]: San Francisco: Holden- Day, 1976.

BROCKWELL, P.; DAVIS, R. *Introduction To Time Series And Forecasting*. Second edition. New York, NY 10013, USA: Springer Texts in Statistics, 2001.

BURBEA, J.; RAO, C. On the convexity of higher order Jensen differences based on entropy functions. *IEEE Transactions on Information Theory*, v. 28, n. 6, p. 961–963, 1982. ISSN 0018-9448.

BUSTOS, O.; OJEDA, S.; VALLEJOS, R. Spatial ARMA models and its applications to image filtering. *Brazilian Journal of Probability and Statistics*, Brazilian Statistical Association, v. 23, p. 141–165, 12 2009.

CHAMPERNOWNE, D. Sampling theory applied to autoregressive sequences. *Journal of the Royal Statistical Society B*, v. 10, p. 204–231, 1948.

- CORDEIRO, G. M. Introdução a teoria assintótica. In: *22º Colóquio Brasileiro de Matemática*. [S.l.: s.n.], 1999.
- CRAIG, F. A.; NEWBOLD, P. Finite sample properties of estimators for autoregressive moving average models. *Journal of Econometrics*, v. 13, p. 159–183, 1980. ISSN 0304-4076.
- CSISZÁR, I. Information type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, v. 2, p. 299–318, 1967.
- DOOB, J. L. *Stochastic Processes*. [S.l.]: London: Chapman & Hall., 1953.
- DUTRA, L. V. *Classificação de Texturas Usando Modelos ARMA e Distâncias da Função de Autocorrelação*. Tese (Doutorado) — Instituto Nacional de Pesquisas Espaciais, 1989.
- EGUCHI, S.; COPAS, J. Interpreting Kullback-Leibler divergence with the Neyman-Pearson lemma. *Journal of Multivariate Analysis*, v. 97, p. 2034–2040, 2006.
- FISHER, R. A. On a test of significance in pearson's biometrika tables. *Royal Statistical Society: Series B*, v. 18, p. 56–60, 1956.
- FRERY, A. C.; NASCIMENTO, A. D. C.; CINTRA, R. J. Analytic expressions for stochastic distances between relaxed complex Wishart distributions. *IEEE Transactions on Geoscience and Remote Sensing*, v. 52, p. 1213–1226, 2013.
- GALBRAITH, R. F.; GALBRAITH, J. I. On the inverse of some patterned matrices arising in the theory of stationary time series. *Journal of applied probability*, v. 11, p. 63–71, 1974.
- GIET, L.; LUBRANO, M. A minimum hellinger distance estimator for stochastic differential equations an application to statistical inference for continuous time interest rate models. *Computational Statistics & Data Analysis*, v. 56, p. 2945–2965, 2008.
- GOUDAIL, F.; RÉFRÉGIER, P.; DELYON, G. Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images. *Journal of the Optical Society of America A*, v. 21, p. 1231–1240, 2004.
- HAMILTON, J. D. *Time series analysis*. [S.l.]: New Jersey: Princeton University Press, 1994.
- HANNAN, E. J.; QUINN, B. The determination of order of autoregression. *Journal of the Royal Statistical Society*, v. 41, p. 190–195, 1979.
- HANNAN, E. J.; RISSANEN, J. Recursive estimation of mixed autoregressive-moving average order. *Biometrika*, v. 69, p. 81–94, 1982.
- HARTLEY, R. V. L. Transmission of information. *Bell System Technical Journal*, v. 7, p. 535–563, 1928.
- HURVICH, C. M.; TSAI, C. Regression and time series model selection in small samples. *Biometrika*, v. 76, p. 297–307, 1989.
- ISHI, N. et al. Computer classification of the EEG time series by Kullback information measure. *International Journal of Systems Science*, v. 6, p. 677–687, 1998.

- KAILATH, T. The divergence and bhattacharyya distance measures in signal selection. *IEEE Transactions on Communication Technology*, v. 15, p. 52–60, 1967.
- KANG, K. A comparison of the least-squares and maximum-likelihood estimation for moving-average processes. *Bureau of Census and Statistics, Canberra.*, 1975.
- KULLBACK, S. An information-theoretic derivation of certain limit relations for a stationary markov chain. *J. SIAM Control*, v. 4, p. 454–459, 1966.
- KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. *Annals of Mathematical Statistics*, v. 22, p. 79–86, 1951.
- KUPPERMAN, M. *Further Applications of Information Theory to Multivariate Analysis and Statistical Inference*. Tese (Doutorado), 1957.
- LAMPORT, L. A document preparation system. In: . [S.l.: s.n.], 1994.
- LEE, J. S.; POTTIER, E. *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton: CRC, 2009.
- LIESE, F.; VAJDA, I. On divergences and informations in statistics and information theory. *IEEE Transactions on Information Theory*, v. 52, p. 4394–4412, 2006.
- LILLIEFORS, H. On the kolmogorovsmirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, v. 62, p. 399–402, 1967.
- LINDLEY, D. V. On a measure of information provided by an experiment. *Annals of Mathematical Statistics*, v. 24, p. 986–1005, 1956.
- MAHALANOBIS, P. C. On the generalized distance in statistics. *National Institute of Science*, v. 12, p. 49–55, 1936.
- MAURICIO, J. A. Exact maximum likelihood estimation of stationary vector ARMA models. *Journal of the American Statistical Association*, v. 90, p. 282–291, 1995.
- MAURICIO, J. A. An algorithm for the exact likelihood of a stationary vector autoregressive-moving average model. *Journal of Time Series Analysis*, Blackwell Publishers, v. 23, p. 473–486, 2002. ISSN 1467-9892.
- MCLEOD, A.; ZHANG, Y. Faster ARMA maximum likelihood estimation. *Computational Statistics & Data Analysis*, v. 52, p. 2166 – 2176, 2008. ISSN 0167-9473.
- MCQUARRIE, A. D.; TSAI, C.-L. *Regression and Time Series Model Selection*. P O Box 128, Farrer Road, Singapore 912805: Word scientific, 1998.
- MENÉNDEZ, M. et al. Asymptotic behaviour and statistical applications of divergence measures in multinomial populations: a unified study. *Statistical Papers*, Springer Berlin / Heidelberg, v. 36, p. 1–29, 1995. ISSN 0932-5026.
- MENTZ, R. P. On the inverse of some covariance matrices. *Mathematical and Computer*, v. 23, p. 127–133, 1972.
- NASCIMENTO, A.; FRERY, A. C.; CINTRA, R. J. Entropy-based statistical analysis of PolSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 51, p. 3733–3743, 2012.

- NASCIMENTO, A. D. C. *Teoria Estatística da Informação para Dados de Radar de Abertura Sintética Univariados e Polarimétricos*. Tese (Doutorado) — Universidade Federal de Pernambuco, 2012.
- NASCIMENTO, A. D. C.; CINTRA, R. J.; FRERY, A. C. Hypothesis testing in speckled data with stochastic distances. *IEEE Transactions on Geoscience and Remote Sensing*, v. 48, p. 373–385, Jan. 2010.
- NASCIMENTO, A. D. C.; FRERY, A. C.; CINTRA, R. J. Bias correction and modified profile likelihood under the wishart complex distribution. *IEEE Transactions on Geoscience and Remote Sensing*, v. 52, p. 1–10, 2013.
- NASCIMENTO, A. D. C. et al. Comparing edge detection methods based on stochastic entropies and distances for PolSAR imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 7, p. 1–16, 2013.
- NYQUIST, H. Certain factors affecting telegraph speed. *Bell System Technical Journal*, v. 2, p. 324–345, 1924.
- NYQUIST, H. Certain topics in telegraph transmission theory. *Transaction of the American Institute of Electrical Engineers*, v. 27, p. 617–644, 1928.
- POLLOCK, D. S. G. *A Handbook of Time-Series Analysis, Signal Processing and Dynamics*. [S.l.]: Signal Processing and its Applications, 1999.
- PRIESTLEY, M. B. Spectral analysis and time series. In: *Academic Press, New York*. [S.l.: s.n.], 1981.
- RENCHE, A. C.; SCHAALE, G. B. *Linear Models in Statistics*. [S.l.]: John Wiley & Sons, 2008.
- RÉNYI, A. On measures of entropy and information. In: *4th Berkeley Symposium on Mathematical Statistics and Probability*. [S.l.: s.n.], 1961. v. 1, p. 547–561.
- SALICRÚ, M. et al. On the applications of divergence type measures in testing statistical hypotheses. *Journal of Multivariate Analysis*, v. 51, p. 372–391, 1994.
- SCHWARZ, G. Estimating the dimensional of a model. *Annals of Statistics, Hayward*, p. 461–464, 1978.
- SEBER, G. A. F. *A Matrix Handbook for Statisticians*. [S.l.]: Hoboken, New Jersey, 2008.
- SEGHOUANE, A.; AMARI, S. The AIC criterion and symmetrizing the Kullback-Leibler divergence. *IEEE Transactions on Neural Networks*, v. 18, p. 971–976, 2007.
- SEGHOUANE, A.-K.; BEKARA, M. A small sample model selection criterion based on Kullbacks symmetric divergence. *IEEE Transactions on Signal Processing*, v. 52, p. 3314–3323, 2004.
- SHAMAN, P. On the inverse of the covariance matrix of a first-order moving average. *Biometrika*, v. 56, p. 595–600, 1969.
- SHAMAN, P. On the inverse in the covariance matrix of a first-order autoregressive-moving average process. *Biometrika*, v. 60, p. 193–196, 1973.

- SHANNON, C. E. A mathematical theory of communication. *The Bell, System Technical Journal*, v. 27, p. 379–423, 1948.
- SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). *Biometrika*, v. 52, p. 591–611, 1965.
- SHIBATA, R. Asymptotic efficient selection of the order of the model for estimating parameters of a linear process. *Annals of Statistics*, v. 8, p. 147–164, 1980.
- SHUMWAY, R. H.; STOFFER, D. S. *Time Series Analysis and Its Applications*. 233 Spring Street, New York, NY 10013, USA: Springer Texts in Statistics, 2005.
- SIDDIQUI, M. M. On the average of the sample covariance matrix in a stationary autoregressive process. *Annals of Mathematical Statistics*, v. 29, p. 585–588, 1958.
- SUGIURA, N. Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, v. 7, p. 1326, 1978.
- TIAO, G.; ALI, M. Analysis of correlated random effects: linear model with two random components. *Biometrika*, v. 58, p. 37–51, 1971.
- UPPULURI, V. R. R.; CARPENTER, J. A. The inverse of a matrix occurring in first-order moving average models. *Sankhyā A*, v. 31, p. 79–82, 1969.
- VERDÚ, S. Fifty years of Shannon theory. *IEEE Transactions Information Theory, Special Commemorative Issue*, v. 44, p. 2057–2078, 1998.
- WIENER, N. The theory of prediction. In: *In: Beckenbach, E. (Ed.), Modern Mathematics for Engineers*. McGraw-Hill, New York. [S.l.: s.n.], 1956.
- YAO, Q.; BROCKWELL, P. J. Gaussian maximum likelihood estimation for ARMA models II: Spatial processes. *Bernoulli*, Bernoulli Society for Mathematical Statistics and Probability, v. 12, p. 403–429, 06 2006.
- YU, J. Empirical characteristic function estimation and its applications. *Econometric Reviews*, v. 23, p. 93–123, 2004.

Apêndice A

Derivadas da Δ_{KL} para o modelo ARMA

Sejam $\mathbf{x}(\boldsymbol{\beta})$ e $\mathbf{y}(\boldsymbol{\beta}^*)$ vetores aleatórios n -dimensionais tais que seus t -ésimos elementos são $X_t(\boldsymbol{\beta}) \sim ARMA(p, q)$: $\boldsymbol{\beta} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2)^\top$ e $Y_t(\boldsymbol{\beta}^*) \sim ARMA(p, q)$: $\boldsymbol{\beta}^* = (\phi_1^*, \dots, \phi_p^*, \theta_1^*, \dots, \theta_q^*, \sigma_*^2)^\top$, respectivamente. Sob a suposição gaussiana,

$$\Delta_{KL}(\mathbf{x}||\mathbf{y}) = \frac{1}{2} [\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)]^\top \boldsymbol{\Omega}^{-1}(\boldsymbol{\beta}^*) [\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)] + \frac{1}{2} \left[\text{tr}(\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta}^*)\boldsymbol{\Omega}(\boldsymbol{\beta})) - n + \log \left(\frac{|\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta}^*)|}{|\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta})|} \right) \right].$$

A derivação desta medida, reduz-se a derivar as quantidades

- (a) $\text{tr}(\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta}^*)\boldsymbol{\Omega}(\boldsymbol{\beta}))$;
- (b) $\log |\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta}^*)|$ e $\log |\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta})|$, e
- (c) $[\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)]^\top \boldsymbol{\Omega}^{-1}(\boldsymbol{\beta}^*) [\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)]$.

Seja $A(\boldsymbol{\beta}) \triangleq \text{tr}[\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta}^*)\boldsymbol{\Omega}(\boldsymbol{\beta})]$, De (SEBER, 2008, p. 55), tem-se que

$$A(\boldsymbol{\beta}) = \sum_{i=1}^n \sum_{j=1}^n \boldsymbol{\Omega}^{ij}(\boldsymbol{\beta}^*) \boldsymbol{\Omega}_{ij}(\boldsymbol{\beta})$$

em que $\boldsymbol{\Omega}^{ij}(\boldsymbol{\beta}^*)$ é a entrada (i, j) do inverso da matriz de covariância de Y_t e $\boldsymbol{\Omega}_{ij}(\boldsymbol{\beta})$ é a entrada (i, j) da matriz de covariância de X_t . Em particular,

(a.1) Para $\{X_t\} \sim AR(1)$:

$$A(\phi_1, \sigma^2) = \left[\frac{\sigma^2}{1 - \phi_1^2} \right] \sum_{i=1}^n \sum_{j=1}^n \phi_1^{|i-j|} \boldsymbol{\Omega}_*^{ij} = \left[\frac{\sigma^2}{1 - \phi_1^2} \right] \left(\sum_{i=1}^n \boldsymbol{\Omega}_*^{ii} + 2 \sum_{j=1}^n \phi_1^{|i-j|} \boldsymbol{\Omega}_*^{ij} \right).$$

(a.2) Para $\{X_t\} \sim AR(p)$:

$$A(\phi_1, \dots, \phi_p, \sigma^2) = \left[\frac{\sigma^2}{1 - \sum_{i=1}^n \phi_i \rho_X(i)} \right] \left\{ \sum_{i=1}^n \Omega_*^{ii} + 2 \sum_{1 \leq i < j \leq n} K(i, j) \Omega_*^{ij} \right\},$$

em que $K(i, j)$ é a entrada (i, j) matriz de correlação de X_t .

(a.3) Para $\{X_t\} \sim MA(1)$:

$$A(\theta_1, \sigma^2) = \sigma^2(1 + \theta_1^2) \left\{ \sum_{i=1}^n \Omega_*^{ii}(\theta_1^*, \sigma_*^2) + 2 \frac{\theta_1}{1 + \theta_1^2} \sum_{i=1}^{n-1} \Omega_*^{i(i+1)} \right\}.$$

(a.4) Para $\{X_t\} \sim ARMA(1, 1)$:

$$A(\phi_1, \theta_1, \sigma^2) = \frac{\sigma^2(1 + 2\theta_1\phi_1 + \theta_1^2)}{1 - \phi_1^2} \left\{ \sum_{i=1}^n \Omega_*^{ii} + 2\rho_X(1) \sum_{1 \leq i < j \leq n} \phi_1^{|i-j|-1} \Omega_*^{ij} \right\},$$

$$\text{em que } \rho_X(1) = \left[\theta_1 + \phi_1 + \frac{\phi_1(\theta_1 + \phi_1)^2}{1 - \phi_1^2} \right] / \left[1 + \frac{(\phi_1 + \theta_1)^2}{1 - \phi_1^2} \right].$$

Defina $B(\beta) \triangleq \log |\mathbf{\Omega}^{-1}(\beta)|$. De (GALBRAITH; GALBRAITH, 1974), seguem-se os resultados:

(b.1) Para $\{X_t\} \sim AR(1)$:

Neste caso, $|\mathbf{\Omega}^{-1}(\phi_1, \sigma^2)| = |\sigma^{-2} \mathbf{V}^{-1}(\phi_1)| = \sigma^{-n/2} (1 - \phi_1^2)$. Portanto,

$$B(\phi_1) = \log(1 - \phi_1^2) - n \log(\sigma^2).$$

(b.2) Para $\{X_t\} \sim AR(p)$:

Neste caso, $|V_p^{-1}(\phi_1, \dots, \phi_p, \sigma^2)| = |\{V_p^{ij}\}_{i,j=1}^p|$. Então,

$$B(\phi_1, \dots, \phi_p, \sigma^2) = \log |\{V_p^{ij}\}_{i,j=1}^p| - n \log(\sigma^2).$$

Em particular para $AR(2)$, $\mathbf{V}_2^{-1} = (1 + \phi_2) \begin{bmatrix} (1 - \phi_2) & -\phi_1 \\ -\phi_1 & (1 - \phi_2) \end{bmatrix}$. Então,

$$B(\phi_1, \phi_2, \sigma^2) = \log |\sigma^{-2} \mathbf{V}_2^{-1}| = \log \{(1 + \phi_2)[(1 - \phi_2)^2 - \phi_1^2]\} - n \log(\sigma^2).$$

(b.3) Para $\{X_t\} \sim MA(1)$: De (GALBRAITH; GALBRAITH, 1974, p.67), $|\mathbf{\Omega}^{-1}(\theta_1)| = \frac{1-\theta_1^2}{1-\theta_1^{2(n+1)}} \sigma^{\frac{-n}{2}}$ e, portanto,

$$B(\theta_1, \sigma^2) = \log(1 - \theta_1^2) - \log(1 - \theta_1^{2(n+1)}) - n \log(\sigma^2).$$

(b.4) Para $\{X_t\} \sim ARMA(1, 1)$: De (GALBRAITH; GALBRAITH, 1974, p.68), $|\mathbf{\Omega}^{-1}(\phi_1, \theta_1, \sigma^2)| = \frac{(1-\phi_1^2)(1-\theta_1^2)}{(1-\phi_1\theta_1)^2 - (\theta_1 - \phi_1)^2 \theta_1^{2n}} \sigma^{\frac{-n}{2}}$ e, portanto,

$$\begin{aligned} B(\phi_1, \theta_1, \sigma^2) &= \log[(1 - \phi_1^2)(1 - \theta_1^2)] - \log[(1 - \phi_1\theta_1)^2 - (\theta_1 - \phi_1)^2 \theta_1^{2n}] - n \log(\sigma^2) \\ &= \log(1 - \phi_1^2) + \log(1 - \theta_1^2) - \log(1 - \phi_1\theta_1)^2 + \log[(\theta_1 - \phi_1)^2 \theta_1^{2n}] - n \log(\sigma^2) \end{aligned}$$

Finalmente, seja

$$C(\beta) \triangleq [\boldsymbol{\mu}(\beta) - \boldsymbol{\mu}(\beta^*)]^\top \mathbf{\Omega}^{-1}(\beta^*) [\boldsymbol{\mu}(\beta) - \boldsymbol{\mu}(\beta^*)],$$

em que $\boldsymbol{\mu}(\beta) = [\cdots |\hat{X}_t(\beta)| \cdots]^\top$ e $\boldsymbol{\mu}(\beta^*) = [\cdots |X_t| \cdots]^\top$ e $\hat{X}_t(\beta)$ é o melhor preditor linear. De (SEBER, 2008, p. 55) e por aplicar propriedades do traço e da inversa da matriz,

$$\begin{aligned} C(\beta) &= \text{tr} \left\{ \mathbf{\Omega}^{-1}(\beta^*) [\boldsymbol{\mu}(\beta) - \boldsymbol{\mu}(\beta^*)] [\boldsymbol{\mu}(\beta) - \boldsymbol{\mu}(\beta^*)]^\top \right\} \\ &= \sum_{i=1}^n \sum_{j=1}^n \mathbf{\Omega}_*^{ij} [x_i - \hat{x}_i(\beta)] [x_j - \hat{x}_j(\beta)] \\ &= \sum_{i=1}^n \mathbf{\Omega}_*^{ii} [x_i - \hat{x}_i(\beta)]^2 + 2 \sum_{1 \leq i < j \leq n} \mathbf{\Omega}_*^{ij} [x_i - \hat{x}_i(\beta)] [x_j - \hat{x}_j(\beta)]. \end{aligned}$$

Particularmente, tem-se

(c.1) Para $\{X_t\} \sim AR(1)$:

$$\begin{aligned} C(\phi_1) &= \mathbf{\Omega}_*^{11} x_1^2 + \sum_{i=2}^n \mathbf{\Omega}_*^{ii} (x_i - \phi_1 x_{i-1})^2 \\ &\quad + 2 \sum_{1 \leq i < j \leq n} \mathbf{\Omega}_*^{ij} (x_i - \phi_1 x_{i-1})(x_j - \phi_1 x_{j-1}), \quad \text{em que } x_0 = 0. \end{aligned}$$

(c.2) Para $\{X_t\} \sim AR(p)$:

$$C(\phi_1, \dots, \phi_p) = \sum_{t=1}^p \Omega_*^{tt} x_t^2 + \sum_{i=p+1}^n \Omega_*^{ii} (x_i - \phi_1 x_{i-1} - \phi_2 x_{i-2} - \dots - \phi_p x_{i-p})^2 \\ + 2 \sum_{1 \leq i < j \leq n} \Omega_*^{ij} (x_i - \phi_1 x_{i-1} - \dots - \phi_p x_{i-p})(x_j - \phi_1 x_{j-1} - \dots - \phi_p x_{j-p}),$$

em que $x_{-p+1} = x_{-p+2} = \dots = x_0 = 0$.

(c.3) Para $\{X_t\} \sim MA(1)$: De (GALBRAITH; GALBRAITH, 1974, p. 67),

$$C(\theta_1) = \sum_{i=1}^n \Omega_*^{ii} [x_i - \hat{x}_i(\theta_1)]^2 + 2 \sum_{1 \leq i < j \leq n} \Omega_*^{ij} [x_i - \hat{x}_i(\theta_1)][x_j - \hat{x}_j(\theta_1)]$$

em que $\Omega_*^{ij}(\theta_1) = \frac{\theta_1^{*(j-1)} [1 - \theta_1^{*2i}] [1 - \theta_1^{*2(n+1-j)}]}{[1 - \theta_1^{*2}] [1 - \theta_1^{*2(n+1)}]}$ e $\hat{X}_{i+1}(\theta_1) = \theta_1 [X_i^* - \hat{X}_i(\theta_1)]/v_{i-1}$

com $v_0 = 1 + \theta_1^2$ e $v_{i+1} = 1 + \theta_1^2 - (\theta_1^2/v_i)$.

(c.4) Para $\{X_t\} \sim ARMA(1,1)$: De (GALBRAITH; GALBRAITH, 1974, p. 68) e (BROCKWELL; DAVIS, 2001),

$$C(\phi_1, \theta_1) = \sum_{i=1}^n \Omega_*^{ii} [x_i - \hat{x}_i(\phi_1, \theta_1)]^2 + 2 \sum_{1 \leq i < j \leq n} \Omega_*^{ij} [x_i - \hat{x}_i(\phi_1, \theta_1)][x_j - \hat{x}_j(\phi_1, \theta_1)] \\ = \sum_{i=1}^n m_*^{ii} [x_i - \hat{x}_i(\phi_1, \theta_1)]^2 + 2 \sum_{1 \leq i < j \leq n} m_*^{ij} [x_i - \hat{x}_i(\phi_1, \theta_1)][x_j - \hat{x}_j(\phi_1, \theta_1)],$$

em que m_*^{ij} definido nas Equações (2.9)-(2.10) é a entrada (i, j) da matriz de covariância inversa de $\{X_t\}$, $\hat{X}_{i+1}(\beta) = \phi_1 X_i^* + \theta_{i,1} [X_i^* - \hat{X}_i(\beta)]$ com $v_0 = (1 + 2\phi_1\theta_1 + \theta_1^2)/(1 - \phi_1^2)$, $\theta_{i,1} = \theta_1/v_{i-1}$ e $v_i = 1 + \theta_1^2 - (\theta_1^2/v_{i-1})$.

Apêndice B

Derivadas da Δ_B para o modelo ARMA

Sejam $\mathbf{x}(\boldsymbol{\beta})$ e $\mathbf{y}(\boldsymbol{\beta}^*)$ vetores aleatórios n -dimensionais tais que seus t -ésimos elementos são $X_t(\boldsymbol{\beta}) \sim ARMA(p, q)$: $\boldsymbol{\beta} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2)^\top$ e $Y_t(\boldsymbol{\beta}^*) \sim ARMA(p, q)$: $\boldsymbol{\beta}^* = (\phi_1^*, \dots, \phi_p^*, \theta_1^*, \dots, \theta_q^*, \sigma_*^2)^\top$, respectivamente. Sob a suposição gaussiana,

$$\Delta_B(\mathbf{x}||\mathbf{y}) = \frac{1}{4} [\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)]^\top \left[\frac{\boldsymbol{\Omega}(\boldsymbol{\beta}^*) + \boldsymbol{\Omega}(\boldsymbol{\beta})}{2} \right]^{-1} [\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)] + \\ - \frac{1}{4} [\log |\boldsymbol{\Omega}(\boldsymbol{\beta}^*)| + \log |\boldsymbol{\Omega}(\boldsymbol{\beta})|] + \frac{1}{2} \log |\boldsymbol{\Omega}(\boldsymbol{\beta}^*) + \boldsymbol{\Omega}(\boldsymbol{\beta})|,$$

em que $\boldsymbol{\mu}(\boldsymbol{\beta}) = [\dots | \hat{X}_t(\boldsymbol{\beta}) | \dots]^\top$, $\boldsymbol{\mu}(\boldsymbol{\beta}^*) = [\dots | X_t | \dots]^\top$ e $\hat{X}_t(\boldsymbol{\beta})$ é o melhor preditor linear. A derivação desta medida, reduz-se a computar as quantidades

$$(a) [\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)]^\top \left[\frac{\boldsymbol{\Omega}(\boldsymbol{\beta}^*) + \boldsymbol{\Omega}(\boldsymbol{\beta})}{2} \right]^{-1} [\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)],$$

$$(b) \log |\boldsymbol{\Omega}(\boldsymbol{\beta}^*)|, \log |\boldsymbol{\Omega}(\boldsymbol{\beta})| \text{ e}$$

$$(c) \log |\boldsymbol{\Omega}(\boldsymbol{\beta}^*) + \boldsymbol{\Omega}(\boldsymbol{\beta})|.$$

Primeiramente, considere as seguintes decomposições de *Cholesky* para as matrizes de covariância, $\boldsymbol{\Omega}(\boldsymbol{\beta}^*)^{-1} = \mathbf{A}_1(\boldsymbol{\beta}^*)\mathbf{A}_1(\boldsymbol{\beta}^*)^\top$ e $\boldsymbol{\Omega}(\boldsymbol{\beta})^{-1} = \mathbf{B}_1(\boldsymbol{\beta})\mathbf{B}_1(\boldsymbol{\beta})^\top$. Adicionalmente, de (RENCER; SCHAALJE, 2008, p. 24), sejam $\mathbf{A}$, $\mathbf{B}$ e $(\mathbf{BCD})$ matrizes não singulares, vale-se o seguinte resultado

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}.$$

Seja $A(\boldsymbol{\beta}) \triangleq [\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)]^\top \left[\frac{\boldsymbol{\Omega}(\boldsymbol{\beta}^*) + \boldsymbol{\Omega}(\boldsymbol{\beta})}{2} \right]^{-1} [\boldsymbol{\mu}(\boldsymbol{\beta}) - \boldsymbol{\mu}(\boldsymbol{\beta}^*)]$. Da penúltima identidade matricial, tem-se

$$\begin{aligned}
[\Omega(\beta^*) + \Omega(\beta)]^{-1} &= [\Omega(\beta) + (\mathbf{A}_1(\beta^*)\mathbf{A}_1(\beta^*)^\top)^{-1}]^{-1} = \mathbf{B}_1(\beta)\mathbf{B}_1(\beta)^\top \\
&\quad - \mathbf{B}_1(\beta)\mathbf{B}_1(\beta)^\top \mathbf{A}_1(\beta^*)^{-\top} [\mathbb{I} + \mathbf{A}_1(\beta^*)^{-1} \\
&\quad \mathbf{B}_1(\beta)\mathbf{B}_1(\beta)^\top \mathbf{A}_1(\beta^*)^{-\top}]^{-1} \mathbf{A}_1(\beta^*)^{-1} \mathbf{B}_1(\beta)\mathbf{B}_1(\beta)^\top
\end{aligned}$$

Tome $\mathbf{H} \triangleq \mathbf{A}_1(\beta^*)^{-1}\mathbf{B}_1(\beta)$, então a equação acima pode ser reescrita como

$$\begin{aligned}
[\Omega(\beta^*) + \Omega(\beta)]^{-1} &= \mathbf{B}_1(\beta)\mathbf{B}_1(\beta)^\top - \mathbf{B}_1(\beta)\mathbf{H}^\top [\mathbb{I} + \mathbf{H}\mathbf{H}^\top]^{-1} \mathbf{H}\mathbf{B}_1(\beta)^\top \\
&= \mathbf{B}_1(\beta) \{ \mathbb{I} - \mathbf{H}^\top [\mathbb{I} + \mathbf{H}\mathbf{H}^\top]^{-1} \mathbf{H} \} \mathbf{B}_1(\beta)^\top.
\end{aligned}$$

Seja $\boldsymbol{\mu}_{11} = [\boldsymbol{\mu}(\beta) - \boldsymbol{\mu}(\beta^*)]$. De (POLLOCK, 1999, p. 682) e por distributividade,

$$\begin{aligned}
[\boldsymbol{\mu}(\beta) - \boldsymbol{\mu}(\beta^*)]^\top [\Omega(\beta^*) + \Omega(\beta)]^{-1} [\boldsymbol{\mu}(\beta) - \boldsymbol{\mu}(\beta^*)] &= \boldsymbol{\mu}_{11}^\top [\Omega(\beta^*) + \Omega(\beta)]^{-1} \boldsymbol{\mu}_{11} \\
&= \boldsymbol{\mu}_{11}^\top \mathbf{B}_1(\beta) \{ \mathbb{I} - \mathbf{H}^\top [\mathbb{I} + \mathbf{H}\mathbf{H}^\top]^{-1} \mathbf{H} \} \mathbf{B}_1(\beta)^\top \boldsymbol{\mu}_{11} \\
&= \{ \mathbf{B}_1(\beta)^\top \boldsymbol{\mu}_{11} - \mathbf{H}^\top [\mathbb{I} + \mathbf{H}\mathbf{H}^\top]^{-1} \mathbf{H}\mathbf{B}_1(\beta)^\top \boldsymbol{\mu}_{11} \}^\top \\
&\quad \{ \mathbf{B}_1(\beta)^\top \boldsymbol{\mu}_{11} - \mathbf{H}^\top [\mathbb{I} + \mathbf{H}\mathbf{H}^\top]^{-1} \mathbf{H}\mathbf{B}_1(\beta)^\top \boldsymbol{\mu}_{11} \} \\
&= \{ \mathbf{B}_1(\beta)^\top \boldsymbol{\mu}_{11} - \mathbf{B}_{11}(\beta, \beta^*) \boldsymbol{\mu}_{11} \}^\top \{ \mathbf{B}_1(\beta)^\top \boldsymbol{\mu}_{11} - \mathbf{B}_{11}(\beta, \beta^*) \boldsymbol{\mu}_{11} \},
\end{aligned}$$

em que $\mathbf{B}_{11}(\beta, \beta^*) = \mathbf{H}^\top [\mathbb{I} + \mathbf{H}\mathbf{H}^\top]^{-1} \mathbf{H}\mathbf{B}_1(\beta)^\top$. Então,

$$A(\beta, \beta^*) \triangleq \{ \mathbf{B}_1(\beta)^\top \boldsymbol{\mu}_{11} - \mathbf{B}_{11}(\beta, \beta^*) \boldsymbol{\mu}_{11} \}^\top \{ \mathbf{B}_1(\beta)^\top \boldsymbol{\mu}_{11} - \mathbf{B}_{11}(\beta, \beta^*) \boldsymbol{\mu}_{11} \} \quad (\text{B.1})$$

Particularmente,

(a.1) Para $\{X_t\} \sim AR(1)$:

$$\begin{aligned}
A(\phi_1, \phi_1^*, \sigma^2, \sigma_*^2) &= \{ \mathbf{B}_1(\phi_1, \sigma^2)^\top \boldsymbol{\mu}_{11} - \mathbf{B}_{11}(\phi_1, \phi_1^*, \sigma^2, \sigma_*^2) \boldsymbol{\mu}_{11} \}^\top \\
&\quad \times \{ \mathbf{B}_1(\phi_1, \sigma^2)^\top \boldsymbol{\mu}_{11} - \mathbf{B}_{11}(\phi_1, \phi_1^*, \sigma^2, \sigma_*^2) \boldsymbol{\mu}_{11} \}
\end{aligned}$$

representa o resultado de (B.1) para $\{X_t\} \sim AR(1)$. O mesmo procedimento é considerado para $\{X_t\} \sim AR(p)$, $\{X_t\} \sim MA(1)$ e $\{X_t\} \sim ARMA(1, 1)$.

Definindo, $B(\beta) \triangleq \log |\Omega(\beta^*)|$, então tem-se

(b.1) Para $\{X_t\} \sim AR(1)$: De (HAMILTON, 1994, Cap. 5),

$$B(\phi_1, \sigma^2) = -\log |\mathbf{\Omega}^{-1}(\phi_1)| = \log |\sigma^{-2}(1 - \phi_1^2)| = -\log(1 - \phi_1^2) + n \log \sigma^2$$

ou, De (ANDERSON; MENTZ, 1977),

$$B(\phi_1, \sigma^2) = \log |\mathbf{\Omega}(\phi_1)| = n \log(\sigma^2) + \log[1 - w_1(\phi_1)w_1(\phi_1)]^{-1}.$$

em que $w_1(\phi_1)$ é solução de $w - \phi_1 = 0$.

(b.2) Para $\{X_t\} \sim AR(p)$: De (HAMILTON, 1994, Cap. 5),

$$B(\phi_1, \dots, \phi_p, \sigma^2) = \log |\mathbf{\Omega}(\phi_1, \dots, \phi_p, \sigma^2)| = n \log \sigma^2 - |V_{ij}^{-1}|.$$

ou, De (ANDERSON; MENTZ, 1977),

$$\begin{aligned} B(\phi_1, \dots, \phi_p, \sigma^2) &= \log |\mathbf{\Omega}(\phi_1, \dots, \phi_p, \sigma^2)| \\ &= n \log(\sigma^2) + \sum_{i,j=1}^p \log[1 - w_i(\phi_1, \dots, \phi_p, \sigma^2)w_j(\phi_1, \dots, \phi_p, \sigma^2)]^{-1} \end{aligned}$$

em que $w_1(\phi_1, \dots, \phi_p), \dots, w_p(\phi_1, \dots, \phi_p)$ são raízes de $\sum_{i=1}^p \phi_i w^{p-i} = 0$.

(b.3) Para $\{X_t\} \sim MA(1)$:

$$B(\theta_1, \sigma^2) = -\log |\mathbf{\Omega}^{-1}(\theta_1, \sigma^2)| = \log(1 - \theta_1^2) - \log(1 - \theta_1^{2(n+1)}) + n \log(\sigma^2).$$

(b.4) Para $\{X_t\} \sim ARMA(1, 1)$:

$$\begin{aligned} B(\phi_1, \theta_1, \sigma^2) &= -\log |\mathbf{\Omega}^{-1}(\phi_1, \theta_1, \sigma^2)| \\ &= n \log \sigma^2 + \log(1 - \phi_1^2) + \log(1 - \theta_1^2) - \log(1 - \phi_1 \theta_1)^2 + \log[(\theta_1 - \phi_1)^2 \theta_1^{2n}]. \end{aligned}$$

Finalmente, seja $C(\boldsymbol{\beta}, \boldsymbol{\beta}^*) \triangleq \log |\mathbf{\Omega}(\boldsymbol{\beta}^*) + \mathbf{\Omega}(\boldsymbol{\beta})|$, De $|AB| = |A||B|$ e $|A^{-1}| = 1/|A|$,

$$|[\mathbf{\Omega}(\boldsymbol{\beta}^*) + \mathbf{\Omega}(\boldsymbol{\beta})]^{-1}| = |\mathbf{\Omega}^{-1}(\boldsymbol{\beta})| |\mathbb{I} + \mathbf{\Omega}(\boldsymbol{\beta}^*)\mathbf{\Omega}^{-1}(\boldsymbol{\beta})|^{-1}.$$

Assim,

$$\begin{aligned} C(\boldsymbol{\beta}, \boldsymbol{\beta}^*) &= -\log |[\boldsymbol{\Omega}(\boldsymbol{\beta}^*) + \boldsymbol{\Omega}(\boldsymbol{\beta})]^{-1}| = -\log |\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta})| + \log |\mathbb{I} + \boldsymbol{\Omega}(\boldsymbol{\beta}^*)\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta})| \\ &= -\log |\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta})| + \sum_{i=1}^n \log \lambda_{12}^{(i)}, \end{aligned}$$

em que $\lambda_{12}^{(1)}, \dots, \lambda_{12}^{(n)}$ são os autovalores de $\mathbb{I} + \boldsymbol{\Omega}(\boldsymbol{\beta}^*)\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta})$. Deste último resultado, Δ_B pode ser determinada para os modelos $\{X_t\} \sim AR(1)$, $\{X_t\} \sim AR(p)$, $\{X_t\} \sim MA(1)$ e $\{X_t\} \sim ARMA(1, 1)$.